To our patients
who inspire us to continually strive to improve their health,
function, and quality of life

To our teachers
who encouraged us to develop a scientific approach to medicine
and instilled in us the need for continuous learning

To our students
who challenge and stimulate us to stay at the cutting edge;
they are our hope for the future

To our colleagues,
who have gone before us, who are with us
and who will follow us

To our families
who provided the support and patience necessary
## PART I
### Principles of Assessment and Evaluation

1. Clinical Evaluation  
   Ibrahim A. Aksoy, John A. Freeman, Kirsten S. Paynter, Bryan K. Ganter, Rolland P. Erickson, Matthew A. Butters, and Stephen F. Noll  

2. The Physical Examination  
   Joseph H. Feinberg and Peter J. Moley  

3. Assessment of Human Muscle Function  
   Walter R. Frontera and Jan Lexell  

4. Electrodiagnostic Evaluation of the Peripheral Nervous System  
   Lawrence R. Robinson  

5. Human Walking  
   William S. Pease and Brian L. Bowyer  

6. Imaging Techniques Relative to Rehabilitation  
   Edgar Colón, Eduardo Labat, Gory Ballester, Jorge Vidal, and Angel Gomez  

7. Diagnostic Ultrasound  
   Henry L. Lew, Tyng-Guey Wang, Wen-Chung Tsai, and Yi-Pin Chiang  

8. Principles and Applications of Measurement Methods  
   Steven R. Hinderer and Kathleen A. Hinderer  

   Charles H. Christiansen, Sandra L. Rogers, and Kristine L. Haertl  

10. Disability Determination  
    Robert D. Rondinelli and Richard T. Katz  

11. Applying the ICF in Rehabilitation Medicine  
    Gerold Stucki, Nenad Kostanjsek, Bedirhan Üstün, Thomas Ewert, and Alarcos Cieza  

### PART II
### The Rehabilitation Team: Systems-Based Practice

12. Systematically Assessing and Improving the Quality and Outcomes of Medical Rehabilitation Programs  
    Mark V. Johnston, Kenneth J. Ottenbacher, James E. Graham, Patricia A. Findley, and Anne C. Hansen  

13. Rehabilitation Team Function and Prescriptions, Referrals, and Order Writing  
    John C. King, Karen J. Blankenship, William Schalla, and Amit Mehta  

14. Psychological Aspects of Rehabilitation  
    Daniel E. Rohe  

15. Speech, Language, Swallowing, and Auditory Rehabilitation  
    Beth Solomon, Carmen Brewer, Martin B. Brodsky, Jeffrey B. Palmer, and Jennifer Hyder  

16. Sexuality and Disability  
    Stacy Elliot and Andrei Krassioukov  

17. Vocational Rehabilitation, Independent Living, and Consumerism  
    Debra Homa and David DeLambo  

18. Community Participation and the Environment: Theoretical, Assessment and Clinical Implications  
    Julie J. Keysor and Daniel K. White  

19. Systems-Based Practice of Physical Medicine and Rehabilitation  
    Bruce M. Gans, Carolyn C. Zollar, Martha Kendrick, and Jared Fleisher  

20. Ethical Issues in Rehabilitation Medicine  
    Janet F. Haas  

21. Interactions with the Medicolegal System  
    Steve R. Geiringer
PART III

Major Conditions 549

23 Stroke Rehabilitation 551  
Joel Stein and Murray E. Brandstater

24 Traumatic Brain Injury 575  
John Whyte, Jennie Ponsford, Thomas Watanabe, and Tessa Hart

25 Multiple Sclerosis 625  
Francis Bethoux and Alexander Rae-Grant

26 Rehabilitation of Persons with Parkinson’s Disease and Other Movement Disorders 645  
Ron Hirschberg, Nutan Sharma, and Donna Moxley Scarborough

27 Rehabilitation of Spinal Cord Injury 665  
Steven Kirshblum and Monifa Brooks

28 Adult Motor Neuron Disease 717  
Lisa S. Krivickas and Gregory T. Carter

29 Peripheral Neuropathy 741  
Mark A. Thomas and Maya Therattil

30 Myopathy 757  
Jay J. Han and David D. Kilmer

31 Osteoarthritis 781  
Todd P. Stitik, Jong-Hyun Kim, Doreen Stiskal, Patrick Foye, Robert Nadler, James Wyss, and Sausheen Heslop

32 Disorders of the Cervical Spine 811  
Jason S. Lipetz and David I. Lipetz

33 Rehabilitation of Lumbar Spine Disorders: An Evidence-Based Clinical Practice Approach 837  
Stefano Negrini, Fabio Zaina, Michele Romano, Salvatore Atanasio, Claudia Fusco, and Carlo Trevisan

34 Scoliosis and Other Spinal Deformities 883  
Scott M. Paul

35 Upper Extremity Soft-Tissue Injuries 907  
Mark D. Klaiman and Kathleen Fink

36 Cumulative Trauma Disorders 923  
Philip Chiou and Joanne Borg-Stein

37 Hand Disorders 937  
Keith A. Bengtson, Jeffrey S. Brault, and Lynn H. Gerber

38 Foot Disorders 953  
Eric L. Kolodin, Thomas Vitale, and Lynn H. Gerber

39 The Prevention and Treatment of Osteoporosis 979  
Patricia Graham, Robert A. Adler, Francis J. Bonner, and Gopi Kasturi

40 Rehabilitation of the Patient with Rheumatic Diseases 1015  
Galen O. Joe, Jeanne E. Hicks, and Lynn H. Gerber

41 Cardiac Rehabilitation 1075  
Matthew Bartels

42 Rehabilitation of the Patient with Respiratory Dysfunction 1099  
John R. Bach and Eric Altschuler

43 Burn Rehabilitation 1125  
Jeffrey C. Schneider and M. Catherine Spires

44 Rehabilitation for Patients with Cancer Diagnoses 1151  
Mary M. Vargo, Justin C. Riutta, and Deborah J. Franklin

45 Vascular Diseases 1179  

46 Transplantation Medicine: A Rehabilitation Perspective 1207  
Ross D. Zafonte and Michael Munin

47 Rehabilitation of the Individual with HIV 1227  
Stephen F. Levinson and Steven M. Fine

PART IV

Secondary Conditions and Complications 1247

48 Physical Inactivity: Physiological and Functional Impairments and Their Treatment 1249  
Eugen M. Halar and Kathleen R. Bell

49 Treatment of the Patient with Chronic Pain 1273  
Nicolas E. Walsh, Daniel Santa Maria, and Maxim Eckmann

50 Spasticity and Muscle Overactivity as Components of the Upper Motor Neuron Syndrome 1319  
Elie P. Elovic, Michal E. Eisenberg, and Neil N. Jasey, Jr
<table>
<thead>
<tr>
<th>Page</th>
<th>Title</th>
<th>Author(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>51</td>
<td>Neurogenic Bladder and Bowel</td>
<td>Todd A. Linsenmeyer, James M. Stone, and Steven A. Steins</td>
</tr>
<tr>
<td>52</td>
<td>Pressure Ulcers</td>
<td>Chester H. Ho and Kath Bogie</td>
</tr>
<tr>
<td>53</td>
<td>Sports Medicine</td>
<td>Gerard A. Malanga, Jose A. Ramirez-Del Toro, Jay E. Bowen, Joseph H. Feinberg, and Garrett S. Hyman</td>
</tr>
<tr>
<td>54</td>
<td>Physical Activity for People with Disabilities</td>
<td>James H. Rimmer and Jennifer A. Gray-Stanley</td>
</tr>
<tr>
<td>55</td>
<td>Performing Arts Medicine</td>
<td>Sonya Rissmiller and Katie Weatherhogg</td>
</tr>
<tr>
<td>56</td>
<td>Children with Disabilities</td>
<td>Martin Diamond and Michael Amento</td>
</tr>
<tr>
<td>57</td>
<td>Adults with Childhood-Onset Disabling Conditions</td>
<td>Joyce Oleszek, Laura Pickler, and Dennis Matthews</td>
</tr>
<tr>
<td>58</td>
<td>Empowering Women with Disabilities to be Self-Determining in their Health Care</td>
<td>Kristi L. Kirschner, Judy Panko Reis, Debjani Mukherjee, and Cassing Hammond</td>
</tr>
<tr>
<td>59</td>
<td>Aging and Rehabilitation</td>
<td>Gary S. Clark, Patrick Kortebein, and Hilary C. Siebens</td>
</tr>
<tr>
<td>60</td>
<td>Primary Care for Persons with Disability</td>
<td>William L. Bodenek, Gerben DeJong, Indira S. Lanig, Michael Friedland, and Amanda Harrington</td>
</tr>
<tr>
<td>61</td>
<td>Manipulation, Massage, and Traction</td>
<td>J. Michael Wierting, Michael T. Andary, Todd G. Holmes, Adam Cugaj, Neal Cross, and Gregory Thompson</td>
</tr>
<tr>
<td>62</td>
<td>Pharmacotherapy of Disability</td>
<td>Todd P. Stitik, Robert Klecz, Brian D. Greenwald, and Jiaxin J. Tran</td>
</tr>
<tr>
<td>63</td>
<td>The Pharmacology of Analgesic Agents</td>
<td>David J. Jones and Antonio Hernandez</td>
</tr>
<tr>
<td>64</td>
<td>Injection Procedures</td>
<td>Nicolas E. Walsh and Maxim Eckmann</td>
</tr>
<tr>
<td>65</td>
<td>Spinal Injection Procedures</td>
<td>Boqing Chen, Todd P. Stitik, Patrick M. Foye, Jong-Hyun Kim, Michael Mehnert, Christopher Castro, Steve Aydin, Glen Chang, and Evish Kamrava</td>
</tr>
<tr>
<td>66</td>
<td>Palliative Care Symptom Management</td>
<td>Andrea L. Cheville</td>
</tr>
<tr>
<td>67</td>
<td>Evolution of Biofeedback in Physical Medicine and Rehabilitation</td>
<td>Steve L. Wolf and He (Helen) Huang</td>
</tr>
<tr>
<td>68</td>
<td>Therapeutic Electrical Stimulation in Neurorehabilitation</td>
<td>Lynne R. Sheffler, Jayme S. Knutson, and John Chae</td>
</tr>
<tr>
<td>69</td>
<td>Functional Neuromuscular Electrical Stimulation</td>
<td>Jayme S. Knutson, Lynne R. Sheffler, and John Chae</td>
</tr>
<tr>
<td>70</td>
<td>Assistive Technology</td>
<td>Cathy Bodine</td>
</tr>
<tr>
<td>71</td>
<td>Upper and Lower Extremity Prosthetics</td>
<td>Nicolas E. Walsh, Gordon Bosker, and Daniel Santa Maria</td>
</tr>
<tr>
<td>72</td>
<td>Upper Limb Orthotics</td>
<td>Heikki Uustal</td>
</tr>
<tr>
<td>73</td>
<td>Lower Extremity Orthotics, Shoes, and Gait Aids</td>
<td>Alberto Esquenazi</td>
</tr>
<tr>
<td>74</td>
<td>Spinal Orthotics</td>
<td>Ferne Pomerantz and Eva Durand</td>
</tr>
<tr>
<td>75</td>
<td>Complementary and Alternative Medicine</td>
<td>Ann C. Cotter, Lisa Bartoli, Jay Rosenfeld, Robert Schulman, and David Jung Seto</td>
</tr>
</tbody>
</table>
### CONTENTS

#### PART VII

**Scientific Advances** 2135

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>80</td>
<td>Evidence-Based Practice in Rehabilitation (Including Clinical Trials)</td>
<td>2137</td>
</tr>
<tr>
<td></td>
<td>Mary D. Slavin and Alan M. Jette</td>
<td></td>
</tr>
<tr>
<td>81</td>
<td>Neural Repair and Plasticity</td>
<td>2155</td>
</tr>
<tr>
<td></td>
<td>Steven C. Cramer</td>
<td></td>
</tr>
<tr>
<td>82</td>
<td>Regenerative Medicine: Implications in Rehabilitation</td>
<td>2173</td>
</tr>
<tr>
<td></td>
<td>Martin K. Childers, Chad Markert, Anthony Atala, Jennifer Cann, George Christ, Mark Furth, and Fabrisia Ambrosio</td>
<td></td>
</tr>
<tr>
<td>83</td>
<td>Rehabilitation Robotics</td>
<td>2187</td>
</tr>
</tbody>
</table>

**Index** I-1
CONTRIBUTORS

Hugo Núñez Bernadet, MD
Punta Del Este, Uruguay

Francois Bethoux, MD
Director, Rehabilitation Services
Mellen Center for MS Treatment and Research
The Cleveland Clinic Foundation
Cleveland, Ohio

Haraldur Bjarnason, MD
Professor
Department of Radiology
Mayo Clinic College of Medicine
Chief, Vascular and Interventional Radiology
Department of Radiology
Mayo Clinic
Rochester, Minnesota

Karen J. Blankenship, MA, RN
Director of Clinical Operations (Retired)
Department of Physical Medicine and Rehabilitation
University Hospital
San Antonio, Texas

William L. Bockenek, MD
Chairman
Department of Physical Medicine and Rehabilitation
Carolina Medical Center
Medical Director
Carolina Rehabilitation
Charlotte, North Carolina

Cathy Bodine, PhD
Assistant Professor
Department of Rehabilitation Medicine
University of Colorado School of Medicine
Denver, Colorado

Kath Bogie, DPhil
Cleveland, Ohio

Michael L. Boninger, MD
Chair and Associate Dean for Medical Student Research
Department of Physical Medicine and Rehabilitation
University of Pittsburgh School of Medicine
Medical Director
Human Engineering Research Laboratories
VA Pittsburgh Healthcare System
Pittsburgh, Pennsylvania

Francis Bonner, Jr., MD
Clinical Professor of Medicine
Drexel University College of Medicine
Philadelphia, Pennsylvania
Medical Director
Rehabilitation Center at South Jersey
Vineland, New Jersey

Joanne Borg-Stein, MD
Assistant Professor
Department of Physical Medicine and Rehabilitation
Boston, Massachusetts
Medical Director
Spaulding Rehabilitation Hospital-Wellesley and Newton Wellesley Hospital
Newton, Massachusetts

Gordon Bosker, MEd, CPO
Instructor/Research
Department of Rehabilitation Medicine
The University of Texas Health Science Center at San Antonio
Prosthetist/Orthotist
Department of Rehabilitation Medicine
University Health System
San Antonio, Texas

Jay E. Bowen, DO
Assistant Professor
Physical Medicine and Rehabilitation
University of Medicine and Dentistry of New Jersey
New Jersey Medical School
Newark, New Jersey
Kessler Institute for Rehabilitation
West Orange, New Jersey

Brian L. Bowyer, MD
Associate Professor
Director, Orthotic and Prosthetic Clinic
Department of Physical Medicine and Rehabilitation
The Ohio State University College of Medicine
Columbus, Ohio

Murray E. Brandstater, MBBS, PhD
Professor and Chair
Department of Physical Medicine and Rehabilitation
Loma Linda University Affiliated Hospitals
Loma Linda, California

Jeffrey S. Brault, DO
Assistant Professor
Department of Physical Medicine and Rehabilitation
Mayo Clinic College of Medicine
Rochester, Minnesota

Carmen Breuer, MD
Münster, Germany

Martin B. Brodsky, MD
Baltimore, Maryland

Monifa Brooks, MD
Staff Physiatrist
Kessler Institute for Rehabilitation
West Orange, New Jersey

Matthew A. Butters, MD
Assistant Professor
Department of Physical Medicine and Rehabilitation
Mayo Clinic College of Medicine
Rochester, Minnesota
Consultant
Department of Physical Medicine and Rehabilitation
Mayo Clinic Scottsdale
Scottsdale, Arizona

Jennifer Cann, BA
Wake Forest Institute for Regenerative Medicine
Wake Forest University Health Sciences
Winston-Salem, North Carolina

Gregory T. Carter, MD, MS
Professor
Co-Director
Muscular Dystrophy Association Amyotrophic Lateral Sclerosis Center
Department of Rehabilitation Medicine
University of Washington
Seattle, Washington

Christopher Castro, MD
Department of Physical Medicine and Rehabilitation
University of Medicine and Dentistry of New Jersey
New Jersey Medical School
Newark, New Jersey
John Chae, MD  
Professor  
Departments of Physical Medicine and Rehabilitation and Biomedical Engineering  
Case Western Reserve University  
Attending Physician  
Department of Physical Medicine and Rehabilitation  
MetroHealth Medical Center  
Cleveland, Ohio

Boqing Chen, MD, PhD  
Department of Physical Medicine and Rehabilitation  
University of Medicine and Dentistry of New Jersey  
New Jersey Medical School  
Newark, New Jersey

Glen Cheng, MD  
Department of Physical Medicine and Rehabilitation  
University of Medicine and Dentistry of New Jersey  
New Jersey Medical School  
Newark, New Jersey

Andrea L. Cheville, MD  
Department of Rehabilitation  
Mayo Clinic  
Rochester, Minnesota

Yi-Pin Chiang, MD  
Mackay Memorial Hospital  
Taipei, Taiwan, ROC

Martin K. Childers, DO, PhD  
Professor  
Department of Neurology and Wake Forest Institute for Regenerative Medicine  
Wake Forest University Health Sciences  
Winston-Salem, North Carolina

Philip Chiou, MD  
Resident Physician  
Department of Physical Medicine and Rehabilitation  
Harvard Medical School  
Resident Physician  
Department of Physical Medicine and Rehabilitation  
Spaulding Rehabilitation Hospital  
Boston, Massachusetts

George Christ, PhD  
Wake Forest Institute for Regenerative Medicine  
Wake Forest University Health Sciences  
Winston-Salem, North Carolina

Charles H. Christiansen, EdD, OTR OT(C), FAOTA  
Clinical Professor  
Division of Rehabilitation Sciences  
University of Texas Medical Branch  
Galveston, Texas

Nicolas Christodoulou, MD  
Nicosia, Cyprus

Alarcos Cieza, MPH  
Munich, Germany

Gary S. Clark, MD, MMM, CPE  
Professor and Chair  
Department of Physical Medicine and Rehabilitation  
Case Western Reserve University  
Medical Director  
MetroHealth Rehabilitation Institute of Ohio  
MetroHealth Medical Center  
Cleveland, Ohio

Andrew J. Cole, MD  
Clinical Associate Professor  
Department of Rehabilitation Medicine  
University of Washington  
Seattle, Washington

Edgar Colón, MD  
Professor  
Chief, Diagnostic Radiology  
Department of Radiological Sciences  
University of Puerto Rico School of Medicine  
San Juan, Puerto Rico

Rory A. Cooper, PhD  
Distinguished Professor and FISA/PVA Endowed Chair  
Department of Rehabilitation Science and Technology  
Senior Career Scientist and Director  
Human Engineering Research Laboratories  
VA Pittsburgh Healthcare System  
Pittsburgh, Pennsylvania

Rosemarie Cooper, MPT, ATP  
Clinical Instructor  
Department of Rehabilitation Science and Technology  
University of Pittsburgh  
Pittsburgh, Pennsylvania

Ann Nancy Cotter, MD  
Associate Clinical Professor  
Department of Physical Medicine and Rehabilitation  
University of Medicine and Dentistry of New Jersey  
New Jersey Medical School  
Newark, New Jersey  
Medical Director  
Integrative Medicine Division  
Atlantic Health  
Morristown and Summit, New Jersey

Steven C. Cramer, MD  
Professor  
Department of Neurology, Anatomy and Neurobiology  
University of California, Irvine  
Irvine, California

Neal A. Cross, PhD  
Chair and Professor  
Department of Anatomy  
Lincoln Memorial University – DeBusk College of Orthopaedic Medicine  
Harrogate, Tennessee

Adam Cugalj, DO  
Clinical Instructor  
Department of Physical Medicine and Rehabilitation  
Michigan State University College of Osteopathic Medicine  
East Lansing, Michigan  
Private Practice  
Granite Physiatry, PLLC  
Concord, New Hampshire

Martin Diamond, MD  
Children’s Specialized Hospital  
Mountainside, New Jersey

David A. DeLambo, RhD, CRC  
Associate Professor  
Department of Rehabilitation Counseling  
University of Wisconsin-Stout  
Menomonie, Wisconsin
CONTRIBUTORS

Gerben DeJong, PhD
Professor
Department of Physical Medicine and Rehabilitation
Georgetown University
Senior Fellow and Director
Center for Post-acute Studies
National Rehabilitation Hospital
Washington, DC

Eva Durand, MD
Private Practice
Valley Stream, New York

Maxim Eckmann, MD
San Antonio, Texas

Michael E. Eisenberg, MD
New York, New York

Stacy L. Elliott, MD
Clinical Professor
Department of Psychiatry and Urological Sciences
University of British Columbia
Medical Director
British Columbia Centre for Sexual Medicine
Department of Psychiatry
Vancouver Coastal Health Authority
Vancouver, British Columbia, Canada

Elie P. Elovic, MD
Professor
Department of Physical Medicine and Rehabilitation
University of Utah School of Medicine
Chief
Department of Physical Medicine and Rehabilitation
University of Utah Health Center
Salt Lake City, Utah

Rolland P. Erickson, MD
Assistant Professor
Department of Physical Medicine and Rehabilitation
Mayo Clinic College of Medicine
Rochester, Minnesota
Consultant
Department of Physical Medicine and Rehabilitation
Mayo Clinic Arizona
Scottsdale, Arizona

Alberto Esquenazi, MD
Professor
Department of Physical Medicine and Rehabilitation
Thomas Jefferson University
Philadelphia, Pennsylvania
Chair and Professor
Department of Physical Medicine and Rehabilitation
Moss Rehabilitation Hospital
Elkins Park, Pennsylvania

Thomas Ewert, PhD
Munich, Germany

Joseph H. Feinberg, MD, MS
Associate Professor
Department of Physiatry
Cornell University
Vice Chairman
Department of Physiatry
Hospital for Special Surgery
New York, New York

Patricia A. Findley, DrPH, MSW, LCSW
Assistant Professor
School of Social Work
Rutgers University
New Brunswick, New Jersey

Steven M. Fine, MD, PhD
Assistant Professor of Medicine
Infectious Diseases Division
University of Rochester Medical Center
Infectious Disease Attending
Department of Internal Medicine
Highland Hospital
Rochester, New York

Kathleen Fink, MD
Clinical Associate Professor
Department of Physical Medicine and Rehabilitation
Georgetown University Hospital
Washington, DC
National Rehabilitation Hospital
Bethesda, Maryland

Jared Fleisher, JD
Attorney
Patton Boggs LLP
Washington, DC

Patrick Foye, MD
Department of Physical Medicine and Rehabilitation
University of Medicine and Dentistry of New Jersey
New Jersey Medical School
Newark, New Jersey

Deborah J. Franklin, PhD, MD
Director
Cancer Rehabilitation
Department of Rehabilitation Medicine
Thomas Jefferson University Hospital
Philadelphia, Pennsylvania

John A. Freeman, DO
Instructor
Department of Physical Medicine and Rehabilitation
Mayo Clinic College of Medicine
Rochester, Minnesota
Consultant
Department of Physical Medicine and Rehabilitation
Mayo Clinic Arizona
Scottsdale, Arizona

Michael B. Friedland, MD
Associate Clinical Professor
Department of Medicine
University of North Carolina at Chapel Hill
Chapel Hill, North Carolina
Staff Physician
Department of Internal Medicine
Charlotte Medical Clinic
Carolinas Medical Center
Charlotte, North Carolina

Mark Furth, PhD
Technical Development Officer
Wake Forest Institute for Regenerative Medicine
Wake Forest University School of Medicine
Winston-Salem, North Carolina

Claudio Fusco, MD
Milan, Italy
CONTRIBUTORS

Gail L. Gamble, MD
Medical Director of Cancer Rehabilitation
Department of Physical Medicine and Rehabilitation
Rehabilitation Institute of Chicago
Chicago, Illinois

Bryan K. Ganter, MD
Assistant Professor
Department of Physical Medicine and Rehabilitation
Mayo Clinic College of Medicine
Rochester, Minnesota
Consultant
Department of Physical Medicine and Rehabilitation
Mayo Clinic Scottsdale
Scottsdale, Arizona

Steve R. Geiringer, MD
Clinical Professor
Department of Physical Medicine and Rehabilitation
Wayne State University
Detroit, Michigan

Lynn H. Gerber, MD
Director
Center for the Study of Chronic Illness and Disability
University Professor
George Mason University
Fairfax, Virginia

Alessandro Giustini, MD
Arezzo, Italy

Angel Gomez, MD
Assistant Professor
Department of Radiological Sciences
University of Puerto Rico School of Medicine
San Juan, Puerto Rico

James E. Graham, PhD, DC
Assistant Professor
Division of Rehabilitation Services
University of Texas Medical Branch
Galveston, Texas

Patricia Graham, MD
Princeton, New Jersey

Jennifer A. Gray-Stanley, PhD, MPP
Assistant Professor
Department of Public Health and Health Education
School of Nursing and Health Studies
Northern Illinois University
Dekalb, Illinois

Brian D. Greenwald, MD
Assistant Professor
Department of Rehabilitation Medicine
Mt. Sinai School of Medicine
New York, New York

Janet F. Haas, MD
William Penn Foundation
Philadelphia, Pennsylvania

Kristine L. Haertl, PhD, ORT/L
Associate Professor
Department of Occupational Science and Occupational Therapy
St. Catherine University
St. Paul, Minnesota

Eugene M. Halar, MD
Professor Emeritus
Department of Rehabilitation Medicine
University of Washington School of Medicine
Seattle, Washington

Cassing Hammond, MD
Assistant Professor
Obstetrics and Gynecology
Northwestern University
Director
Department of Obstetrics and Gynecology
Northwestern Memorial Hospital
Chicago, Illinois

Jay J. Han, MD
Director
Neuromuscular Medicine Fellowship
Department of Physical Medicine and Rehabilitation
University of California Davis Medical Center
Sacramento, California

Anne C. Hansen, MS/CCC-SP
Stroke and Neurovascular Program Coordinator
Froedtert and the Medical College of Wisconsin
Milwaukee, Wisconsin

Amanda Harrington, MD
SCIM Fellow
Department of Physical Medicine and Rehabilitation
Case Western Reserve University
MetroHealth Rehabilitation Institute of Ohio
Cleveland, Ohio

Tessa Hart, PhD
Institute Scientist
Moss Rehabilitation Research Institute
MossRehab
Elkins Park, Pennsylvania

Antonio Hernandez, MD
Associate Professor
Department of Anesthesiology
University of Texas Health Science Center at San Antonio
Director of Cardiothoracic and Transplant Anesthesiology
Department of Anesthesiology
University Hospital
San Antonio, Texas

Sausheen Heslop, MD
New York Hospital
Queens, New York

Stefan Hesse, MD
Professor
Department of Neurological Rehabilitation
Charité-University Medicine Berlin
Chief Consultant
Department of Neurology
Medical Park Berlin
Berlin, Germany

Jeanne E. Hicks, MD, FACR
Former Adjunct Associate Professor of Rehabilitation Medicine
Department of Orthopaedic Surgery
Georgetown University Medical Center
Washington, DC
Former Deputy Chief
Department of Rehabilitation Medicine
Warren G. Magnuson Clinical Center
National Institutes of Health
Bethesda, Maryland
Kathleen A. Hinderer, PhD, MPT, PT
Ann Arbor, Michigan

Stephen R. Hinderer, MD, MS, PT
Associate Professor
Department of Physical Medicine and Rehabilitation
Wayne State University School of Medicine
Rehabilitation Institute of Michigan
Detroit, Michigan

Ron Hirschberg, MD
Instructor
Department of Physical Medicine and Rehabilitation
Harvard Medical School
Spaulding Rehabilitation Hospital
Boston, Massachusetts

Chester Ho, MD
Assistant Professor
Department of Physical Medicine & Rehabilitation
Case Western Reserve University
Chief, Spinal Cord Injury Service
Louis Stokes Department of Veterans Affairs Medical Center
Cleveland, Ohio

Martin D. Hoffman, MD
Professor of Clinical Physical Medicine and Rehabilitation
University of California—Davis School of Medicine
Sacramento, California
Chief, Physical Medicine and Rehabilitation Service
VA Northern California Health Care System
Mather, California

Neville Hogan, PhD
Professor
Department of Mechanical Engineering
Massachusetts Institute of Technology
Newman Laboratory for Biomechanics and Human Rehabilitation
Veterans Health Administration
Department of Rehabilitation and Development
Cambridge, Massachusetts

Debra Homa, PhD, CVE, CRC
Associate Professor
Department of Rehabilitation and Counseling
University of Wisconsin-Stout
Menomonie, Wisconsin

He (Helen) Huang, PhD
Assistant Professor
Director of Neuromuscular Rehabilitation Engineering Laboratory
Department of Electrical, Computer, and Biomedical Engineering
University of Rhode Island
Kingston, Rhode Island

Garrett S. Hyman, MD
Northwest Spine and Sports Physicians
Bellevue, Washington

Neil N. Jasey, Jr.
New York, New York

Alan M. Jette, PhD
Director
Health and Disability Research Institute
Boston University School of Public Health
Boston, Massachusetts

Galen O. Joe, MD
Deputy Chief
Department of Rehabilitation
National Institutes of Health
Bethesda, Maryland

Mark V. Johnston, PhD
Professor
Department of Occupational Therapy
College of Health Sciences
University of Wisconsin-Milwaukee
Milwaukee, Wisconsin

David J. Jones, PhD
Professor
Departments of Anesthesiology and Pharmacology
University of Texas Health Science Center at San Antonio
San Antonio, Texas

Daniel A. Judelson, MD
Storrs, Connecticut

Evish Kamrava, MD
Department of Physical Medicine and Rehabilitation
University of Medicine and Dentistry of New Jersey
New Jersey Medical School
Newark, New Jersey

Gopi Kasturi, MD
Royal Oak, Michigan

Richard T. Katz, MD
St. Louis, Missouri

Martha Kendrick, JD
Partner
Patton Boggs LLP
Washington, DC

Julie J. Keysor, PhD, PT
Associate Professor
Department of Physical Therapy and Athletic Training
Boston University School of Medicine
Boston, Massachusetts

David D. Kilmer, MD (deceased)
Professor and Chair
Department of Physical Medicine and Rehabilitation
University of California Davis School of Medicine
Davis, California

Jong-Hyun Kim, MD
Department of Physical Medicine and Rehabilitation
University of Medicine and Dentistry of New Jersey
New Jersey Medical School
Newark, New Jersey

John C. King, MD
Professor
Department of Rehabilitation Medicine
The University of Texas Health Science Center at San Antonio
Director
Reeves Rehabilitation Center
University Hospital
San Antonio, Texas

Steven Kirshblum, MD
Associate Professor
Department of Physical Medicine and Rehabilitation
University of Medicine and Dentistry of New Jersey
New Jersey Medical School
Newark, New Jersey
Associate Medical Director and Director of Spinal Cord Injury Services
Kessler Institute for Rehabilitation
West Orange, New Jersey
Kristi L. Kirschner, MD
Professor
Clinical Medical Humanities and Bioethics and PM&R
Northwestern University
Feinberg School of Medicine
Attending Physician
Schwab Rehabilitation Hospital
Chicago, Illinois

Mark D. Klaiman, MD
Consultant
Department of Rehabilitation Medicine
National Institutes of Health
Active Associate
Department of Rehabilitation
Suburban Hospital
Bethesda, Maryland

Robert Klecz, MD
Clinical Chief of Orthopedic Rehabilitation
Kessler Institute for Rehabilitation, Saddle Brook Campus
Saddle Brook, New Jersey

Jayme Knutson, PhD
Researcher
Department of Biomedical Engineering
Case Western Reserve University
Bioscientific Staff
Department of Physical Medicine and Rehabilitation
MetroHealth Medical Center
Cleveland, Ohio

Eric L. Kolodin, DPM
Assistant Attending
Department of Surgery, Section of Podiatric Medicine
Saint Barnabas Medical Center
Livingston, New Jersey

Patrick Kortebein, MD
Assistant Professor
Physical Medicine & Rehabilitation, and Geriatrics
University of Arkansas for Medical Sciences
Little Rock, Arkansas

Nenad Kostanjsek, MD
Geneva, Switzerland

William J. Kraemer, PhD
Professor of Kinesiology, Physiology, and Neurobiology
The University of Connecticut School of Medicine
Storrs, Connecticut

Andrei Krasioiov, MD, PhD, FRCPC
Associate Professor
Division of Physical Medicine and Rehabilitation
Associate Director and Scientist, ICORD
Director of Autonomic Research Unit
Staff Physician
Spinal Cord Program
GF Strong Rehabilitation Center
Department of Medicine
University of British Columbia
Vancouver, British Columbia, Canada

Hermano Igo Krebs, PhD
Principal Research Scientist and Lecturer
Massachusetts Institute of Technology
Department of Mechanical Engineering
Cambridge, Massachusetts

Lisa S. Krivickas, MD (deceased)
Associate Professor
Harvard Medical School
Associate Chief
Physical Medicine and Rehabilitation
Director of EMG
Spaulding Rehabilitation Hospital
Boston, Massachusetts

Eduardo J. Labat, MD
Neuroradiology Fellow
Department of Radiology
Mallinckrodt Institute of Radiology
Washington University
St. Louis, Missouri

Indira S. Lanig, MD
Department of Physical Medicine and Rehabilitation
University of Colorado Denver School of Medicine
Denver, Colorado

Stephen F. Levinson, MD, PhD
Associate Professor
Department of Physical Medicine and Rehabilitation
Wayne State University
Detroit, Michigan
Physiatrist
Kessler Institute for Rehabilitation
West Orange, New Jersey

Henry L. Low, MD, PhD
Professor, Department of Physical Medicine and Rehabilitation
Virginia Commonwealth University, School of Medicine
National Consultant for Research and Education
Defense and Veterans Brain Injury Center (DVBIC)
Richmond, Virginia

Jan Lexell, MD, PhD
Professor
Division of Rehabilitation Medicine
Department of Clinical Sciences, Lund University
Medical Director
Department of Rehabilitation Medicine
Skåne University Hospital
Lund, Sweden

Leonard S.W. Li, MD
Hong Kong, China

Todd A. Linsenmeyer, MD
Associate Professor
Department of Physical Medicine and Rehabilitation
Department of Surgery (Division of Urology)
University of Medicine and Dentistry of New Jersey
New Jersey Medical School
Newark, New Jersey
Director of Urology
Kessler Institute for Rehabilitation
West Orange, New Jersey

David Lipetz, MSPT, OCS, Cert. MDT
Private Practice
Garden City, New York

Jason S. Lipetz, MD
Assistant Professor
Department of Rehabilitation Medicine
Albert Einstein College of Medicine
Bronx, New York
Attending Physician
Department of Physical Medicine and Rehabilitation
North Shore Long Island Jewish Hospital
Manhasset/Glen Cove/New Hyde Park, New York
Albert C. Lo, MD, PhD  
Assistant Professor  
Departments of Neurology, Community Health, Engineering  
Brown University  
Staff Physician  
Department of Internal Medicine  
Providence Veterans Affairs Medical Center  
Providence, Rhode Island

Gerard A. Malanga, MD  
Associate Professor  
Department of Physical Medicine and Rehabilitation  
University of Medicine and Dentistry of New Jersey  
New Jersey Medical School  
Newark, New Jersey  
Director, Sports, Spine, and Orthopedic Rehabilitation  
Kessler Institute for Rehabilitation  
West Orange, New Jersey

Chad Markert, PhD  
Postdoctoral Research Fellow  
Wake Forest Institute for Regenerative Medicine  
Wake Forest University  
Winston-Salem, North Carolina

Dennis Matthews, MD  
Professor and Chair  
Department of Physical Medicine and Rehabilitation  
University of Colorado Denver School of Medicine  
Chairman  
Department of Rehabilitation Medicine  
The Children’s Hospital  
Aurora, Colorado

Laura McClure, MPT, PhD  
Research Scientist  
Human Engineering Research Laboratories  
Department of Veterans Affairs  
University of Pittsburgh  
Seating Specialist  
Center for Assistive Technology  
University of Pittsburgh Medical Center  
Pittsburgh, Pennsylvania

Michael Mehnert, MD  
Department of Physical Medicine and Rehabilitation  
University of Medicine and Dentistry of New Jersey  
New Jersey Medical School  
Newark, New Jersey

Amit Mehta, PT, MBA  
Director, Therapy Service  
Reeves Rehabilitation Center  
University Hospital  
San Antonio, Texas

Peter J. Moley, MD  
Assistant Professor  
Department of Rehabilitation Medicine  
Weill Cornell Medical College  
Assistant Attending Physiatrist  
Department of Physiatry  
The Hospital for Special Surgery  
New York, New York

Debjani Mukherjee, PhD  
Chicago, Illinois

Michael Munin, MD  
Associate Professor  
Department of Physical Medicine and Rehabilitation  
University of Pittsburgh  
Pittsburgh, Pennsylvania

Robert Nadler, CPO  
Hanger Prosthetics and Orthotics  
Hamilton, New Jersey

Stefano Negriini, MD  
Milan, Italy

Stephen F. Noll, MD  
Assistant Professor  
Department of Physical Medicine and Rehabilitation  
Mayo Clinic College of Medicine  
Rochester, Minnesota  
Medical Director, Rehabilitation  
Department of Physical Medicine and Rehabilitation  
Mayo Clinic Scottsdale  
Scottsdale, Arizona

Gustavo S. Oderich, MD  
Assistant Professor of Surgery  
Division of Vascular Surgery  
Mayo Clinic College of Medicine

Joyce Oleszek, MD  
Assistant Professor  
Department of Rehabilitation  
The Children’s Hospital/University of Colorado Hospital  
Aurora, Colorado

Kenneth J. Ottenbacher, PhD, OTR  
Professor and Director  
Division of Rehabilitation Sciences  
University of Texas Medical Branch  
Galveston, Texas

Jeffrey B. Palmer, MD  
The Laurence Cardinal Sheehan Professor and Chair  
Department of Physical Medicine and Rehabilitation  
Johns Hopkins School of Medicine  
Physiatrist-in-Chief  
Johns Hopkins Hospital  
Baltimore, Maryland

Chang-il Park, MD  
Department and Research Institute of Rehabilitation Medicine  
Yonsei University College of Medicine  
Seoul, Korea

Kirsten S. Paynter, MD  
Assistant Professor  
Department of Physical Medicine and Rehabilitation  
Mayo Clinic College of Medicine  
Rochester, Minnesota

William S. Pease, MD  
Ernest W. Johnson Professor and Chairman  
Department of Physical Medicine and Rehabilitation  
The Ohio State University College of Medicine and Public Health  
Medical Director, Dodd Hall Rehabilitation Program  
The Ohio State University Hospital  
Columbus, Ohio

Laura Pickler, MD, MPH  
Assistant Professor  
Department of Family Medicine and Pediatrics  
University of Colorado Health Sciences Center  
Aurora, Colorado
Ferne R. Pomerantz, MD  
Assistant Professor  
Department of Rehabilitation and Regenerative Medicine  
Columbia University College of Physicians and Surgeons  
Director  
Spinal Cord Injury Service  
Helen Hayes Hospital  
West Haverstraw, New York

Jennie L. Ponsford, PhD  
Professor of Neuropsychology  
School of Psychology and Psychiatry  
Monash University  
Clayton, Victoria, Australia  
Director  
Monash-Epworth Rehabilitation Research Centre  
Epworth Hospital  
Richmond, Victoria, Australia

Alexander Rae-Grant, MD  
Clinical Associate Professor  
Department of Medicine  
The Cleveland Clinic  
Lerner College of Medicine of Case Western Reserve University  
Staff Neurologist  
Neurological Institute  
Cleveland Clinic  
Cleveland, Ohio

Jose A. Ramirez-del Toro, MD  
Pittsburgh, Pennsylvania

Judy Panko Reis, MA, MS  
Health Care Access Policy Analyst  
Access Living of Metropolitan Chicago  
Adjunct Instructor  
College of Nursing  
University of Illinois Chicago  
Formerly Director of the Women with Disabilities Center at the Rehabilitation Institute of Chicago 1991–2009  
Chicago, Illinois

James Rimmer, PhD  
Professor  
Department of Disability and Human Development  
University of Illinois at Chicago  
Chicago, Illinois

Haim Ring, MD (deceased)  
Professor of Physical Medicine and Rehabilitation  
Deputy Director  
Loewenstein Rehabilitation Center  
Tel Aviv, Israel

Sonya Rissmiller, MD  
Charlotte, North Carolina

Justin C. Riutta, MD  
Director  
Breast Cancer Rehabilitation and Lymphedema Program  
William Beaumont Hospital  
Royal Oak, Michigan

Sandra L. Rogers, PhD, OTR/L  
Associate Professor  
School of Occupational Therapy  
College of Health Professions  
Pacific University  
Hillsboro, Oregon

Daniel E. Rohe, PhD, LP  
Rehabilitation Psychologist  
Department of Physical Medicine and Rehabilitation  
Mayo Clinic  
Rochester, Minnesota

Michele Romano, MD  
Vigevano, Italy

Robert D. Rondinelli, MD  
Des Moines, Iowa

Jay Rosenfeld, MD  
Staff Physiatrist  
Rehabilitation Hospital of the Cape and Islands  
East Sandwich, Massachusetts

Jennifer Ryder, MA, CCC-SLP  
Speech Language Pathologist  
National Institutes of Health  
Bethesda, Maryland

Daniel L. Santa Maria, MD  
Assistant Professor  
Department of Rehabilitation Medicine  
The University of Texas Health Science Center at San Antonio  
Medical Director  
Medical Arts and Research Center  
University of Texas Medicine  
San Antonio, Texas

Donna Moxley Scarborough, PhD  
Miami, Florida

William Schalla, OTR, BS, BA  
Inpatient Rehabilitation Team Leader  
Reeves Rehabilitation Center  
University Hospital  
San Antonio, Texas

Mark Schmeler, PhD, OTR/L, ATP  
Assistant Professor  
Rehabilitation Science and Technology  
University of Pittsburgh  
Occupational Therapist  
Center for Assistive Technology  
University of Pennsylvania Medical Center Health System  
Pittsburgh, Pennsylvania

Robert Schulman, MD  
Physiatrist  
New York, New York

Lynne Sheffler, MD  
Assistant Professor  
Department of Physical Medicine and Rehabilitation  
Case Western Reserve University  
Physician  
Department of Physical Medicine and Rehabilitation  
MetroHealth Medical Center  
Cleveland, Ohio

Jeffrey C. Schneider, MD  
Assistant Professor  
Department of Physical Medicine and Rehabilitation  
Harvard Medical School  
Medical Director  
Trauma, Burn, and Orthopaedic Program  
Department of Physical Medicine and Rehabilitation  
Spaulding Rehabilitation Hospital  
Boston, Massachusetts
Nutan Sharma, MD
Boston, Massachusetts

Hilary C. Siebens, MD
Professor
Department of Physical Medicine and Rehabilitation
University of Virginia
Charlottesville, Virginia
Associate Professor of Medicine
University of California at Los Angeles
Los Angeles, California

David L. Skaggs, MD
Associate Professor
Department of Orthopaedic Surgery
University of Southern California
Endowed Chair of Spinal Disorders
Department of Orthopaedic Surgery
Children’s Hospital of Los Angeles
Los Angeles, California

Mary D. Slavin, PT, PhD
Health and Disability Research Institute
Boston University
Boston, Massachusetts

Beth Solomon, MS, CCC-SLP
Chief
Speech, Language, and Pathology Section
Department of Rehabilitation Medicine
National Institutes of Health
Bethesda, Maryland

Mary Catherine Spires, MD
Professor
Department of Physical Medicine and Rehabilitation
University of Michigan
Associate Chair for Clinical Affairs
Department of Physical Medicine and Rehabilitation
University of Michigan Hospital System
Ann Arbor, Michigan

Joel Stein, MD, PhD
Simon Baruch Professor
Department of Rehabilitation and Regenerative Medicine
Columbia University College of Physicians and Surgeons
Professor
Division of Rehabilitation Medicine
Weill Cornell Medical College
Chief
Department of Rehabilitation Medicine
New York-Presbyterian Hospital
New York, New York

Steven A. Stiens, MD, MS
Associate Professor
Department of Rehabilitation Medicine
Director of Fellowship Program
Spinal Cord Medicine
University of Washington
Attending Physician
Spinal Cord Unit
VA Puget Sound Health Care System
Seattle, Washington

Doreen Stiskal, PT, PhD
Associate
Department of Physical Medicine and Rehabilitation
University of Medicine and Dentistry of New Jersey
New Jersey Medical School
Newark, New Jersey
Assistant Chair
Graduation Programs in Health Sciences
School of Graduate Medical Education
Seton Hall University
West Orange, New Jersey

Todd P. Stitik, MD
Professor and Director
Occupational/Musculoskeletal Medicine
Physical Medicine and Rehabilitation
New Jersey Medical School
Newark, New Jersey

James M. Stone, MD
Private Practice
Northern California Surgical Group A Medical Corp.
Redding, California

David M. Strick, PhD, PT
Assistant Professor
Department of Physiology
Physical Therapist
Department of Physical Medicine and Rehabilitation
Mayo Clinic
Rochester, Minnesota

Gerold Stucki, MD, MS
Professor and Chair
Department of Health Services and Health Policy
University of Lucerne
Lucerne, Switzerland
Director
Swiss Paraplegic Research
Nottwil, Switzerland

Maya Therattil, MD
Assistant Professor
Department of Physical Medicine and Rehabilitation
Albert Einstein College of Medicine
Director, Spinal Cord Medicine
Department of Physical Medicine and Rehabilitation
Montefiore Medical Center
Bronx, New York

Mark A. Thomas, MD
Associate Professor
Department of Physical Medicine and Rehabilitation
Albert Einstein College of Medicine
Associate Chairman and Program Director
Department of Physical Medicine and Rehabilitation
Montefiore Medical Center
Bronx, New York

Jiaxin J. Tran, MD
Physical Medicine and Rehabilitation Resident
University of Medicine and Dentistry of New Jersey
New Jersey Medical School
Newark, New Jersey

Carlo Trevisan, MD
Padua, Italy

Wen-Chung Tsai, MD
Professor
Chang Gung Memorial Hospital
Taipei, Taiwan, ROC
CONTRIBUTORS

Gregory A. Thompson, DO
Chair of Osteopathic Principles and Practice
Department of Clinical Medicine
Lincoln Memorial University – DeBusk College of Osteopathic Medicine
Harrogate, Tennessee

Norman S. Turner III, MD
Associate Professor and Consultant
Department of Orthopaedic Surgery
Mayo Clinic College of Medicine
Rochester, Minnesota

Bedirhan Üstün, MD
Geneva, Switzerland

Heikki Uustal, MD
Associate Professor
Department of Medicine
University of Medicine and Dentistry of New Jersey
New Brunswick, New Jersey
Medical Director
Prosthetic/Orthotic Team
Department of Rehabilitation Medicine
JFK-Johnson Rehabilitation Institute
Edison, New Jersey

Mary M. Vargo, MD
Associate Professor
Department of Physical Medicine and Rehabilitation
Case Western Reserve University
Cleveland, Ohio

Jorge Vidal, PhD
Adjunct Professor
Department of Radiological Sciences
University of Puerto Rico School of Medicine
San Juan, Puerto Rico

Thomas D. Vitale, DPM
Instructor, Department of Surgery
New York College of Podiatric Medicine
New York, New York
Assistant Attending
Section of Podiatry
Saint Barnabas Medical Center
Livingston, New Jersey

B.T. Volpe, MD
Professor of Neurology and Neuroscience
Cornell University
New York, New York

Tyng-Guey Wang, MD
Associate Professor
Department of Physical Medicine and Rehabilitation
National Taiwan University Hospital
College of Medicine
National Taiwan University
Taipei, Taiwan, ROC

Thomas K. Watanabe, MD
Clinical Director
Drucker Brain Injury Center
Department of Physical Medicine and Rehabilitation
MossRehab at Elkins Park/Albert Einstein Medical Center
Elkins Park, Pennsylvania

Katie Weatherhogg, MD
Springfield, Missouri

Daniel K. White, PT, ScD
Research Assistant Professor
Department of Physical Therapy and Athletic Training
Boston University
College of Health and Rehabilitation Science
Sargent College
Boston, Massachusetts

John Whyte, MD, PhD
Professor
Department of Rehabilitation Medicine
Thomas Jefferson University
Philadelphia, Pennsylvania
Director
Moss Rehabilitation Research Institute
MossRehab
Elkins Park, Pennsylvania

J. Michael Wieting, DO
Professor of Physical Medicine and Rehabilitation
Director of Sports Medicine
Lincoln Memorial University – DeBusk College of Osteopathic Medicine
Harrogate, Tennessee

Steve L. Wolf, PhD
Professor
Department of Medicine
University of Pennsylvania
Philadelphia, Pennsylvania

John Wyss, MD, PT
Department of Physical Medicine and Rehabilitation
University of Medicine and Dentistry of New Jersey
New Jersey Medical School
Newark, New Jersey

Vincent J. Yacyshyn, MD
Clinical Associate Professor
Department of Cardiology
University of Calgary
Consultant
Cardiovascular Associates
Calgary, Alberta, Canada

Mark Young, MD
Baltimore, Maryland

Ross D. Zafonte, DO
Earl P. and Ida S. Charlton Professor and Chair of Physical Medicine and Rehabilitation
Department of Physical Medicine and Rehabilitation
Harvard Medical School
Boston, Massachusetts

Fabio Zaina, MD
Vigevano, Italy

Carolyn C. Zollar
Washington, DC
The field of Physical Medicine and Rehabilitation focuses on the restoration of health and function and reintegration of the patient into the community. The goal of Physical Medicine and Rehabilitation: Principles and Practice is to organize, summarize, discuss, and make available knowledge in the field to assist the developing or established practitioner in these endeavors. This fifth edition is divided in two volumes for ease of use by the practitioner and to accommodate the dramatic increase in information and knowledge in the field of physical medicine and rehabilitation since the publication of the fourth edition. This edition also introduces a classification of the chapters in new sections.

The content of the book has been extensively revised and expanded. There are six new chapters and over thirty major revisions in this edition. Our goal is to provide a comprehensive, thorough, and multidisciplinary discussion covering the depth and breadth of the science of physical medicine and rehabilitation and the evidence that supports current best practice. Chapters cover the scientific fundamentals of our field as well as the state-of-the-art clinical interventions used in the treatment and rehabilitation of patients with a wide variety of diseases and disabilities. Authors for each chapter were chosen for their experience and expertise in their given topic. This text reflects the efforts of over two hundred contributing authors representing all parts of the world.

The editorial board for this edition has changed. Joel A. DeLisa, MD, Editor-in-Chief of the first four editions of the book has become Editor Emeritus. Walter R. Frontera, MD, PhD has assumed the role of Editor-in-Chief. In addition, a wonderful group of three Editors and nine Associate Editors have worked diligently to make this fifth edition a world class treatise.

The editors of this edition would like to express their appreciation to each of the editors and authors of previous editions. Their work has contributed in a special way to this current version. We also would like to acknowledge the hard work of the authors of this current edition; they have helped create an excellent source of knowledge for those interested in physical medicine and rehabilitation. Their commitment and dedication have made this an exciting and productive effort.

We hope this fifth edition of Physical Medicine and Rehabilitation: Principles and Practice contributes significantly to the advancement of the field. As an essential resource for the training and continuing education of medical rehabilitation professionals, this text will help ensure that the care they provide to people with disabling conditions is of the highest quality, resulting in improvement in their health, function, and quality of life.

Walter R. Frontera
Joel A. DeLisa
Bruce M. Gans
Nicolas E. Walsh
Lawrence R. Robinson
Jeffrey R. Basford
William L. Bockenek
Gregory T. Carter
John Chae
Lynn H. Gerber
Alan M. Jette
Todd P. Stitik
Gerold Stucki
Ross D. Zafonte
Principles of Assessment and Evaluation
Clinical Evaluation

OVERVIEW

Physical medicine and rehabilitation focuses on the restoration of function and the subsequent reintegration of the patient into the community. As with other branches of medicine, the cornerstone of physical medicine and rehabilitation is a meticulous and thorough clinical evaluation of the patient. Therapeutic intervention by physiatrists must be based on proper assessment of the patient. Impaired function cannot be isolated from preexisting and concurrent medical problems or from the social circumstances of the individual patient.

Evaluation of Function

Medical diagnosis focuses on the historical clues and physical findings that lead the examiner to the correct identification of disease. After a medical diagnosis is established, the physiatrist must ascertain functional consequences of the disease. Appropriate clinical evaluation requires the examiner to have a clear understanding of the distinctions among the disease, body functions, activity limitations, and participation restrictions, as discussed in Chapter 19.

If a disease cannot be eliminated or its severity cannot be reduced through medical or surgical means, measures are used to minimize its impact on functioning. For example, a weak muscle can be strengthened or a hearing impairment can be minimized with the use of an electronic aid. When disease process that leads to impaired functioning cannot be eliminated or its severity cannot be reduced, physiatric intervention must address the limitations on activity and the restrictions on participation. For successful rehabilitation, the physiatrist not only must address the consequences of impaired functioning directly but also must identify intact functional capabilities. When intact capabilities and their use are augmented and adapted to new uses, functional independence can be enhanced.

Case 1

AW had gained much enjoyment and self-esteem as a competitive runner before a spinal cord injury that left him with paraplegia. During and after inpatient rehabilitation, he vigorously pursued a cardiovascular and upper extremity conditioning program. He obtained an ultralightweight sport wheelchair and resumed competitive athletics as a wheelchair racer, winning several regional races.

Comment: AW’s intact capabilities included normal arm strength, a competitive spirit, and self-discipline. Through augmentation and adaptation and with the use of an appropriate wheelchair, he regained enjoyment and self-esteem in his athletic endeavors.

Sometimes it is not possible to ascertain the specific disease responsible for a patient’s constellation of historical, physical, and laboratory findings. Medical management must then address the symptoms of the patient. Although diagnosis is highly desirable, it is not a prerequisite to the identification and subsequent management of functional loss. To determine the expectations of future activity in relation to past activity, the physiatrist should attempt to characterize the temporal nature of the disease process over time.

Case 2

FZ, a 62-year-old woman, had difficulty climbing stairs. When questioned, she revealed that she and her husband had been in the habit of taking a 30-minute evening walk for many years. However, 2 years earlier, fatigue began to limit her walk to no more than a few blocks. During the previous year, she had had difficulty rising from low seating, and 6 months previously, she reluctantly quit taking walks. During the preceding few weeks, she had found that climbing stairs was a burden, and she had started taking showers because she needed assistance getting out of the bathtub.

FZ reported no sensory deficits. Physical examination showed hypotonic muscle stretch reflexes and predominantly proximal muscle weakness. Electrodiagnostic studies and muscle biopsy demonstrated a noninflammatory myopathy; however, further extensive evaluation failed to determine a cause. FZ was provided with a bath bench, a toilet seat riser, a lightweight folding wheelchair for long-distance mobility, and a cane for short distances. She was instructed in safe ambulation with the cane, operation of the wheelchair, energy conservation techniques, and the proper placement of bathroom safety bars. Safe automobile operation was documented, and she was provided with documentation to obtain a handicapped parking permit. The functional impact of potentially progressive muscle weakness was discussed with her, and she was given supportive counseling.

When FZ returned for a follow-up examination 1 month later, muscle testing showed only a slight progression of her weakness, and her functional capabilities were unchanged. Another follow-up examination was scheduled for 6 weeks later.

Comment: Although a specific diagnosis was not established, rehabilitation intervention addressed the patient’s specific functional losses. Serial evaluations at regular follow-up intervals allowed the physiatrist to identify and minimize future functional loss.
Comprehensiveness of Evaluation
The scope of physical medicine and rehabilitation encompasses more than a single organ system. Attention to the whole person is paramount. The objective of the physiatrist is to eliminate disability and restore functioning. The goal is to empower the individual to attain the fullest possible physical, mental, social, and economic independence by maximizing activity and participation. Consequently, the evaluation must assess not only the disease but also the way the disease affects and is affected by the person’s family and social environment, vocational responsibilities and economic state, avocational interests, hopes, and dreams.

Cases 3 and 4
CC, a 63-year-old piano tuner, had a left cerebral infarction manifested only as minimal dysfunction of the dominant right hand. Despite demonstrating discrete function of the digits of the involved hand on physical examination, he was psychologically devastated to find that he could no longer accomplish the fine and precise motor patterns necessary to continue in his profession.

BD, a 63-year-old corporate attorney, had a left cerebral infarction resulting in severe spastic weakness of his nondominant upper extremity. He completed some paperwork every day during his inpatient rehabilitation and returned to full-time employment shortly after finishing treatment.

Comment: For each person, the degree of functional impairment is uniquely and disproportionately related to the extent of resultant limitations in activity and restrictions in participation.

Interdisciplinary Nature of Evaluation
Although most of this chapter addresses the patient history and physical examination as related to the rehabilitation evaluation, these are only part of the comprehensive rehabilitation assessment. This statement is not meant to deprecate the usefulness of these traditional tools for the physician. Both are of critical importance and serve as the basis for further evaluation; yet, by their very nature, they are also limited. Speech and language disorders can inhibit communication. Subjective interpretation of the facts by the patient and the family can cloud the objective assessment of function. Performance cannot be optimally assessed by interview and physical examination alone.

For example, asking the patient about ambulation skills during the interview may identify a potential problem, but such skills can only be assessed objectively and reliably by having the physician and physical therapist observe the patient’s ambulation in various situations. Likewise, the occupational therapist must assess the performance of activities of daily living, and the rehabilitation nurse must assess the safety and judgment of the hospitalized patient. The speech therapist furnishes a measured assessment of language function and, through special communication skills, may obtain information from the patient that was missed during the interview. The rehabilitation psychologist provides a quantified and standardized assessment of cognitive and perceptual function and a skilled assessment of the patient’s current psychological state. Through interaction with the patient’s family and employer, the social worker can provide useful information that is otherwise unavailable regarding the patient’s social support system and economic resources. The concept of the physical medicine and rehabilitation team applies not only to evaluation of the patient but also to ongoing management of the rehabilitation process in the outpatient as well as the inpatient practice setting.

Setting and Purpose
Because of the expanding scope of physical medicine and rehabilitation, the evaluation setting can be diverse.

<table>
<thead>
<tr>
<th>Setting</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hospital</td>
<td></td>
</tr>
<tr>
<td>Inpatient rehabilitation unit</td>
<td>Comprehensive evaluation by the rehabilitation team</td>
</tr>
<tr>
<td>Off-service consultation</td>
<td>Assessment by physiatrist of potential for rehabilitation benefit</td>
</tr>
<tr>
<td>Clinic</td>
<td></td>
</tr>
<tr>
<td>General physiatry clinic</td>
<td>Comprehensive evaluation by the team</td>
</tr>
<tr>
<td>Assessment by physiatrist of potential for rehabilitation benefit</td>
<td>Thorough evaluation of musculoskeletal or spine disorder</td>
</tr>
<tr>
<td>Special clinic</td>
<td>Thorough evaluation of specific disease group (e.g., muscular dystrophy or sports injury)</td>
</tr>
<tr>
<td>Day rehabilitation program</td>
<td>Comprehensive evaluation by rehabilitation team</td>
</tr>
<tr>
<td>Impairment or disability clinic</td>
<td>Evaluation determined by requirement of referring agency (e.g., workers’ compensation or Social Security Administration)</td>
</tr>
<tr>
<td>Community nursing home</td>
<td>Comprehensive evaluation by rehabilitation team</td>
</tr>
<tr>
<td>Limited assessment by selected members of rehabilitation team</td>
<td>Limited assessment by selected members of rehabilitation team</td>
</tr>
<tr>
<td>School</td>
<td>Assessment by physiatrist of potential for rehabilitation benefit</td>
</tr>
<tr>
<td>Transitional living facilities</td>
<td>Limited evaluation of functioning</td>
</tr>
<tr>
<td>Comprehensive evaluation by rehabilitation team</td>
<td>Limited evaluation for participation in sports</td>
</tr>
<tr>
<td>Limited assessment of specific problem</td>
<td>Comprehensive evaluation by rehabilitation team</td>
</tr>
</tbody>
</table>
A necessary corollary to the setting is the purpose of the evaluation. Both the setting and the purpose will affect the format and extent of the evaluation. Traditionally, the inpatient rehabilitation unit or the outpatient physiatry clinic has been the optimal setting for a comprehensive evaluation by the entire rehabilitation team. However, in these days of increasing medical costs and intervention by the government and other third-party payers, creative approaches may be required to accomplish comprehensive rehabilitation evaluations in the clinic and elsewhere in the community (Table 1-1).

PATIENT HISTORY

Ordinarily, the patient history is obtained in an interview of the patient by the physiatrist. If communication disorders and cognitive deficits are encountered during the evaluation, additional corroborative information must be obtained from significant others accompanying the patient. The spouse and family members can be valuable resources. The physiatrist may also find it necessary to interview other caregivers, such as paid attendants, public health nurses, and home health agency aides.

The major components of the patient history are the chief report of symptoms, the history of the present illness, the functional history, the past medical history, a review of systems, the patient profile, and the family history.

Chief Report of Symptoms

The goal in assessing the chief report of symptoms is to document the patient's primary concern in his or her own words. Patients often report an impairment in the form of a symptom that may suggest a certain disease or group of diseases. A report of "chest pain when I walk up a flight of stairs" suggests cardiac disease, whereas a report that "my hands ache and go numb when I drive" hints at carpal tunnel syndrome.

Of equal importance is the recognition that a chief report of impaired function may also be the first implication of activity limitation or participation restriction. A homemaker's report that "my balance has been getting worse and I've fallen several times" may be related to disease involving the vestibular system and to disability created by unsafe ambulation. Similarly, a farmer's declaration that "I can no longer climb up onto my tractor" not only suggests a neuromuscular or orthopedic disease but also conveys that the disorder has resulted in disability because the patient is not able to fulfill vocational expectations.

History of Present Illness

The history of the present illness is obtained when the patient relates the development of the present illness. When necessary, patients should be asked to define the specific words they use. Specific questions relating to a particular symptom can also help focus the interview. Using these techniques, the physician can gently guide the patient to follow a chronological sequence and to fully describe the symptoms and their consequences. Above all, the patient should be allowed to tell the story. More than one symptom may be elicited during the interview, and the physician should document each problem in an orderly fashion (Table 1-2) (1).

A complete list of the patient's current medications should be obtained. Polypharmacy is commonly encountered in people with chronic disease, at times with striking adverse effects. Side effects of medications can further impede cognition, psychological state, vascular reflexes, balance, bowel and bladder control, muscle tone, and coordination already impaired by the present illness or injury.

The history of the present illness should include a record of handedness, which is important in many areas of rehabilitation.

Functional History

The physiatric evaluation of a patient with chronic disease often reveals impaired function. The functional history enables the physiatrist to characterize the disabilities that have resulted from disease and to identify remaining functional capabilities. Some physicians consider the functional history to be part of the history of the present illness, whereas others view it as a separate segment of the patient interview. The examiner must know not only the functional status associated with the present illness but also the premorbid level of function.

Although the specific organization of the activities of daily living varies somewhat, the following elements of personal independence are constant: communication, eating, grooming, bathing, toileting, dressing, bed activities, transfers, and mobility.

When obtaining the functional history, the physician may record in a descriptive paragraph the patient's level of independence in each activity. However, functional stability is best communicated, followed over time, and made accessible for study when the physician uses a standard functional assessment scale, as discussed in Chapter 28.

### TABLE 1.2 Analysis of Symptoms

1. Date of onset
2. Character and severity
3. Location and extension
4. Time relationships
5. Associated symptoms
6. Aggravating and alleviating factors
7. Previous treatment and effects
8. Progress, noting remission and exacerbations

Communication

Because a major component of physical medicine and rehabilitation practice is patient education, effective communication is essential. The interviewer must assess the patient's communication options. In the clinical situation, this aspect of the evaluation blurs the distinction between history and physical examination. It is difficult to interact with the patient in a meaningful way without coincidentally examining his or her ability to communicate; significant speech and language deficiencies become obvious. However, for purposes of discussion, certain facets of the assessment relate more specifically to the history and will be discussed here. Additional facets are presented below in the section on the physical examination.

Speech and language pathology has provided clinicians with numerous classification systems for speech and language disorders (see Chapter 15). From a functional view, the elements of communication hinge on four abilities related to speech and language (2):

1. Listening
2. Reading
3. Speaking
4. Writing

By assessing these factors as well as comprehension and memory, the examiner can determine a patient's communication abilities. Representative questions include the following:

1. Do you have difficulty hearing?
2. Do you use a hearing aid?
3. Do you have difficulty reading?
4. Do you need glasses to read?
5. Do others find it hard to understand what you say?
6. Do you have problems putting your thoughts into words?
7. Do you have difficulty finding words?
8. Can you write?
9. Can you type?
10. Do you use any communication aids?

Eating

The abilities to present solid food and liquids to the mouth, to chew, and to swallow are basic skills taken for granted by able-bodied people. However, in patients with neurologic, orthopedic, or oncologic disorders, these tasks can be formidable. Dysfunctional eating is associated with far-reaching consequences, such as malnutrition, aspiration pneumonitis, and depression. As in the assessment of other skills for activities of daily living, eating function should be assessed specifically and methodically.

Representative questions include the following:

1. Can you eat without help?
2. Do you have difficulty opening containers or pouring liquids?
3. Can you cut meat?
4. Do you have difficulty handling a fork, knife, or spoon?
5. Do you have problems bringing food or beverages to your mouth?
6. Do you have problems chewing?
7. Do you have difficulty swallowing solids or liquids?
8. Do you ever choke?
9. Do you regurgitate food or liquids through your nose?

Patients with nasogastric or gastrostomy tubes should be asked who helps them prepare and administer their feedings. The type, quantity, and schedule of feedings should be recorded.

Grooming

Grooming may not be considered as important as feeding. However, impaired functioning that leads to deficits in grooming can have deleterious effects on hygiene as well as on body image and self-esteem. Consequently, grooming skills should be of real concern to the rehabilitation team.

Representative questions include the following:

1. Can you brush your teeth without help?
2. Can you remove and replace your dentures without help?
3. Do you have problems fixing or combing your hair?
4. Can you apply your makeup independently?
5. Do you have problems shaving?
6. Can you apply deodorant without assistance?

Bathing

The ability to maintain cleanliness also has far-reaching physical and psychosocial implications. Deficits in cleaning can result in skin maceration and ulceration, skin and systemic infections, and the spread of disease to others. Patients should be questioned about their ability to bathe independently.

Representative questions include the following:

1. Can you take a tub bath or shower without assistance?
2. Do you feel safe in the tub or shower?
3. Do you use a bath bench or shower chair?
4. Can you accomplish a sponge bath without help?
5. Are there parts of your body that you cannot reach?

For patients with sensory deficits, bathing is also a convenient time for skin inspection, and inquiry about the patient’s inspection habits should be made. For patients using a wheelchair, walker, or other mobility device, architectural barriers to bathroom entry should be determined.

Toileting

To the cognitively intact person, incontinence of stool or urine can be a psychologically devastating deficit of personal independence. Ineffective bowel or bladder control has an adverse impact on self-esteem, body image, and sexuality, and it can lead to participation restriction. The soiling of skin and clothing can result in ulceration, infection, and urologic complications. The physiatrist should vigorously but sensitively pursue questioning about toileting dependency.
Representative questions include the following:

1. Can you use the toilet without assistance?
2. Do you need help with clothing before or after using the toilet?
3. Do you need help with cleaning after a bowel movement?

For patients with indwelling urinary catheters, the usual management of the catheter and leg bag should be examined. If bladder emptying is accomplished by intermittent catheterization, the examiner should determine who performs the catheterization and should have a clear understanding of his or her technique. For patients who have had ostomies for urine or feces, the examiner should determine who cares for the ostomy and should ask the patient to describe the technique.

Feminine hygiene is generally performed while on or near the toilet, so at this point in the interview, it may be appropriate to ask about problems with the use of sanitary napkins or tampons.

Dressing
We dress to go out into the world to be employed in the workplace, to dine in restaurants, to be entertained in public places, and to visit friends. Even at home, convention dictates that we dress to entertain anyone except close friends and family. We dress for protection, warmth, self-esteem, and pleasure. Dependency in dressing can result in a severe limitation to personal independence and should be investigated thoroughly during the rehabilitation interview.

Representative questions include the following:

1. Do you dress daily?
2. What articles of clothing do you regularly wear?
3. Do you require assistance putting on or taking off your underwear, shirt, slacks, skirt, dress, coat, stockings, panty hose, shoes, tie, or coat?
4. Do you need help with buttons, zippers, hooks, snaps, or shoelaces?
5. Do you use clothing modifications?

Bed Activities
The most basic stage of functional mobility is independence in bed activities. The importance of this functional level should not be underestimated. Persons who cannot turn from side to side to redistribute pressure and periodically expose their skin to the air are at high risk for development of pressure sores over bony prominences and skin maceration from heat and occlusion. For the person who cannot stand upright to dress, bridging (lifting the hips off the bed in the supine position) will allow the donning of underwear and slacks. Independence is likewise enhanced by an ability to move between a recumbent position and a sitting position. Sitting balance is required to accomplish many other activities of daily living, including transfers.

Representative questions include the following:

1. Can you move to and from the wheelchair to the bed, toilet, bath bench, standard seating, or car seat without assistance?
2. Can you get out of bed without difficulty?
3. Do you require assistance to rise to a standing position from either a low or a high seat?
4. Can you get on and off the toilet without help?

Transfers
The second stage of functional mobility is independence in transfers. Being able to move between a wheelchair and the bed, toilet, bath bench, shower chair, standard seating, or car seat often serves as a precursor to independence in other areas. Although a male patient can use a urinal to void without having to transfer, a female patient cannot be independent in bladder care without the ability to transfer to the toilet and will probably require an indwelling catheter. Travel by airplane or train is difficult without the ability to transfer from the wheelchair to other seating. Bathing or showering is not independent without the ability to move to the bath bench or shower chair. The inability to transfer to a car seat precludes the use of a motor vehicle with standard seats. Also included in this category is the ability to rise from a seated position to a standing position. Low seats without arm supports present a much greater problem than straight-backed chairs with arm supports.

Representative questions include the following:

1. When lying down, can you turn onto your front, back, and sides without assistance?
2. Can you lift your hips off the bed when lying on your back?
3. Do you need help to sit or lie down?
4. Do you have difficulty maintaining a seated position?
5. Can you operate the bed controls on an electric hospital bed?

Wheelchair Mobility
Although wheelchair independence is more likely than walking to be inhibited by architectural barriers, it provides excellent mobility for the person who is not able to walk. With efficiently engineered, lightweight manual wheelchairs, the energy expenditure required to wheel on flat ground is only slightly greater than that of walking. With the addition of a motorized drive, battery power, and controls for speed and direction, a wheelchair can be propelled even by a person who lacks the upper extremity strength necessary to propel a manual wheelchair, and it can thus help maintain independence in mobility.

Quantification of manual wheelchair skills can be accomplished in several ways. Patients may report in feet, yards, meters, or city blocks the distance that they are able to traverse before resting. Alternatively, the number of minutes they can continuously propel the chair can be specified, or the environment in which they are able to use the chair can be described (e.g., within a single room, around the house, or throughout the community).
Representative questions include the following:
1. Do you propel your wheelchair yourself?
2. Do you need help to lock the wheelchair brakes before transfers?
3. Do you require assistance to cross high-pile carpets, rough ground, or inclines in your wheelchair?
4. How far or how many minutes can you walk before you must rest?
5. Can you move independently about your living room, bedroom, and kitchen?
6. Do you go out to stores, to restaurants, and to friends’ homes?

With any of these functional levels of wheelchair mobility, patients should be asked what keeps them from going farther afield and whether they need help lifting the wheelchair into and out of an automobile.

**Ambulation**

The final level of mobility is ambulation. In the narrowest sense of the word, ambulation is walking, and we have used this definition to simplify the following discussion. However, within the sphere of rehabilitation, ambulation may be any useful means of movement from one place to another. In the view of many rehabilitation professionals, the person with a bilateral above-knee amputation ambulates with a manual wheelchair, the patient with C4 tetraplegia ambulates with a motorized wheelchair, and the survivor of polio in an underdeveloped country might ambulate by crawling. Driving a motor vehicle may also be considered a form of ambulation. Ambulation ability can be quantified the same way wheelchair mobility is quantified. Persons may report the distance they are able to walk, how long they can walk before they require a rest period, and the scope of the environment within which they walk.

Representative questions include the following:
1. Do you walk unaided?
2. Do you use a cane, crutches, or a walker to walk?
3. How far or how many minutes can you walk before you must rest?
4. What stops you from going farther?
5. Do you feel unsteady or do you fall?
6. Can you go upstairs and downstairs unassisted?
7. Do you go out to stores, to restaurants, and to friends’ homes?
8. Can you use public transportation (e.g., the bus or subway) without assistance?

**Operation of a Motor Vehicle**

In the perception of many patients, full independence in mobility is not attained without the ability to operate a motor vehicle on one’s own. Although driving skills are by no means necessary for urban dwellers with readily available public transportation, they may be essential to persons living in a suburban or rural environment. Driving skills should always be assessed in patients of driving age.

Representative questions include the following:
1. Do you have a valid driver’s license?
2. Do you own a car?
3. Do you drive your car to stores, to restaurants, and to friends’ homes?
4. Do you drive in heavy traffic or over long distances?
5. Do you drive in low light or after sunset?
6. Do you use hand controls or other automobile modifications?
7. Have you been involved in any motor vehicle accidents or received any citations for improper operation of a motor vehicle since your illness or injury?

**Past Medical History**

The past medical history is a record of any major illness, trauma, or health maintenance since the patient’s birth. The effects of certain past conditions will continue to affect the present level of function. Identifying these conditions affords an opportunity to better characterize the patient’s baseline level of function before the present disorder. The examiner must take special care to decipher whether the patient’s diagnostic terms accurately represent the true diagnoses. Although many past conditions associated with extensive immobilization, deconditioning, and disability are themselves amenable to rehabilitation measures, such conditions tend to affect the goals for future rehabilitation efforts.

**Case 5**

PB, a 66-year-old woman, was referred for rehabilitation after an above-knee amputation of her right leg because of vascular disease. Her past history was notable for a right cerebral infarction 7 years earlier. Despite comprehensive rehabilitation after the stroke, PB was able to walk only one block with a quadruped cane and an ankle-foot orthosis because of spastic left hemiparesis.

*Comment*: After prosthetic fitting and training, most people in this age group with an above-knee amputation regain ambulation skills, albeit with a cane or other gait aid. However, because PB had a preexisting ambulation disability due to the left hemiparesis that had occurred before amputation, her rehabilitation goals included a wheelchair prescription, with consideration of a hemichair if she could not wheel with her left arm, and training in wheelchair activities. Although ambulation beyond a few yards was not feasible, a preparatory prosthesis with a manual knee lock was provided on a trial basis to determine whether it aided in transfers. For PB, ambulation disability was dictated more by her previous impairments than by impairments associated with her present illness.

All elements of the standard past medical history should be completed; however, a history of neurologic, cardiopulmonary, or musculoskeletal disease should alert the physiatrist to special needs of the patient. Psychiatric disorders are also of
special interest to the physiatrist and are discussed below in the section on the psychological and psychiatric history.

**Neurologic Disorders**
Most frequently encountered in older populations but possibly present in any age group, a past history of neurologic disease can have a tremendous impact on the rehabilitation outcome of an unrelated current illness. Whether congenital or acquired, preexisting cognitive impairment places restrictions on educationally oriented rehabilitation intervention. Disorders with sensory manifestations such as loss of touch, pain, or joint position or affections that are characterized by perceptual dysfunction can retard the patient’s ability to monitor performance during the acquisition of new functional skills. These maladies also render patients more likely to be unresponsive to soft-tissue injury from prolonged or excessive skin surface pressures during long periods of immobility. When these conditions are coupled with preexisting visual or auditory impairment, the function is further encumbered. Likewise, new motor learning can be inhibited by a residual motor deficit that results in spasticity, weakness, or decreased endurance. A diligent search for antecedent neurologic disease should be a fundamental part of the rehabilitation evaluation.

**Cardiopulmonary Disorders**
For patients with motor disabilities, the activities of daily living require more than the normal expenditure of energy. When preexisting cardiopulmonary disorders limit the patient’s capacity to tolerate the greater energy expenditures imposed by motor impairment, they can result in additional functional deficits. This is also the case with many forms of hematologic, renal, and hepatic dysfunction. The physician should gather as much cardiopulmonary data as needed to estimate cardiac reserve accurately. Only when disease of the cardiopulmonary system is identified and addressed can medical intervention be initiated and rehabilitation tailored to maximize cardiac reserve.

**Musculoskeletal Disorders**
Weakness, joint ankylosis, or instability from previous trauma or arthritis, amputation, and other musculoskeletal dysfunctions can all affect functional capacity deleteriously. A search for such disorders is a necessary prerequisite to a complete physiatric evaluation.

**Review of Systems**
The systems should be thoroughly reviewed to screen for clues to disease not otherwise identified in the history of the present illness or in the past medical history. Many diseases have the potential to cause adverse effects on rehabilitation outcomes. However, as described previously, certain disorders are of special interest to the physiatrist. This part of the evaluation considers constitutional, head and neck, respiratory, cardiovascular, gastrointestinal, genitourinary, musculoskeletal, neurologic, psychiatric, endocrine, and dermatologic symptoms.

**Constitutional Symptoms**
Of particular interest to the examiner are suggestions of infection and nutritional deficiency. Fatigue can be a prominent symptom in patients with neurologic and neuromuscular conditions such as stroke, multiple sclerosis, amyotrophic lateral sclerosis, or poliomyelitis sequelae, or with other conditions such as obstructive sleep apnea or chronic pain syndromes.

**Head and Neck Symptoms**
Vision, hearing, and swallowing deficits must be identified.

**Respiratory Symptoms**
Any pulmonary condition that inhibits delivery of oxygen to the tissues will adversely affect endurance. Symptoms such as dyspnea, cough, sputum, hemoptysis, wheezing, and pleuritic chest pain should be identified.

**Cardiovascular Symptoms**
The manifestations of heart disease restrict cardiac reserve and endurance. When identified, many cardiovascular conditions can be ameliorated through medical management. Identifying arrhythmias may help prevent recurrent strokes of embolic cause. The presence of chest pain, dyspnea, orthopnea, palpitations, or light-headedness should be determined.

Peripheral vascular disease is the leading cause of amputation. The potential for ulceration and gangrene caused by bed rest, orthoses, pressure garments, and other rehabilitation equipment can be minimized if peripheral disease is recognized. The patient should be asked about claudication, foot ulcers, and varicosities.

**Gastrointestinal Symptoms**
Almost any form of gastrointestinal tract disease can result in nutritional deficiency, a particularly insidious condition that limits rehabilitation efforts more frequently than previously realized (3). Bowel control is of special interest for patients with neurologic disorders. These patients should be asked about incontinence, bowel care techniques, and use of laxatives.

**Genitourinary Symptoms**
Manifestations of neurogenic bladder must be sought. Questions should be asked about specific fluid intake, voiding schedules, specific bladder-emptying techniques, urgency, frequency, incontinence, retention and incomplete emptying, sensation of fullness and voiding, dysuria, pyuria, infections, flank pain, hematuria, and renal stones.

For female patients, a menstrual and pregnancy history should be obtained, and inquiries should focus on dyspareunia, vaginal and clitoral sensation, and orgasm. Male patients should be asked about erection, ejaculation, progeny, and pain during intercourse.

**Musculoskeletal Symptoms**
The musculoskeletal system review must be thorough because of the likelihood of musculoskeletal dysfunction in patients in a rehabilitation program. The examiner should ask about
muscle pain, weakness, fasciculation, atrophy, hypertrophy, skeletal deformities and fractures, limited joint motion, joint stiffness, joint pain, and swelling of soft tissues and joints.

**Neurologic Symptoms**
Because of the increased prevalence of neurologic disorders in patients in a rehabilitation program, a methodical neurologic review should always be performed. The following areas should be addressed: sense of smell, diplopia, blurred vision, visual field cuts, imbalance, vertigo, tinnitus, weakness, tremors, involuntary movements, convulsions, depressed consciousness, ataxia, loss of touch, pain, temperature, dysesthesia, hyperpathia, and changes in memory and thinking.

Chewing, swallowing, hearing, reading, and speaking may be addressed in either the functional history or the review of systems.

**Psychiatric Symptoms**
Psychological and psychiatric issues can be discussed during the review of symptoms. However, we prefer to explore this area while obtaining the psychosocial history for the patient profile.

**Endocrine Symptoms**
Screening questions should be presented to address intolerance to hot or cold, excessive sweating, increase in urine, increase in thirst, and changes in skin, hair distribution, and voice.

**Dermatologic Symptoms**
Rash, itching, pigmentation, moisture or dryness, texture, changes in hair growth, and nail changes should be questioned.

**Patient Profile**
The patient profile provides the interviewer with information about the patient’s present and past psychological state, social milieu, and vocational background.

**Personal History**

**Psychological and Psychiatric History**
Any present illness accompanied by functional loss can be psychologically challenging. A quiescent major psychiatric disturbance may resurface during such stressful times and may hinder or halt rehabilitation efforts. When the examiner is able to identify a history of psychiatric dysfunction, the necessary support systems to lessen the likelihood of recrudescence can be applied prophylactically during rehabilitation. The examiner should seek a history of previous psychiatric hospitalization, psychotropic pharmacologic intervention, or psychotherapy. The patient should be screened for past or current anxiety, depression and other mood changes, sleep disturbances, delusions, hallucinations, obsessive and phobic ideas, and past major and minor psychiatric illnesses. A review of the patient’s prior and current responses to stress often helps the rehabilitation team to better understand and modify behavioral responses to catastrophic illness or trauma. Therefore, it is important to know the patient’s emotional responses to previous illness and family troubles and to know how the stress of the current illness is being addressed. If initial screening suggests any abnormality, a clinical psychologist can conduct tests to clarify psychological symptoms or to identify a personality disturbance.

**Lifestyle**
Leisure activities can promote both physical health and emotional health. The patient’s leisure habits should be reviewed to identify special rehabilitation measures that might return independence in these activities. Examples of questions to consider include the following (4):

1. What sorts of interests do you have?
2. Are you more interested in intellectual pursuits (i.e., symbol oriented) than physical endeavors?
3. Do you derive the most pleasure from social interactions, organizations, and group functions (i.e., interpersonally oriented)?
4. Have you been actively pursuing any of these interests?

The work-oriented person without avocational interests before the present illness will need recreational counseling during rehabilitation.

**Diet**
Inadequate nutrition may inhibit rehabilitation efforts. In addition, even after initial myocardial and cerebrovascular events due to atherosclerosis, some secondary prevention can be accomplished through dietary intervention. The examiner should determine the patient’s ability to prepare meals and snacks, as well as the patient’s usual dietary habits and special diets.

**Alcohol and Drugs**
Drug, alcohol, and nicotine use must be assessed. Patients with cognitive, perceptual, and motor deficits can be further impaired to a dangerous degree through substance abuse. The use of alcohol or drugs is frequently a factor in head and spinal cord injuries. Identifying abuse and dependency provides an opportunity to help the patient modify future behavior through counseling. The CAGE questionnaire is a brief but useful screening vehicle for assessing alcohol abuse and dependency (Table 1-3); a single affirmative answer should initiate further investigation (5).

**TABLE 1.3 The CAGE Questionnaire**

<table>
<thead>
<tr>
<th>Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Have you ever felt you ought to cut down on your drinking?</td>
</tr>
<tr>
<td>2. Have people annoyed you by criticizing your drinking?</td>
</tr>
<tr>
<td>3. Have you ever felt bad or guilty about your drinking?</td>
</tr>
<tr>
<td>4. Have you ever had a drink first thing in the morning to steady your nerves or get rid of a hangover (eye-opener)?</td>
</tr>
</tbody>
</table>

Social History

Family

Catastrophic illness in a family member places enormous stress on the rest of the family. When the family is already facing other problems with interaction, health, or substance abuse, the potential is greater for disintegration of the family unit. This tendency is unfortunate because the availability of a sturdy support system of family and friends can be as predictive of disposition as it is of functional outcome. The examiner should determine the patient's marriage history and marital status and should obtain the names and ages of other family members who live in the home. The established roles of each member should be understood clearly (e.g., who handles the finances, the cooking, the cleaning, or the discipline). The examiner should also determine whether other family members live nearby. To ascertain the availability of all potential assistants, the examiner should inquire about their willingness and ability to participate in the care of the patient and about their work or school schedule.

Home

The patient’s home design should be reviewed to identify architectural barriers. The examiner should determine whether the patient owns or rents the home, the location of the home (e.g., urban, suburban, or rural), the distance between the home and rehabilitation services, the number of steps into the home, the presence of (or room for) entry ramps, and the accessibility of the kitchen, bath, bedroom, and living room.

Vocational History

Education and Training

Although education does not predict intellectual function, the educational level achieved by the patient may suggest intellectual skills upon which the rehabilitation team can draw during the patient’s convalescence. In addition, when coupled with the assessment of physical function, the educational background will dictate future educational and training needs. After determining the years of education completed by the patient and whether high school, undergraduate, or graduate degrees were obtained, the examiner should review the patient's performance. The acquisition of special skills, licenses, and certifications should be noted. Future vocational goals are always important to address but are of particular concern with adolescent patients. A discussion of these goals should indicate the need for and the type of interest, aptitude, and skills testing and any appropriate vocational counseling.

Work History

An understanding of the patient's work experience can help determine the need for further education and training. It also provides an idea of the patient’s motivation, reliability, and self-discipline. The duration and type of previous jobs and the reason for job changes should be recorded. Not only titles but also actual job descriptions must be obtained, and the patient should be asked about architectural barriers within the workplace. These principles also apply to the patient who works at home. In addition, the examiner should define the specific expectations related to meal preparation, shopping, home maintenance, cleaning, child rearing, and discipline. Finally, the examiner should ask where clothes are washed and whether architectural barriers prevent the patient from reaching appliances or areas in the home and yard.

Finances

The physical medicine and rehabilitation team, in particular the social worker or case manager, should have a basic understanding of the patient’s income, investment, and insurance resources, disability classifications, and debts. This financial information is important in determining the services and assistance to which an individual patient may be entitled.

Family History

The family history can be used to identify hereditary disease in the family and to assess the health of people in the patient's home support system. Knowledge of the health and fitness of the spouse and other family members can aid dismissal planning.

PHYSICAL EXAMINATION

The physical examination performed by the physiatrist has much in common with the general medical examination. Of necessity, it is a well-practiced art. Through perceptions gleaned from observation, palpation, percussion, and auscultation, the examining physician seeks physical findings to support and formulate the diagnosis and to screen for other conditions not suggested by the patient history. The physical examination also differs from the general medical examination. After investigating the physical findings that help to establish the medical diagnosis, the physiatrist still has two principal tasks:

1. To scrutinize the patient for physical findings that can help define the functional impairments emanating from the disease
2. To identify the patient’s remaining physical, psychological, and intellectual strengths that can serve as the base for reestablishing functioning

Physical medicine and rehabilitation emphasizes the orthopedic and neurologic examinations and makes assessment of function an integral part of the overall physical examination. Severe motor, cognitive, and communication impairments make it difficult or impossible for some patients to follow the directions of the physician, and these impairments limit certain traditional physical examination maneuvers. Thus, creativity is often required to accomplish the examination. Expert examination skills are particularly necessary in such situations.

We assume that the reader is competent in the performance of the general medical examination (6). The following discussion places priority on the aspects of the physical examination that have special relevance to physical medicine

CHAPTER 1 | CLINICAL EVALUATION
and rehabilitation. The major segments of the physical examination are vital signs and general appearance, integument and lymphatics, head, eyes, ears, nose, mouth and throat, neck, chest, heart and peripheral vascular system, abdomen, genitourinary system and rectum, musculoskeletal system, neurologic examination, and functional examination.

**Vital Signs and General Appearance**

The recording of blood pressure, pulse, temperature, weight, and general observations is important. The identification of hypertension may be meaningful to the secondary prevention of stroke and myocardial infarction. Supine, sitting, and standing blood pressures should be obtained to rule out orthostasis in any patient who has had unexplained falls, light-headedness, or dizziness. Tachycardia can be the initial manifestation of sepsis in a patient with high-level tetraplegia, or it can suggest pulmonary embolism in an immobilized patient. Initial weight recordings are invaluable to identify and follow up malnutrition, obesity, and fluid and electrolyte disorders common after various forms of brain injury. A notation should be made if patients act hostile, tense, or agitated or if their behavior is uncooperative, inappropriate, or preoccupied.

**Integument and Lymphatics**

Skin disorders are frequently encountered in patients undergoing physical rehabilitation. Prolonged pressure in patients with peripheral vascular disease, sensory disorders, immobility, and altered consciousness often results in damage to skin and underlying tissues. Many diseases common to disabled persons, and their treatments, render the skin more prone to trauma and infection. Skin problems that are only somewhat bothersome to able-bodied people can be devastating to persons with disabilities when they interfere with the use of prostheses, orthoses, and other devices.

The patient’s skin should be inspected in appropriate lighting. By considering the skin as each separate body region is examined, the physiatrist can study the entire body surface without total exposure of the patient. In particular, the skin over bony prominences and in contact with prosthetic and orthotic devices should be examined for lichenification, erythema, or breakdown. Intertriginous areas should be inspected for maceration and ulceration; the distal lower extremities in patients with vascular disease should be examined for pigmentation, hair loss, and breakdown; and the hands and feet in insensate patients should be observed for unrecognized trauma. All common lymph node sites should be palpated for enlargement and tenderness, and areas of edema should be palpated for pitting.

**Head**

The head should be inspected for signs of past or present trauma. Gentle palpation should be performed for evidence of previous trauma or neurosurgical procedures, shunt pumps, and other craniofacial abnormalities. Auscultation for bruits should be done when considering vascular malformations.

**Eyes**

Unrecognized acuity errors can hamper rehabilitation efforts, especially in patients needing adequate eyesight to compensate for disorders of other sensory systems. With the patient’s usual eyewear in place, far and near vision should be tested with the use of standard charts. If charts are not available, the patient’s vision can be compared with the examiner’s vision by object identification and description for far vision and by reading materials of several print sizes for near vision. Findings can be substantiated with refraction when circumstances permit. An ophthalmoscopic examination should be performed; if dilatory agents are necessary, one with a short duration can be used; notation should be made in the patient’s chart of the time of administration and the name of the preparation. Evidence of erythema and inflammation of the globe or conjunctiva should be sought; aphasic patients and those with altered consciousness may not adequately express the pain of acute glaucoma or the discomfort of conjunctivitis. The eyes of comatose patients should be inspected for inadequate lid closure; deficient lubrication should be compensated for to prevent corneal ulcerations.

**Ears**

Unrecognized hearing impairment can limit rehabilitation efforts. Hearing acuity can be checked with the “watch test” or by having the patient listen to and repeat words that are whispered. If a unilateral hearing deficit is identified, the Weber test and the Rinne test can be used to determine whether it is a nerve or conductive loss. Findings can be substantiated with an audiogram. An otoscopic examination should be performed. If otorrhea is present in a head-injured patient, Benedict solution can be used to assess the presence of sugar, which indicates cerebrospinal fluid.

**Nose**

A routine examination of the nose, including olfactory function, is generally sufficient. Clear or blood-tinged drainage in a head-injured patient indicates the presence of cerebrospinal fluid.

**Mouth and Throat**

The oral and pharyngeal mucosa should be inspected for poor hygiene and infections (e.g., candidiasis in patients taking corticosteroids or broad-spectrum antibiotics), the teeth for disrepair, and the gums for gingivitis or hypertrophy. Dentures should be checked for fit and maintenance. In patients with arthritis or trauma, the temporomandibular joints should be inspected and palpated for crepitation, tenderness, swelling, or limited motion. Any of these problems can impair food and fluid intake, resulting in poor nutrition.

**Neck**

A routine examination of the neck is generally sufficient. The examiner should listen for carotid bruits in patients with atherosclerosis and cerebrovascular disorders. In patients with musculoskeletal disorders, range of motion (ROM) should be assessed. However, neck motion need not be checked in
patients with recent trauma or chronic polyarthritis until radiographic studies have ruled out fracture or instability.

**Chest**
Tolerance to exercise is considerably affected by pulmonary function. For patients whose exercise tolerance is already compromised by neurologic or musculoskeletal disease, the examiner must search rigorously for pulmonary dysfunction to minimize the deficit. The standard medical maneuvers are usually sufficient; however, certain aspects of the chest examination merit mention.

The chest wall should be inspected to note the rate, amplitude, and rhythm of breathing. The presence of cough, hiccups, labored breathing, accessory muscle activity, and chest wall deformities should be noted. Respiration may be restricted by rheumatologic disorders such as advanced spondyloarthropathies and scleroderma. Likewise, restrictive pulmonary disease with hypoventilation is common in muscular dystrophy and other neuromuscular diseases, severe kyphoscoliosis, and chronic spinal cord injuries. Tachypnea and tachycardia may be the only readily apparent manifestations of pulmonary embolism, pneumonia, or sepsis after a high-level spinal cord injury. The finding of a barrel chest may lead the examiner to identify obstructive pulmonary disease so that medical management can minimize its effect on functioning.

The patient should be instructed to cough, and the force and efficiency of this action should be noted. If the cough is weak, the patient can be assisted by exerting manual pressure over the abdomen coincidentally with the cough to observe the effect. The chest wall should be palpated for tenderness, deformity, and transmitted sounds. During the acute care of a head-injured patient, rib fractures may be missed. Percussion should be performed to document diaphragmatic level and excursion. Auscultation should be performed to characterize breath sounds and to identify wheezes, rubs, rhonchi, and rales. Pneumonitis can be especially insidious in the immunosuppressed patient.

When pulmonary disease is suspected, further investigation with pulmonary function tests and a determination of blood gas levels may need to be undertaken. If the patient has a tracheostomy, the skin around the opening should be examined; the type of apparatus recorded, and cuff leaks noted. Screening for breast malignancy may be necessary in women and men alike.

**Heart and Peripheral Vascular System**
As with pulmonary disease, cardiovascular dysfunction can adversely affect exercise tolerance already encumbered by neurologic or musculoskeletal disease. When cardiovascular disorders are identified, intervention can relieve or reduce deleterious effects on exercise tolerance and general health. Implementation of appropriate secondary prevention measures for embolic stroke is contingent upon the identification of arrhythmias, valvular disease, and congenital anomalies.

In the clinical situation, peripheral circulation is usually assessed during examination of the patient’s limbs. When bracing is being considered, the examiner should search for the pallor and cool dystrophic skin of arterial occlusive disease; inappropriate devices may lead to edema and subsequent skin breakdown. Deep venous thrombosis is a major risk to immobilized patients, who should be examined for varicose and incompetent veins. Bedside Doppler studies should be used as necessary to help delineate arterial or venous concerns such as Raynaud phenomenon.

**Abdomen**
In many patients, the general medical examination of the abdomen is the only necessary screen to identify abnormality and to assess gastrointestinal tract symptoms. In patients with widespread spasticity (e.g., due to multiple sclerosis or myelopathy), inspection and auscultation should precede palpation and percussion. Manipulation of the abdominal wall often results in a wave of increased tone that will temporarily impede the rest of the abdominal examination. Vigorous abdominal palpation in patients with disordered peristalsis from certain central nervous system diseases may initiate regurgitation of stomach contents. Such patients should be examined gently when they are in a partially reclined position.

**Genitourinary System and Rectum**
During any comprehensive evaluation, the genitalia should be examined. A thorough evaluation of the male and female genitalia is particularly necessary for patients with disorders of continence, micturition, and sexual function. Incontinence in patients of either sex and in male patients using an external collecting device such as the condom catheter can result in maceration and ulceration. Thus, the penile skin in male patients, the periurethral mucosa in female patients, and all intertriginous perineal areas should be examined. The scrotal contents should be palpated for orchitis and epididymitis in male patients with indwelling catheters. Incontinence from neurogenic causes is common in patients undergoing rehabilitation; however, the examiner should check for a cystocele or other structural cause of incontinence that can be remediated. Patients with long-term use of indwelling catheters should be checked for external urethral meatal ulceration, and male patients should be checked for penile fistulas. If urinary retention is suspected, the physical examination should be followed by an in-and-out catheterization to measure residual urine.

The rehabilitation assessment is not complete without digital examination of the rectum and anus to check anal tone and perineal sensation. In any patient with suspected central nervous system, autonomic, or pelvic disease, the bulbocavernosus reflex should be evaluated to monitor sphincter tone by firmly compressing the glans of the penis or clitoris with one hand while inserting the index finger of the other hand into the anus. Sphincter tone is increased with many upper motor neuron lesions, whereas it is decreased or absent with lesions peripheral to the sacral cord (S2-4).

**Musculoskeletal System**
Disorders of the musculoskeletal system are a major portion of the pathologic conditions addressed by the rehabilitation
physician. The examiner must possess expert skills in the evaluation of all musculoskeletal components and should systematically assess the bone, joint, cartilage, ligament, tendon, and muscle in each body region. Accomplishing this task requires full familiarity with surface landmarks and underlying anatomical features.

The assignment of many examination components to the neurologic or musculoskeletal examination is arbitrary because the neuromusculoskeletal function is so integrated. Examination of the musculoskeletal system is divided into inspection, palpation, ROM assessment, joint stability assessment, and muscle strength testing.

**Inspection**

Musculoskeletal inspection should be performed for scoliosis, abnormal kyphosis, and lordosis; joint deformity, amputation, absence and asymmetry of body parts (leg-length discrepancy); soft-tissue swelling, mass, scar, and defect; and muscle fasciculations, atrophy, hypertrophy, and rupture. At times, the dysfunction may be subtle and decipherable only through careful observation. While proceeding with the examination, the physician should note any wary and tentative movements of the patient indicative of pain, any exaggerated and inconsistent conduct indicative of malingering, and any bizarre behavior indicative of conversion reaction.

**Palpation**

Localized abnormalities (e.g., areas of tenderness or deformity) identified through inspection and any body regions of concern to the patient should be palpated to ascertain their structural origins. For an abnormality, it is important to first determine whether its basic consistency is that of soft tissue or bone and whether it is of normal anatomical structure. An attempt should be made to further identify soft-tissue abnormalities as pitting or nonpitting edema, synovitis, or mass lesions.

All skeletal elements near areas of hemorrhage and ecchymosis in patients with altered consciousness should be palpated. The elderly patient with traumatic subdural hematoma may have an extremity fracture associated with a fall. During the critical care of a motorcyclist with a head injury, an incidental fracture may have been missed. Likewise, any in-hospital fall by a confused patient warrants a search for occult bony trauma.

**ROM Assessment**

Human joint motion is measured during clinical evaluation by many health care professionals for various reasons, including initial evaluation, evaluation of treatment procedure, feedback to the patient, assessment of work capacity, or research studies. When identifying a starting point for measuring the ROM of a joint, we prefer to regard the anatomical position as the baseline (zero starting point). If rotation is being measured, the midway point between the normal rotation range should be the zero starting point (7).

Considerable variation exists among the ROM measurements of different persons. Factors such as age, sex, conditioning, obesity, and genetics can influence the normal ROM. The American Academy of Orthopaedic Surgeons has reported the average ROM measurements for joints in the human body (8).

When the patient does not assist the examiner during an assessment, the measurement is a passive ROM. If the patient performs the ROM maneuver without assistance, then the range is an active ROM. If comparisons are made between active and passive ROMs, the starting position, stabilization, goniometer, alignment, and type of goniometer should be the same.

Different methods are available for recording the results of ROM measurements. Graphic recordings are often helpful for providing feedback to the patient or a third party. Sometimes the difference between the patient’s ROM and a normal ROM is of special interest to the examiner, such as when the surgeon wants to evaluate finger motion periodically as a guide to recovery after a hand operation.

The goniometer position, starting position, and average ROM of the more commonly measured joints are shown in Figures 1-1 through 1-26.

(Text continues on page 21)
FIGURE 1-2. Shoulder hyperextension. Starting position: Prone, arm at side, with hand pronated. Measurement: Sagittal plane: Avoid substitutions of lifting shoulder from table or rotating trunk. Goniometer: same as in Figure 1-1. (Courtesy of J.F. Lehmann, MD.)


FIGURE 1-4. Shoulder internal rotation. Starting position: Supine, arm abducted to 90° and elbow off table, elbow flexed to 90 degrees and hand pronated, forearm perpendicular to floor. Measurement: Transverse plane: Avoid substitutions of protracting shoulder, rotating trunk, changing angle at shoulder or elbow. Goniometer: Axis through longitudinal axis of humerus; shaft perpendicular to floor; shaft parallel to midline or forearm. (Courtesy of J.F. Lehmann, MD.)

FIGURE 1-5. Shoulder external rotation. Starting position: Same as in Figure 1-2. Measurement: Transverse plane: Avoid substitutions of arching back, rotating trunk, changing angle at shoulder or elbow. Goniometer: same as in Figure 1-4. (Courtesy of J.F. Lehmann, MD.)


FIGURE 1-8. Forearm pronation. Starting position: Sitting (or standing), arm at side with elbow held close to trunk, elbow bent to 90 degrees, forearm in neutral position between pronation and supination, wrist in neutral position, pencil held securely in midpalmar crease. Measurement: Transverse plane. Avoid substitutions of rotating trunk, moving arm, changing angle at elbow, angulating wrist. Goniometer: Axis through longitudinal axis of forearm; shaft parallel to midline of humerus; shaft parallel to pencil (on thumb side). (Courtesy of J.F. Lehmann, MD.)

FIGURE 1-9. Forearm supination. Starting position: Same as in Figure 1-8. Measurement: Same as in Figure 1-8. (Courtesy of J.F. Lehmann, MD.)

FIGURE 1-11. Wrist extension. Starting position: Same as in Figure 1-10. Measurement: Sagittal plane. Goniometer: Axis on ventral surface of wrist (in line with third metacarpal bone); shaft on mid-dorsal surface of forearm; shaft on midpalmar surface of hand. (Courtesy of J.F. Lehmann, MD.)


FIGURE 1-13. Wrist ulnar deviation. Starting position: Same as in Figure 1-12. Measurement: Same as in Figure 1-12. (Courtesy of J.F. Lehmann, MD.)
**FIGURE 1-14.** First metacarpophalangeal flexion. Starting position: Elbow slightly flexed, hand supinated, fingers and thumb extended. Measurement: Frontal plane. Goniometer: Axis on lateral aspect of metacarpophalangeal joint; shaft parallel to midline of first metacarpal bone; shaft parallel to midline of proximal phalanx. (Courtesy of J.F. Lehmann, MD.)

**Starting position**
- Elbow slightly flexed
- Hand supinated
- Fingers and thumb extended

**Measurement**
- Frontal plane
- Goniometer:
  - Axis on lateral aspect of metacarpophalangeal joint
  - Shaft parallel to midline of first metacarpal bone
  - Shaft parallel to midline of proximal phalanx

---


**Starting position**
- Elbow flexed
- Hand pronated
- Wrist in neutral position

**Measurement**
- Sagittal plane
- Goniometer:
  - Axis on mid-dorsum of joint
  - Shaft on mid-dorsum of metacarpal bone
  - Shaft on mid-dorsum of proximal phalanx
CHAPTER 1 | CLINICAL EVALUATION


FIGURE 1-18. Hip extension. See Figure 1-19. (Courtesy of J.F. Lehmann, MD.)


FIGURE 1-20. Hip flexion. Starting position: Lying on side or supine (may flex lower knee slightly for support). Measurement: Sagittal plane. Relocate greater trochanter and redraw C–D, as described in Figure 1-19. Goniometer placement is the same as in Figure 1-19. (Courtesy of J.F. Lehmann, MD.)
FIGURE 1-21. Hip abduction. (Courtesy of J.F. Lehmann, MD.)

FIGURE 1-22. Hip adduction. Starting position: Supine, leg extended and in neutral position. Measurement: Frontal plane: Mark both anterosuperior iliac spines and draw a line between them. Goniometer: Axis over hip joint; shaft parallel to line between spines of ilium; shaft along shaft of femur. (Courtesy of J.F. Lehmann, MD.)

FIGURE 1-23. Hip internal rotation (left) and hip external rotation (right). Starting position: Prone, sitting, or supine (indicate position on record), knee flexed to 90 degrees. Measurement: Transverse plane: Avoid substitutions of rotating trunk and lifting thigh from table. Goniometer: Axis through longitudinal axis of femur; shaft parallel to table; shaft parallel to lower part of leg. (Courtesy of J.F. Lehmann, MD.)
Joint Stability Assessment

Joint stability is the capacity of the structural elements of a joint to resist the forces of an inappropriate vector. It is determined by the degree of bony congruity, cartilaginous and capsular integrity, and ligament and muscle strength and by the forces applied to the joint. For example, the ball-and-socket arrangement of the hip joint is inherently stable because of bony congruity, whereas the glenohumeral joint must rely on musculoligamentous support because of the incongruity of the spherical humeral head in relation to the shallow curve of the glenoid fossa.

Joint stability is often compromised by disorders commonly treated by physical medicine and rehabilitation staff. For example, inflammatory synovitis associated with polyarthritis weakens the joint capsule and surrounding ligaments, and the resulting pain inhibits muscle contraction. This inhibition renders the involved joint susceptible to trauma from normal and abnormal forces and leads to joint instability. Similarly, traumatic and neurogenic conditions may commonly result in instability of peripheral and axial joints.

Excessive joint motion is often identified during the ROM assessment. However, several specialized physical examination maneuvers (e.g., the Larson test, the Lachman test, or the pivot shift test) can be used to assess individual joint integrity. Although a discussion of these tests is beyond the scope of this chapter, excellent texts are available on them (6,9,10).
The stability of each joint should be assessed in an orderly fashion. A routine series of individual joint maneuvers should be used as part of the general examination, and additional tests should be performed as necessary to identify more subtle instability when indicated by the history or general examination.

If joint instability is recognized or suspected during the physical examination, subsequent radiographic studies can often be helpful for quantifying the extent of instability. At times, flexion-extension views of the spine and stressed joint views of extremity joints can be informative; however, these should not be considered until the physical examination and nonstressed films have determined that such maneuvers are safe.

**Muscle Strength Testing**

Manual muscle testing provides an important means of assessing strength but also can be used to assess weakness. The examiner should keep in mind many factors that can affect the effort a patient is able to put into the testing. These factors include age, sex, pain, fatigue, low motivation, fear, misunderstanding of the test, and the presence of lower or upper motor neuron disease.

Lower motor neuron disease results in patterns of motor loss that depend on the location of the disease. For example, peripheral neuropathy results in a pattern of weakness in the muscles supplied by the affected nerve, whereas poliomyelitis results in residual weakness that is often scattered. The flaccid characteristic of a paretic muscle or muscle group in lower motor neuron disease allows the testing procedure to be uncomplicated by the spasticity or rigidity of upper motor neuron disease. Knowledge of the appearance of the muscle surface when a muscle undergoes atrophy from lower motor neuron disease also can be helpful to the clinician. If the joint crossed by the muscle being tested is unstable because of a chronic flaccid state, the grade of weakness may be difficult to estimate.

Upper motor neuron disease frequently results in spastic muscles that make manual testing challenging. For example, the antagonist muscle may be spastic and resist the action of the muscle being tested, or contractures may have developed that complicate the testing by limiting the available ROM.

Detailed discussions of the technique of manual muscle testing can be found in the publications of Kendall et al. (11) and Hislop and Montgomery (12). The anatomical basis for manual muscle testing of the major groups of muscles is discussed below (1).

**ANATOMICAL INFORMATION REQUIRED TO TEST INDIVIDUAL MUSCLE STRENGTH**

In the description of each test below, which is based on the format used in Mayo Clinic Examinations in Neurology (1), the name of each muscle is followed by the corresponding peripheral nerve and spinal segmental supply. Different authorities give considerable variability in segmental supply, particularly for certain muscles. Anatomical variation also exists both in the plexus and in the peripheral nerves.

Therefore, the segments listed here cannot be regarded as absolute. The "action" sections identify only the principal and important secondary or accessory functions that are useful in testing. The positions and the movements in each test refer first to the patient. When the movement is adequately indicated by the action of the muscle, it has been omitted here. Unless otherwise stated, "resistance" refers to pressure applied by the examiner in the opposite direction of the patient’s movement. For brevity and uniformity, we have given methods of testing that involve the patient initiating action against the resistance of the examiner, except when the other method is distinctly more applicable, but we do not mean to imply a preference for this method. We have often given the location of the belly of the muscle and its tendon so as to stress the importance of observation and palpation in identifying the muscle’s function. We have listed only those participating muscles with a definite action in the movement being tested that may substitute in part or in whole for the muscle being reviewed.

The following text has been adapted with permission from the Mayo Foundation for Medical Education and Research (1).

**Trapezius**

Spinal accessory nerve (Fig. 1-27).

**Action**

Elevation, retraction (adduction), and rotation (lateral angle upward) of the scapula, providing fixation of the scapula during movement of the arm.

---

CHAPTER 1  |  CLINICAL EVALUATION

Test
1. Elevation (shrugging) of the shoulder against resistance tests the upper portion, which is readily visible.
2. Bracing the shoulder (backward movement and adduction of the scapula) primarily tests the middle portion.
3. Abduction of the arm against resistance intensifies the winging of the scapula that may be present in paresis of the trapezius muscle (as in spinal accessory neuropathy).

In isolated trapezius palsy with the shoulder girdle at rest, the scapula is displaced downward laterally and rotated so that the superior angle is farther from the spine than the inferior angle. The lateral displacement is due in part to the unopposed action of the serratus anterior. The vertebral border, particularly at the inferior angle, is flared. These changes are accentuated when the arm is abducted from the side against resistance. On flexion (forward elevation) of the arm, however, this flaring of the inferior angle virtually disappears. These features are important in distinguishing paresis of the trapezius from that of the serratus anterior, because both conditions produce winging of the scapula.

Participating Muscles
1. Elevation: levator scapulae (third and fourth cervical nerves and dorsal scapular nerve, C3-5).
2. Retraction: rhomboids.

Rhomboids (Fig. 1-28)
Dorsal scapular nerve from the anterior ramus, C4, C5 (Fig. 1-29).

Action
Retraction (adduction) of the scapula and elevation of its vertebral border.

Test
The hand is on the hip; the arm is held backward and rotated medially. The examiner attempts to force the patient’s elbow laterally and forward while observing and palpating the muscle bellies medial to the scapula.

 Participating Muscles
Trapezius; levator scapulae with elevation of the medial border of the scapula.

Serratus Anterior (see Fig. 1-27)
Long thoracic nerve from the anterior rami, C5-7 (see Fig. 1-29). See Appendix A.

Action
1. Protraction (lateral and forward movement) of the scapula while it is kept close to the thorax.
2. Assistance in the upward rotation of the scapula.

**FIGURE 1-28.** Musculature of the shoulder, posterior view. (From Jenkins DB. Hollinshead’s Functional Anatomy of the Limbs and Back. 8th ed. Philadelphia, PA: WB Saunders; 2002:89. Used with permission of Mayo Foundation for Medical Education and Research.)

Test
The patient’s outstretched arm is thrust forward against a wall or against resistance provided by the examiner.

Isolated palsy results in comparatively little change in the appearance of the shoulder girdle at rest. However, there is slight winging of the inferior angle of the scapula and a slight shift medially toward the spine. When the outstretched arm is thrust forward, the entire scapula, particularly its inferior angle, shifts backward away from the thorax, producing the characteristic wing effect. Abduction of the arm laterally, however, produces comparatively little winging, compared with the manifestations of paralysis of the trapezius.

**Supraspinatus (Fig. 1-30)**
Suprascapular nerve from the upper trunk of the brachial plexus, C5, C6 (see Fig. 1-29). See Appendix B.

**Action**
Initiation of abduction of the arm from the side of the body.

**Test**
The above action is tested against resistance.

Atrophy may be detected just above the spine of the scapula, but the trapezius overlies the supraspinatus, and atrophy of either muscle will produce a depression in this area. Scapular fixation is important in this test.

**Participating Muscle**
Deltoid.

---

**Infraspinatus (Fig. 1-31)**
Suprascapular nerve from the upper trunk of the brachial plexus, C5, C6 (see Fig. 1-29). See Appendix B.

**Action**
Lateral (external) rotation of the arm at the shoulder.

**Test**
The elbow is at the side and flexed 90 degrees. The patient resists the examiner’s attempt to push the hand medially toward the abdomen.

The muscle is palpable, and atrophy may be visible below the spine of the scapula.

**Participating Muscles**
Teres minor (axillary nerve); deltoid (posterior fibers).

---

**Pectoralis Major (Fig. 1-32)**
See Figure 1-29 and Appendix A.

1. Clavicular portion (lateral pectoral nerve from the lateral cord of the plexus, C5-7).
2. Sternoternal portion (medial pectoral nerve from the medial cord of the plexus, lateral pectoral nerve, C6-8, T1).
Deltoid (Fig. 1-33C; see Fig. 1-32)
Axillary nerve from the posterior cord of the plexus, C5, C6 (see Fig. 1-29). See Appendix D.

Action
1. Abduction of the arm.
2. Flexion (forward movement) and medial rotation of the arm: anterior fibers.
3. Extension (backward movement) and lateral rotation of the arm: posterior fibers.

Test
1. The arm is almost horizontal in abduction. The patient resists the examiner’s efforts to depress the elbow. Paralysis of the deltoid leads to conspicuous atrophy and serious disability, because the other muscles that participate in abduction of the arm (the supraspinatus, trapezius, and serratus anterior—the last two by rotating the scapula) cannot compensate for the deltoid’s lack of function.
2. Flexion and extension of the arm are tested against resistance.

Participating Muscles
1. Abduction: given above.
2. Flexion: pectoralis major (clavicular portion); biceps.
3. Extension: latissimus dorsi; teres major.

Subscapularis (Fig. 1-33B)
Upper and lower subscapular nerves from the posterior cord of the plexus, C5-7 (see Fig. 1-29). See Appendix C.

Action
Medial (internal) rotation of the arm at the shoulder.
Test
The patient’s elbow is at the side and flexed 90 degrees. The patient resists the examiner’s attempt to pull the hand laterally. Because this muscle is not accessible to observation or palpation, it is necessary to gauge the activity of other muscles that produce this movement. The pectoralis major is the most powerful medial rotator of the arm; hence, paralysis of the subscapularis alone results in relatively little weakness of this movement.

Participating Muscles
Pectoralis major; deltoid (anterior fibers); teres major; latissimus dorsi.

**Biceps; Brachialis (Fig. 1-34)**
Musculocutaneous nerve from the lateral cord of the plexus, C5, C6 (see Fig. 1-29).

**Action**
1. Biceps: flexion and supination of the forearm and assistance in flexion of the arm at the shoulder.
2. Brachialis: flexion of the forearm at the elbow.

**Test**
Flexion of the forearm is tested against resistance. The forearm should be in supination to decrease participation of the brachioradialis.

**Triceps (Fig. 1-35)**
Radial nerve, which is a continuation of the posterior cord of the plexus, C6-8 (see Fig. 1-29). See Appendix C.
CHAPTER 1 | CLINICAL EVALUATION

**Brachioradialis (Fig. 1-36)**

Radial nerve, C5, C6 (see Fig. 1-29). See Appendix E.

**Action**

Flexion of the forearm at the elbow.

**Test**

Flexion of the forearm is tested against resistance, with the forearm midway between pronation and supination. The belly of the muscle stands out prominently on the upper surface of the forearm, tending to bridge the angle between the forearm and the arm.

**Supinator (see Fig. 1-36)**

Posterior interosseous nerve from the radial nerve, C5, C6 (see Fig. 1-29). See Appendix E.

**Action**

Supination of the forearm.

**Test**

The forearm is in full extension and supination. The patient tries to maintain supination while the examiner attempts to pronate the forearm and palpate the biceps.

Resistance to pronation by the intact supinator can usually be felt before there is appreciable contraction of the biceps.
Extensor Carpi Radialis Brevis (see Fig. 1-37)
Posterior interosseous nerve from the radial nerve, C6, C7 (see Fig. 1-29). See Appendix E.

Action
Extension (dorsiflexion) of the hand at the wrist.

Test
The forearm is in complete pronation. Dorsiflexion of the wrist is tested against resistance applied straight downward to the dorsum of the hand.

The tendon is palpable just proximal to the base of the third metacarpal bone. To minimize participation of the extensors of the digits, the patient should relax the fingers and thumb but flex them somewhat.

Extensor Carpi Ulnaris (see Fig. 1-37)
Posterior interosseous nerve from the radial nerve, C7, C8 (see Fig. 1-29). See Appendix E.

Action
Extension (dorsiflexion) and ulnar deviation of the hand at the wrist.

Test
The forearm is in pronation. Dorsiflexion and ulnar deviation of the wrist are tested against resistance applied to the dorsum of the hand downward and toward the radial side.

The tendon is palpable just below or above the distal end of the ulna. To minimize participation of the extensors of the digits, the patient should relax the fingers but flex them somewhat.

Extensor Digitorum Communis (see Fig. 1-37)
Posterior interosseous nerve from the radial nerve, C7, C8 (see Fig. 1-29). See Appendix E.

Action
1. Extension of the fingers, principally at the metacarpophalangeal joints.
2. Assistance in extension (dorsiflexion) of the wrist.

Test
The forearm is in pronation. The wrist is stabilized in a straight position. Extension of the fingers at the metacarpophalangeal joints is tested against resistance applied to the proximal phalanges.

The distal portions of the fingers may be somewhat relaxed but slightly flexed. The tendons are visible and palpable over the dorsum of the hand.

Extension at the interphalangeal joints is a function primarily of the interossei (ulnar nerve) and the lumbricals (median and ulnar nerves).

To minimize the action of the common extensor, the physiatrist can individually test the extensor digiti quinti and extensor indicis (posterior interosseous nerve, C7, C8), which

Extensor Carpi Radialis Longus (Fig. 1-37)
Radial nerve, C6, C7 (see Fig. 1-29). See Appendix E.

Action
Extension (dorsiflexion) and radial abduction of the hand at the wrist.

Test
The forearm is in almost complete pronation. Dorsiflexion of the wrist is tested against resistance applied to the dorsum of the hand downward and toward the ulnar side.

The tendon is palpable just above its insertion into the base of the second metacarpal bone. To minimize participation of the extensors of the digits, the patient should relax the fingers and thumb but flex them somewhat.

are proper extensors of the little finger and the index finger, respectively, while the other fingers are flexed. In a thin person's hand, these tendons can usually be identified.

**Abductor Pollicis Longus (see Fig. 1-36)**
Posterior interosseous nerve from the radial nerve, C7, C8. See Appendix E.

**Action**
1. Radial abduction of the thumb (in the same plane as that of the palm, in contradistinction to palmar abduction, which is movement perpendicular to the plane of the palm).
2. Assistance in radial abduction and flexion of the hand at the wrist.

**Test**
1. The hand is on the edge (the forearm is midway between pronation and supination).
2. Radial abduction of the thumb is tested against resistance applied to the metacarpal.

The tendon, which forms the anterior (volar) boundary of the “anatomic snuffbox,” is palpable just above its insertion into the base of the metacarpal bone.

**Participating Muscle**
Extensor pollicis brevis.

**Extensor Pollicis Brevis**
Posterior interosseous nerve from the radial nerve, C7, C8. See Appendix E.

**Action**
1. Extension of the proximal phalanx of the thumb.
2. Assistance in radial abduction and extension of the metacarpal of the thumb.

**Test**
The hand is on the edge. The wrist and the metacarpal of the thumb are stabilized by the examiner, with the thumb close to the palm at its radial border. Extension of the distal phalanx is tested against resistance.

If the patient is permitted to flex the wrist or abduct the thumb away from the palm, some extension of the phalanges results simply from lengthening the path of the extensor tendon. At the wrist, the tendon forms the posterior (dorsal) boundary of the anatomic snuffbox.

The characteristic result of radial nerve palsy is wrist-drop. Extension of the fingers at the interphalangeal joints is still possible by virtue of the action of the interossei and lumbricals, but extension of the thumb is lost.

**Pronator Teres (Fig. 1-38)**
Median nerve, C6, C7 (see Fig. 1-29). See Appendix F.

![Figure 1-38. Pronators of the forearm. (From Jenkins DB. Hollinshead's Functional Anatomy of the Limbs and Back. 8th ed. Philadelphia, PA: WB Saunders; 2002:172. Used with permission of Mayo Foundation for Medical Education and Research.)](image-url)
Action
Pronation of the forearm.

Test
The elbow is at the side of the trunk, the forearm is in flexion to the right angle, and the arm is in lateral rotation at the shoulder to eliminate the effect of gravity, which favors pronation in most positions. Pronation of the forearm is tested against resistance, starting from a position of moderate supination.

Participating Muscle
Pronator quadratus (anterior interosseous branch of the median nerve, C7, C8, T1).

Flexor Carpi Radialis (Figs. 1-38 and 1-39)
Median nerve, C6, C7 (see Fig. 1-29). See Appendix F.

Action
1. Flexion (palmar flexion) of the hand at the wrist.
2. Assistance in radial abduction of the hand.

Test
1. Flexion of the hand is tested against resistance applied to the palm.
2. The fingers should be relaxed to minimize participation of their flexors. The tendon is the more lateral (radial) of the two conspicuous tendons on the volar aspect of the wrist.

In complete median nerve palsy, flexion of the wrist is considerably weakened but can still be performed by the flexor carpi ulnaris (ulnar nerve), assisted to some extent by the abductor pollicis longus (radial nerve). In this event, ulnar deviation of the hand usually accompanies flexion.

Palmaris Longus (see Fig. 1-39)
Median nerve, C7, C8, T1 (see Fig. 1-29). See Appendix F.

Action
Flexion of the hand at the wrist.

Test
Same as that for flexor carpi radialis. The tendon is palpable at the ulnar side of the tendon of the flexor carpi radialis.

Flexor Carpi Ulnaris (see Fig. 1-39)
Ulnar nerve, C7, C8, T1 (see Fig. 1-29). See Appendix G.

Action
1. Flexion and ulnar deviation of the hand at the wrist.
2. Fixation of the pisiform bone during contraction of the abductor digiti quinti.

Test
Flexion and ulnar deviation of the hand are tested against resistance applied to the ulnar side of the palm in the direction of extension and radial abduction. The fingers should be relaxed. The tendon is palpable proximal to the pisiform bone.

Flexor Digitorum Superficialis (see Fig. 1-39)
Median nerve, C7, C8 (see Fig. 1-29). See Appendix F.

Action
1. Flexion primarily of the middle phalanges of the fingers at the first interphalangeal joints; flexion secondarily of the proximal phalanges at the metacarpophalangeal joints.
2. Assistance in flexion of the hand at the wrist.

Test
The wrist is in a neutral position; the proximal phalanges are stabilized. Flexion of the middle phalanx of each finger is tested against resistance applied to that phalanx, with the distal phalanx relaxed.

Flexor Digitorum Profundus
See Figure 1-39 and Appendices F and G.
1. Radial portion: usually to digits II and III (median nerve and its anterior interosseous branch C7, C8, T1).
2. Ulnar portion: usually to digits IV and V (ulnar nerve, C7, C8, T1).

**Action**
1. Flexion primarily of the distal phalanges of the fingers; flexion secondarily of other phalanges.
2. Assistance in flexion of the hand at the wrist.

**Test**
1. Flexion of the distal phalanges is tested against resistance, with the proximal and middle phalanges stabilized in extension.
2. With the middle and distal phalanges folded over the edge of the examiner’s hand, the patient resists the examiner’s attempt to extend the distal phalanges.

**Flexor Pollicis Longus (see Fig. 1-39)**
Anterior interosseous branch of the median nerve, C7, C8, T1. See Appendix F.

**Action**
1. Flexion of the thumb, particularly the distal phalanx.
2. Assistance in ulnar adduction of the thumb.

**Test**
Flexion of the distal phalanx is tested against resistance, with the thumb in the position of palmar adduction and with stabilization of the metacarpal and the proximal phalanx.

**Abductor Pollicis Brevis (Fig. 1-40)**
Median nerve, C8, T1 (see Fig. 1-29). See Appendix F.

**Action**
1. Palmar abduction of the thumb (perpendicular to the plane of the palm).
2. Assistance in opposition and in flexion of the proximal phalanx of the thumb.

**Test**
Palmar abduction of the thumb is tested against resistance applied at the metacarpophalangeal joint.

| The muscle is readily visible and palpable in the thenar eminence. |

**Participating Muscle**
Flexor pollicis brevis (superficial head).

**Opponens Pollicis (see Fig. 1-40)**
Median nerve, C8, T1 (see Fig. 1-29). See Appendix F.
**Action**
Movement of the first metacarpal across the palm, rotating it into opposition.

**Test**
The thumb is in opposition. The examiner attempts to rotate and draw the thumb back to its usual position.

**Participating Muscles**
Abductor pollicis brevis; flexor pollicis brevis.

**Flexor Pollicis Brevis (Fig. 1-40)**
Superficial head (median nerve, C8, T1); deep head (ulnar nerve, C8, T1) (see Fig. 1-29). See Appendix F.

**Action**
1. Flexion of the proximal phalanx of the thumb.
2. Assistance in opposition, ulnar adduction (entire muscle), and palmar abduction (superficial head) of the thumb.

**Test**
1. The thumb is in the position of palmar adduction, with stabilization of the metacarpal.
2. Flexion of the proximal phalanx is tested against resistance applied to that phalanx, while the distal phalanx is as relaxed as possible.

**Participating Muscles**
Flexor pollicis longus; abductor pollicis brevis; adductor pollicis.

Severe median nerve palsy produces the “simian” hand, wherein the thumb tends to lie in the same plane as the palm, with the volar surface facing more anteriorly than normal. Atrophy of the muscles of the thenar eminence is usually conspicuous.

Three muscles supplied, at least in part, by the ulnar nerve have already been described: the flexor carpi ulnaris, the flexor digitorum profundus, and the flexor pollicis brevis. The remaining muscles supplied by this nerve are described below.

**Hypothenar Muscles (see Fig. 1-40)**
Ulnar nerve, C8, T1 (see Fig. 1-29). See Appendix G.

**Action**
1. Abductor digiti minimi and flexor digiti minimi: abduction and flexion (proximal phalanx) of the little finger.
2. Opponens digiti minimi: opposition of the little finger toward the thumb.
3. All three muscles: palmar elevation of the head of the fifth metacarpal, helping to cup the palm.

**Test**
The action usually tested is abduction of the little finger (against resistance).

The abductor digiti minimi is readily observed and palpated at the ulnar border of the palm. Opposition of the thumb and the little finger can be tested together by gauging the force required to separate the tips of the two digits when opposed or by attempting to withdraw a piece of paper clasped between the tips of the digits.

**Interossei (Figs. 1-41 and 1-42)**
Ulnar nerve, C8, T1 (see Fig. 1-29). See Appendix G.

**Action**
1. Dorsal: abduction of the index, middle, and ring fingers from the middle line of the middle finger (double action on the middle finger: both radial and ulnar abduction; radial abduction of the index finger; ulnar abduction of the ring finger).
2. First dorsal: adduction (especially palmar adduction) of the thumb.
3. Palmar: adduction of the index, ring, and little fingers toward the middle finger.
4. Both sets: flexion of metacarpophalangeal joints and simultaneous extension of the interphalangeal joints.

**Test**
1. Abduction and adduction of the individual fingers are tested against resistance, with the fingers extended. Adduction can be tested by retention of a slip of paper between the fingers and between the thumb and the index finger, as the examiner attempts to withdraw it.
2. Ability of the patient to flex the proximal phalanges and simultaneously extend the distal phalanges.
3. Extension of the middle phalanges of the fingers against resistance while the examiner stabilizes the proximal phalanges in hyperextension.

The long extensors of the fingers (radial nerve) and the lumbrical muscles (median and ulnar nerves) assist in extension of the middle and distal phalanges. The first dorsal interosseus is readily observed and palpated in the space between the index finger and the thumb.

**Adductor Pollicis (see Fig. 1-42)**
Ulnar nerve, C8, T1 (see Fig. 1-29). See Appendix G.

**Action**
Adduction of the thumb in both the ulnar and palmar directions (in the plane of the palm and perpendicular to the palm, respectively). Assistance in flexion of the proximal phalanx.

**Test**
Adduction in each plane is tested against resistance by retention of a slip of paper between the thumb and the radial border of the hand and between the thumb and the palm, without flexion of the distal phalanx.

FIGURE 1-42. The chief adductors of the digits. (From Jenkins DB. Hollinshead’s Functional Anatomy of the Limbs and Back. 8th ed. Philadelphia, PA: WB Saunders; 2002:213. Used with permission of Mayo Foundation for Medical Education and Research.)
It is often possible to palpate the edge of the adductor pollicis just volar to the proximal part of the first dorsal interosseous.

**Participating Muscles**
1. Ulnar adduction: first dorsal interosseous; flexor pollicis longus; extensor pollicis longus; flexor pollicis brevis.
2. Palmar adduction: first dorsal interosseous, in particular; extensor pollicis longus.

**Flexors of the Neck**
Cervical nerves, C1-6.

**Test**
Sitting or supine: flexion of the neck, with the chin on the chest, is tested against resistance applied to the forehead.

**Extensors of the Neck**
Cervical nerves, C1-T1.

**Test**
Sitting or prone: extension of the neck is tested against resistance applied to the occiput.

**Diaphragm**
Phrenic nerves, C3-5.

**Action**
Abdominal respiration (inspiration) is distinguished from thoracic respiration (inspiration), which is produced principally by the intercostal muscles.

**Test**
1. The patient is observed for protrusion of the upper portion of the abdomen during deep inspiration when the thoracic cage is splinted.
2. The patient is observed for ability to sniff.
3. Litten phenomenon (successive retraction of the lower intercostal spaces during inspiration) may be observed if the patient’s body habitus permits, such as in thin individuals.
4. Diaphragmatic movements are observed fluoroscopically. Unilateral phrenic nerve palsy leads to diaphragmatic movement on one side but not the other (Litten sign).

Weakness of the diaphragm should be suspected in disease of the spinal cord when the deltoid or biceps is paralyzed, because these muscles are supplied by neurons situated close to those that innervate the diaphragm.

**Intercostal Muscles**
Intercostal nerves, T1-11.

**Action**
Expansion of the thorax anteroposteriorly and transversely, producing thoracic inspiration.

**Test**
1. Observation and palpation of the expansion of the thoracic cage during deep inspiration while maintaining pressure against the thorax.
2. Observation for asymmetric movement of the thorax, particularly during deep inspiration.
3. Other more general tests of function of the respiratory muscles are as follows:
   (a) Observation of the patient for rapid shallow respiration, flaring of ala nasi, and the use of accessory muscles of respiration.
   (b) Ability of the patient to repeat three or four numbers without pausing for breath.
   (c) Ability of the patient to hold his or her breath for 15 seconds.

**Anterior Abdominal Muscles**
Upper (T6-9); lower (T10-L1).

**Test**
1. Supine: flexion of the neck is tested against resistance applied to the forehead by the examiner. Contraction of the abdominal muscles can be observed and palpated. Upward movement of the umbilicus is associated with weakness of the lower abdominal muscles (Beevor sign).
2. Supine: hands on the occiput. Flexion of the trunk by anterior abdominal muscles is followed by flexion of the pelvis on the thighs by the hip flexors (chiefly iliopsoas) to reach a sitting position. The examiner holds down the patient’s legs.

Completion of this test can exclude significant weakness of either the abdominal muscles or the flexors of the hips. Weak abdominal muscles, in the presence of strong hip flexors, result in hyperextension of the lumbar spine during attempts to elevate the legs or to rise to a sitting position.

**Extensors of the Back**
See Figure 1-43.

**Test**
Prone: with the hands clasped over the buttocks, the patient elevates his or her head and shoulders off the table while the examiner holds down the patient’s legs.

The gluteal and hamstring muscles fix the pelvis on the thigh.

**Iliopsoas (Fig. 1-44)**
Psoas major (lumbar plexus, L2-4); iliacus (femoral nerve, L2-4). See Appendix H.

**Action**
Flexion of the thigh at the hip.

**Test**
1. Sitting: flexion of the thigh is tested by raising a knee against resistance by the examiner.
2. **Supine:** flexion of the thigh is tested by raising an extended leg off the table and maintaining it against downward pressure applied by the examiner just above the knee.

**Participating Muscles**
Rectus femoris and sartorius (both: femoral nerve, L2-4); tensor fasciae latae (superior gluteal nerve, L4, L5).

**Adductor Magnus, Longus, Brevis (see Fig. 1-44)**
Obturator nerve, L2-4; part of adductor magnus is supplied by sciatic nerve, L5, and functions with hamstrings. See Appendix I.

**Action**
Adduction of the thigh principally.

**Test**
Sitting or supine: the knees are held together while the examiner attempts to separate them.

The legs also can be tested separately and the muscles palpated.

---

**Participating Muscles**
Gluteus maximus; gracilis (obturator nerve, L2-4).

**Abductors of the Thigh (Fig. 1-45)**
Superior gluteal nerve, L4, L5, S1.

1. Gluteus medius and gluteus minimus principally.
2. Tensor fasciae latae to a lesser extent.

**Action**
1. Abduction and medial rotation of the thigh.
2. Tensor fasciae latae assists in flexion of the thigh at the hip.
PART I | PRINCIPLES OF ASSESSMENT AND EVALUATION

36

Test
1. Sitting: the knees are separated against resistance by the examiner. In this position, the gluteus maximus and some of the other lateral rotators of the thigh function as abductors, hence diminishing the accuracy of the test.
2. Supine: same test as for abductors, above, but more exact.
3. Lying on opposite side: the hip is abducted (moved upward) while the examiner presses downward on the lower leg and stabilizes the pelvis.

Medial Rotators of the Thigh (see Fig. 1-45)
Same as abductors; superior gluteal nerve, L4, L5, S1.

Test
Sitting or supine: the knee is flexed to 90 degrees. Medial rotation of the thigh is tested against resistance applied by the examiner at the knee and the ankle in an attempt to rotate the thigh laterally.

Lateral Rotators of the Thigh (Fig. 1-46)
L4, L5, S1, S2.
1. Gluteus maximus (inferior gluteal nerve, L5, S1, S2) chiefly.
2. Obturator internus and gemellus superior (nerve to obturator internus, L5, S1, S2).
3. Quadratus femoris and gemellus inferior (nerve to quadratus femoris, L4, L5, S1).

Test
Sitting or supine: the knee is flexed to 90 degrees. Lateral rotation of the thigh is tested against an attempt by the examiner to rotate the thigh medially.

The gluteus maximus is the muscle principally tested, and it can be observed and palpated in the prone position.

Gluteus Maximus (see Fig. 1-46)
Inferior gluteal nerve, L5, S1, S2.

Action
1. Extension of the thigh at the hip.
2. Lateral rotation of the thigh.
3. Assistance in adduction of the thigh.

The tensor fasciae latae and, to a lesser extent, the gluteus medius can be palpated.
CHAPTER 1 | CLINICAL EVALUATION

Test
1. Sitting or supine: starting with the thigh slightly raised, extension (downward movement) of the thigh is tested against resistance applied by the examiner under the distal part of the thigh. In this rather crude test, the muscle cannot be observed or readily palpated.
2. Prone: the knee is well flexed to minimize the participation of the hamstrings. Extension of the thigh is tested by raising the knee from the table against downward pressure applied by the examiner to the distal part of the thigh. The muscle is accessible to observation and palpation in this position.

Quadriceps Femoris (Fig. 1-47)
Femoral nerve, L2-4. See Appendix H.

Action
1. Extension of the leg at the knee.
2. Rectus femoris assists in flexion of the thigh at the hip.

Test
1. Sitting or supine: the lower leg is in moderate extension.
2. Maintenance of extension is tested against effort by the examiner to flex the patient’s leg at the knee.
   
   Atrophy is easily noted.

Hamstrings (Fig. 1-48)
Sciatic nerve, L4, L5, S1, S2. See Appendix J. Biceps femoris: external hamstring (L5, S1, S2). Semitendinosus and semimembranosus: internal hamstrings (L4, L5, S1, S2).

Action
1. Flexion of the leg at the knee.
2. All but the short head of the biceps femoris assist in extension of the thigh at the hip.

Test
1. Sitting: flexion of the lower leg is tested against resistance.
2. Prone: the knee is partly flexed. Further flexion is tested against resistance.

   Observation and palpation of the muscles and the tendons are important for proper interpretation.

Anterior Tibialis (Figs. 1-49 and 1-50)
Deep peroneal nerve, L4, L5, S1. See Appendix K.
Action
Dorsiflexion and inversion (particularly in the dorsiflexed position) of the foot.

Test
Dorsiflexion of the foot is tested against resistance applied to the dorsum of the foot downward and toward eversion.

The belly of the muscle just lateral to the shin and the tendon medially on the dorsal aspect of the ankle should be observed and palpated to be certain that dorsiflexion is not being accomplished by the extensor digitorum longus without contraction of the anterior tibialis. Atrophy is conspicuous.

Participating Muscles
1. Dorsiflexion: extensor hallucis longus; extensor digitorum longus.
2. Inversion: posterior tibialis.

Extensor Hallucis Longus (see Figs. 1-49 and 1-50)
Deep peroneal nerve, L5, S1. See Appendix K.

Action
Extension of the great toe and dorsiflexion of the foot.

Test
Extension of the great toe is tested against resistance, while the foot is stabilized in a neutral position.

The tendon is palpable between the tendons of the anterior tibialis and the extensor digitorum longus.

Extensor Digitorum Longus (see Figs. 1-49 and 1-50)
Deep peroneal nerve, L4, L5, S1. See Appendix K.

Action
Extension of the lateral four toes and dorsiflexion of the foot.
Test
Extension of the lateral four toes and dorsiflexion of the foot are tested against resistance.

The tendons are visible and palpable on the dorsal aspect of the ankle and the foot lateral to the tendon of the extensor hallucis longus.

**Extensor Digitorum Brevis (see Fig. 1-49)**
Deep peroneal nerve, L4, L5, S1. See Appendix K.

**Action**
Assists in the extension of all the toes except the little toe.

Test
The belly of the muscle is observed and palpated on the lateral aspect of the dorsum of the foot during toe extension.

**Peroneus Longus, Brevis (Fibularis Longus, Brevis) (Fig. 1-51)**
Superficial peroneal nerve, L5, S1. See Appendix L.

**Action**
1. Eversion of the foot.
2. Assistance in plantar flexion of the foot.

Test
The foot is in plantar flexion. Eversion is tested against resistance applied by the examiner to the lateral border of the foot.

The tendons are palpable just above and behind the external malleolus. Atrophy may be visible over the anterolateral aspect of the lower extremity.

**Gastrocnemius; Soleus (Fig. 1-52)**
Tibial nerve, L5, S1, S2. See Appendix J.

**Action**
1. Plantar flexion of the foot.
2. The gastrocnemius also flexes the knee and cannot act effectively to plantar flex the foot when the knee is well flexed.

Test
1. The knee is extended to test both muscles. The knee is flexed principally to test the soleus.
2. Plantar flexion of the foot is tested against resistance.

---


The muscles and tendon should be observed and palpated. Atrophy is readily visible. The gastrocnemius and soleus are strong muscles, and leverage in testing favors the patient rather than the examiner. For this reason, slight weakness is difficult to detect by resisting flexion of the ankle or by pressing against the flexed foot in the direction of extension. Consequently, the strength of these muscles should be tested against the weight of the patient’s body. The patient should stand on one foot and plantar flex the foot to lift himself or herself directly and fully upward. It may be necessary for the examiner to hold the patient steady while this test is performed.

**Participating Muscles**
Long flexors of the toes; posterior tibialis, fibularis longus, and fibularis brevis (particularly near the extreme plantar flexion).

**Posterior Tibialis (Fig. 1-53)**
Posterior tibial nerve, L5, S1. See Appendix J.

**Action**
1. Inversion of the foot.
2. Assistance in plantar flexion of the foot.

**Test**
The foot is in complete plantar flexion. Inversion is tested against resistance applied to the medial border of the foot and directed toward eversion and slightly toward dorsiflexion.

This maneuver virtually eliminates participation of the anterior tibialis in inversion. The toes should be relaxed to prevent participation of the long flexors of the toes.

**Long Flexors of the Toes (see Fig. 1-53)**
Posterior tibial nerve, L5, S1. See Appendix J.
1. Flexor digitorum longus.
2. Flexor hallucis longus.

**Action**
1. Plantar flexion of the toes, especially at the distal interphalangeal joints.
2. Assistance in plantar flexion and inversion of the foot.

**Test**
1. The foot is stabilized in a neutral position. Plantar flexion of the toes is tested against resistance applied to the distal phalanges.
2. Weakness of the long toe flexors results in inability to curl the tips of the toes under the foot against resistance. (See intrinsic foot muscle testing below.)

**Intrinsic Muscles of the Foot**
Virtually all muscles except the extensor digitorum brevis (medial and lateral plantar nerves from the posterior tibial nerve, L5, S1, S2).

**Action**
The flexion of the proximal phalanges during extension of the distal phalanges is an action comparable to that of the intrinsic muscles of the hand.

**Test**
The patient’s foot is stabilized in a neutral position, and plantar flexion of the toes is tested against resistance applied to the distal phalanges. (Same maneuver as that in the test of the long toe flexors.)

**Neurologic Examination**
With the exception of the musculoskeletal examination, no other component of the standard physical examination is more important to the physical medicine and rehabilitation assessment than the neurologic examination. Although often conducted to identify disease, the neurologic examination provides the physiatrist with an opportunity to identify both the neurologic impairments to be addressed and the residual abilities to be used in maximizing the functional outcome for the patient.

Although it is customary to record the results of the neurologic examination in a separate portion of the examination report, the neurologic examination is rarely performed all at one time. The examiner often finds it convenient to integrate the appropriate
portions of the neurologic examination into the assessment of a specific region of the body. For example, cranial nerve assessment often is performed with other components of the head and neck examination, because the patient is positioned appropriately for both. For purposes of discussion, the neurologic examination is addressed separately and is divided into assessments of mental status, speech and language function, cranial nerves, reflexes, central motor integration, sensation, and perception. Muscle strength is discussed in the section on examination of the musculoskeletal system. The assessment of complex motor activities is discussed in the section “Functional Examination.” The reader is referred to *Mayo Clinic Examinations in Neurology* for a comprehensive discussion of the neurologic evaluation (1).

**Mental Status**

**Level of Consciousness**

Before performing a formal mental status examination, the examiner should determine the patient’s level of consciousness. Qualitative terms such as “drowsy,” “lethargic,” and “stuporous” are useful in a descriptive sense, but they suffer from a lack of precise definition. “Stuporous” to one examiner may mean “lethargic” to another. A definitive classification of mental status requires a standardized approach (13–15). In the Glasgow Coma Scale, the examiner classifies the patient’s eye, motor, and verbal responses to verbal and physical stimuli according to a numerical scale that is quantifiable and reproducible (Table 1-4) (16). Such a standardized scale is necessary to assess changes over time and to facilitate communication among physicians, nurses, therapists, and family members. In patients with traumatic brain injury, other aspects of the neurologic assessment, such as pupillary responses, ocular movements, and respiration, will provide information about the cause of altered consciousness but do not quantifiably relate in a statistical sense to eventual outcome.

**Cognitive Evaluation**

With the conscious patient, assessment of mental status begins when the physician enters the room and continues throughout the examination. However, as with the assessment of the level of consciousness, a formal approach to the mental status examination can help the examiner to identify and quantify specific impairments and residual capacity, to recognize subtle temporal changes, and to facilitate communication among caregivers. A commonly used clinical tool for evaluating the patient’s mental status is the Folstein Mini-Mental Status Examination (13). Excellent systems have been developed to assess intellectual performance in specific populations (10,16). Although systems may or may not include perceptual testing, speech and language assessment, or an inventory of thought processing, certain components of the evaluation remain constant.

**Orientation**

The patient is asked to report his or her name, address, and telephone number and the building (e.g., hospital or clinic), city, state, year, month, and day.

**Attention**

Attention is assessed with digit repetition; the patient is asked to repeat a series of random numbers. Two numbers are used initially (e.g., 4 and 9); if the patient answers correctly, the sequence is increased by one more digit with each additional repetition until the patient either repeats seven digits correctly or makes a mistake. The number of digits repeated correctly should be noted.

**Recall**

Three numbers or three objects are listed, and the patient is asked to remember them for repetition later. In 5 minutes, the patient is asked to recall the list, and the number of correct responses is recorded. If all responses are correct, recall responses are obtained again at 10 minutes and at 15 minutes.

**General Fund of Information**

Questions are asked appropriate to the patient’s age, cultural interests, and educational background. For example, the names of the past five US presidents or leaders of other countries, the current US vice president, and the governor of the patient’s home state can be requested, or inquiries can be made for information about current events and other nearly universal subjects (e.g., world wars and basic scientific principles).

**Calculations**

The patient is asked to count by sevens, and the last correct response is recorded. Arithmetic calculations of increasing difficulty are presented.

---

**TABLE 1.4 Glasgow Coma Scale**

<table>
<thead>
<tr>
<th>Response</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eye opening, E</td>
<td>E 4</td>
</tr>
<tr>
<td>Spontaneous</td>
<td>E 3</td>
</tr>
<tr>
<td>To speech</td>
<td>E 2</td>
</tr>
<tr>
<td>To pain</td>
<td>E 1</td>
</tr>
<tr>
<td>Nil</td>
<td>E 1</td>
</tr>
<tr>
<td>Best motor response, M</td>
<td>M 6</td>
</tr>
<tr>
<td>Obey</td>
<td>M 5</td>
</tr>
<tr>
<td>Localizes</td>
<td>M 4</td>
</tr>
<tr>
<td>Withdraws</td>
<td>M 3</td>
</tr>
<tr>
<td>Abnormal flexion</td>
<td>M 2</td>
</tr>
<tr>
<td>Extensor response</td>
<td>M 1</td>
</tr>
<tr>
<td>Nil</td>
<td>M 1</td>
</tr>
<tr>
<td>Verbal response, V</td>
<td>V 5</td>
</tr>
<tr>
<td>Orientated</td>
<td>V 4</td>
</tr>
<tr>
<td>Confused conversation</td>
<td>V 3</td>
</tr>
<tr>
<td>Inappropriate words</td>
<td>V 2</td>
</tr>
<tr>
<td>Incomprehensible sounds</td>
<td>V 1</td>
</tr>
<tr>
<td>Nil</td>
<td>V 1</td>
</tr>
</tbody>
</table>

*Coma score (E + M + V) = 3 to 15.

Proverbs
An explanation of three common proverbs is requested. The patient is assessed as to whether he or she can abstract the principle from the adage and explain it in concrete terms.

Similarities
The patient is asked to describe what is common to an orange and an apple, to a desk and a bookcase, and to a cup and a fork. The number of correct responses is recorded.

Judgment
The patient is presented with three problems (e.g., smelling smoke in a movie theater, finding a stamped and addressed envelope on the sidewalk, and locating a friend in an unfamiliar city) and asked how to handle each situation.

Speech and Language Function
As with the assessment of mental status, the analysis of communicative function occurs throughout the entire examination. The patient should be evaluated for the presence and extent of aphasia, apraxia, and dysarthria and for any residual communicative skills (17,18). At times, effort is required to discriminate among the disorders of aphasia, apraxia of speech, and language dysfunction associated with a more generalized cognitive deficit. Expert assessment of speech production and language processing can be valuable for diagnosis of neurologic disease (see Chapter 15). However, as described in the preceding section, assessment of the four basic elements of communication (i.e., listening, reading, speaking, and writing) provides a practical framework for functional evaluation.

Listening
After first determining that the patient does not have a significant hearing loss, had been able to speak the examiner’s language before the onset of disease, and has the requisite motor and visual skills, the physician should test the patient’s auditory comprehension, noting the extent of his or her ability to follow specific directions without gestures from the examiner. Often, it is useful to characterize the degree of impairment with stepped commands. First, the patient’s ability to follow one-step commands is assessed by asking him or her to perform three different single motor activities, such as “take off your glasses,” “touch your nose,” and “open the book.” Each command should be given separately, and a prolonged pause should be allowed to observe the response. These responses are rated and notation is made of whether the patient requires pantomime of the activity before performing the task. If two of the three responses are correct, an assessment should be made of the patient’s skill at following two-step commands, such as “touch your nose, then take off your glasses,” “point to the window, then close the book,” and “touch my hand, then touch your knee.” If the patient can follow two-step commands, then an assessment of his or her ability to follow three-step commands is conducted in a similar fashion. A simple object such as a toothbrush is held up, and the patient is asked to demonstrate its use. This request is repeated at least two more times with different objects. If speech is functional, the patient is asked to repeat a short phrase that is spoken. The response should be observed for perseveration and jargon.

Reading
It is important to verify that the patient had reading skills before the onset of the neurologic disorder. The patient should be asked to read a short written command and perform the activity; the patient also can be asked to follow written two-step and three-step commands. If writing is otherwise functional, the patient should be asked to read aloud what he or she has written.

Speaking
If auditory comprehension is adequate, language production can be tested in several ways. An object is indicated, and the patient is asked to name it and state its function; at least three objects are used. The patient is asked to report his or her name, hometown, telephone number, or another simple verifiable fact. A picture can be shown and the patient asked to describe it. Tests for phonation and resonance deficits are performed by asking the patient to say a prolonged “aaah” and by observing for force and steadiness of pitch and tone. The patient should be asked to say “pa-pa-pa” to test lip closure, “ta-ta-ta” to test tongue function, and “ka-ka-ka” to test speed, regulatory, and posterior pharyngeal function. If reading is otherwise functional, articulation can be further assessed by having the patient read aloud a short passage containing various vowels and consonants.

Writing
The patient should be asked to write his or her name, address, telephone number, and a brief paragraph.

Cranial Nerves
Cranial Nerve I (Olfactory)
Olfactory function should be evaluated routinely. Deficits are common after head trauma.

Cranial Nerve II (Optic)
Visual field testing of each eye should be performed, with a temporary patch over the contralateral eye. It is best to test each quadrant diagonally to identify any quadrantanopia. Although visual double simultaneous stimulation may be more correctly classified under cortical sensation, it is convenient to assess for extinction during visual field testing, after full fields have been verified. Visual acuity is discussed in the section on “Eye Examination.”

Cranial Nerves III (Oculomotor), IV (Trochlear), and VI (Abducens)
Visual pathways are assessed by evaluating the pupil size, pupillary reactions, and extraocular movements. Strabismus is evaluated by testing corneal light reflections.
**Cranial Nerve V (Trigeminal)**
The muscles of mastication and facial sensation should be tested.

**Cranial Nerve VIII (Vestibulocochlear)**
The patient should be examined for nystagmus. Auditory function is discussed in the section on “Ear Examination.”

**Cranial Nerves VII (Facial), IX (Glossopharyngeal), X (Vagus), and XII (Hypoglossal)**
Isolating individual cranial nerve function emanating from the lower part of the brainstem is difficult. Cranial nerves are often grouped by function. Evaluation should be conducted of taste (nerves VII, IX, and X), muscles of facial expression (nerve VII) and articulation (nerves VII, IX, X, and XII), and swallowing function (nerves IX, X, and XII).

**Cranial Nerve XI (Accessory)**
The function of sternocleidomastoid and trapezius is frequently assessed during manual muscle testing.

Brainstem- and visual-evoked responses, electromyography and other forms of electrodiagnostic testing, and swallowing videofluoroscopy are often necessary for better delineation of dysfunction of cranial nerves and their brain-stem interactions.

**Reflexes**

**Muscle Stretch Reflexes**
Muscle stretch reflexes should be tested when the patient is relaxed. The commonly tested muscles are the biceps (C5, C6), triceps (C6-8), brachioradialis (C5, C6), quadriceps (L2-4), and gastrocnemius-soleus (L5, S1, S2). The reflexes of the masseter (cranial nerve V), internal hamstring (L4, L5, S1, S2), and external hamstring (L5, S1, S2) are tested in select cases. The patient should be observed for clonus.

**Superficial Reflexes**
Segmental reflexes are often helpful for localizing the lesion. These include the corneal (cranial nerves V and VII), gag (cranial nerves IX and X), anal (S3-5), and plantar (L5, S1, S2) reflexes. At times, it is useful to include the epigastric (T6-9), midabdominal (T9-11), hypogastric (T11, T12, L1), and cremasteric (L1, L2) reflexes.

**Pathologic Reflexes**
Elicitation of the Babinski reflex should be attempted. In questionable cases, the confirmatory Chaddock, Oppenheim, and Stransky reflexes should be tested.

**Central Motor Integration**

**Muscle Tone**
Spasticity, rigidity, and hypotonicity can be assessed by evaluating the patient’s resistance to passive movement, pendulousness, and ability to posturally fixate.

**Coordination**
Coordination in the upper extremities can be assessed with the finger-nose, finger-nose-finger, and knee-pat tests. Coordination in the lower extremities can be evaluated with the toe-finger and heel-knee-shin test.

**Alternate Motion Rate**
The tongue-wiggle, finger-wiggle, and foot-pat tests can be used to identify subtle spasticity, rigidity, and incoordination.

**Involuntary Movements**
The patient should be observed for tremors, chorea, athetosis, ballismus, dystonia, myoclonus, asterixis, and tics. If present, these should be described in the neurologic report.

**Apraxia**
Apraxia is the failure of motor planning and execution without deficits in strength, coordination, or sensation; however, deficits of strength, coordination, and sensation are often also present because of the extent of the lesion. Automatic motor activities can be observed while the patient manipulates a pen or pencil, handles clothing, and moves about the examination room; then, the patient’s ability to perform some of the same maneuvers on command can be assessed. The patient should be asked to touch his or her nose, drink from a glass, put a pencil in the glass, and use scissors. The patient should then be asked to perform these activities without the objects with each hand. Inefficient or fumbling movements or inability to accomplish the task should be noted. Dressing apraxia can be assessed by asking the patient to put on a coat. To assess for more subtle deficits, the examiner should first turn one sleeve of the coat inside out. Constructional apraxia can be evaluated by asking the patient to copy a geometric design or draw the face of a clock.

**Sensation**

**Superficial Sensation**
Light touch can be tested with a wisp of cotton, superficial pain with a single-use pin, and temperature with two test tubes, one with hot tap water and the other with cold tap water. Abnormal findings should be recorded on a drawing of the human figure and compared with standard charts of spinal dermatomes and peripheral nerves (1).

**Deep Sensation**
The evaluation of joint position sense begins with the distal joints of the hand and foot and moves proximally until normal sensation is identified. Testing for deep pain in the upper extremities can be done by hyperextension of small finger joints and in the lower extremities by firm compression of the calf muscles or Achilles tendon. Vibration sense is often evaluated, but its isolated absence does not result in functional deficit.

**Cortical Sensation**
If superficial and deep sensations are intact, two-point discrimination, graphesthesia, stereognosis, and double simultaneous stimulation can be evaluated.
Perception
Disorders of perception are most common with lesions of the nondominant parietal lobe but also can occur with lesions on the dominant side.

Agnosia
Agnosia is the failure to recognize familiar objects despite intact vision, hearing, sensation, and language function (although language is also often deficient because of the extent of the lesion). Pictures of common objects or the objects themselves are shown, and the patient is asked to identify them and describe their components. Agnosia of body parts can be assessed by asking the patient to identify his or her (or the examiner’s) arm, finger, or eye. Unilateral environmental neglect can be assessed by observing ambulation or wheelchair operation for difficulty clearing corners and doorjambs, extinction on double simultaneous stimulation, and failure to scan the complete page width when asked to read a passage or cross out all the occurrences of the letter E. Body scheme agnosias can be evaluated by searching for denial of obvious physical impairments when the patient is asked to describe them.

Right-Left Disorientation
If agnosia of body parts is not present, the patient should be asked to indicate various body parts on the right and the left sides.

Other Perceptual Tests
If perceptual deficits are identified with the maneuvers described above, the examiner should test for additional deficits, such as impaired geographic and spatial orientation and figure-ground relationships. Comprehensive, formal, and quantitative testing of perception by a psychologist and an occupational therapist is warranted if any deficits are found during the physical examination.

Functional Examination
After impairments have been identified, the consequences of each impairment for the function of the patient must be appraised. Prediction of functional status should not be attempted from the history and physical examination; instead, function should be examined. For a comprehensive assessment, the patient must be evaluated by individual physical medicine and rehabilitation team members in settings where the activities are actually performed. Bathing skills should be observed by the occupational therapist or the rehabilitation nurse in the bathroom while the patient attempts to bathe; eating skills should be analyzed by the occupational therapist while the patient eats a meal; and car transfer skills should be assessed by the physical therapist with the use of the patient’s car. Each team member will use unique skills to contribute to a comprehensive determination of functional status. Many functional evaluative processes cannot be accomplished at a single point. Safety and judgment can be assessed only by observing the patient in varying situations within both the rehabilitation environment and the community.

However, in many instances, the physiatrist must glean a basic view of the functional status at the time of the initial evaluation. For instance, in the clinic, the physician may be consulted to determine a patient’s need for rehabilitation services. It is unlikely that the physician will be able to observe the patient during a meal, in the bathroom, or in the process of transferring to and from a car. In such cases, the physician must use creativity to place the patient in situations similar to those of daily life. Examples are given below. Components of the communication assessment were discussed in the sections “History and Physical Examination” and will not be repeated here.

Eating
The patient should be asked to use examining equipment in place of feeding utensils to demonstrate proficiency in bringing food to the mouth. If aspiration has not already been identified, the patient should be provided with a glass of water and asked to drink.

Grooming
The patient should be asked to comb his or her hair and to mimic the activities of brushing teeth or putting on makeup.

Bathing
The patient should be asked to mimic the activities of bathing. It is important to note if any body parts cannot be reached by the patient, particularly the back, the scalp, and the axilla and arm contralateral to hemiparesis.

Toileting
The patient must have adequate unsupported sitting balance, must have the requisite wrist and hand motion to reach the perineum adequately, must be able to handle toilet paper, and must be able to rise from low seating.

Dressing
The patient should be observed during undressing before the examination and dressing after completion of the examination. The examiner should explain the purpose of the observation and should be accompanied by a nurse or aide.

Bed Activities
During the physical examination, the examiner should note whether the patient has difficulty moving between the seated and the supine positions. It should also be determined whether the patient can roll from front to back and back to front and whether the patient can raise the pelvis off the examining table while supine.

Transfers
The patient should be observed rising from seating with and without armrests and moving between the bed and a chair.
TABLE 1.5  Gait Analysis

Standing balance
Observe for steadiness of position; push the patient off balance and note the patient’s attempts to regain balanced posture.

Individual body part movements during walking
Observe for fixed or abnormal posture and inadequate, excessive, or asymmetrical movement of body parts.
Head and trunk: listing or tilting, shoulder dipping, elevation, depression, protraction, and retraction.
Arm swing: protecting positioning or posturing.
Pelvis and hip: hip hiking, dropping (Trendelenburg), or lateral thrust.
Knee: genu valgum, varum, or recurvatum.
Foot and ankle: excessive inversion or eversion.

Gait cycle factors
Cadence: rate, symmetry, fluidity, and consistency.
Stride width: narrow or broad based; knee and ankle clearance.
Stride length: shortened, lengthened, or asymmetrical.
Stance phase: initial contact, loading response, toe off; knee stability during all components of stance; coordination of knee and ankle movements.
Swing phase: adequate and synchronized knee flexion and ankle dorsiflexion during swing, abduction, or circumduction.

Wheelchair Mobility
The patient should be asked to demonstrate wheeling straight ahead and turning, on both carpeted and noncarpeted floors, if available; locking the brakes; and manipulating the leg rests.

Ambulation
To adequately recognize disturbances of gait, the examiner must be able to view body parts. If the examining room is secluded, the assessment can be performed with the patient wearing only the underwear. If privacy is not possible, the patient should have

FUNCTIONAL STATUS

<table>
<thead>
<tr>
<th>NAME</th>
<th>John Doe</th>
<th>3/4/18</th>
<th>Hemi Paralysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACTIVITY</td>
<td>Independent</td>
<td>Independent with aids</td>
<td>Requires Assistance</td>
</tr>
<tr>
<td>Listening</td>
<td>Audible</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reading</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speaking</td>
<td>Aerithoria</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Writing</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eating</td>
<td>Set up meal, needs no utensils, feeds self, simple plate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grooming</td>
<td>Washes hair, body, washes teeth, uses own toothbrush,application of dental cream, changes nails, can dress, completed dressing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bathing</td>
<td>Washes face, body, washes teeth</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Toileting</td>
<td>bladder with bowel</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dressing</td>
<td>Needs assistance for transfer to bathroom, can dress</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bed Activities</td>
<td>Hospital bed with bed rails</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transfers</td>
<td>Vertical lift, lock brakes, uses own judgment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wheel Chair</td>
<td>Vertical lift to transport, uses own brakes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ambulation</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Driving</td>
<td>Est. completed</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

If the activity is independent or dependent, mark with a check. If the activity is independent with aids, list the aids needed. If the activity requires assistance, describe the assistance and list the aids needed.

FIGURE 1-54. Sample of a functional status record.
access to washable or disposable shorts. If the examiner does not already have knowledge of the patient’s ambulation skills, the patient should be provided with a safety belt before gait is assessed. To discern specific gait abnormalities, the examiner must study both the individual components and the composite activity. The patient should be observed from the front, back, and each side. If the patient experiences pain during ambulation, its temporal relationship to the gait cycle should be noted. This analysis must be approached in an orderly fashion. One routine for gait analysis is outlined in Table 1-5 (4,19). See Chapter 10 for a comprehensive discussion of gait.

**Operation of a Motor Vehicle**

Driving ability can be assessed best in an automobile. However, the examiner can gain some information about the patient’s motor abilities for driving by asking the patient to demonstrate the motions of operating the pedals and hand controls.

**Quantitation of Function**

Several scales can be used to document and quantify functional status in activities of daily living. These are extremely useful in assessing a patient's rehabilitation progress (see Chapter 28). When validated and standardized, these scales are essential tools for analysis of rehabilitation outcome for a series of patients participating in a specific intervention program. When they are used by multiple rehabilitation centers to share data, relevant information can be obtained to advance the field and to assess the cost versus the benefit of rehabilitation. Physicians should develop expertise in the use of these valuable tools.

However, the data collection for most validated functional scales requires additional time because they require interdisciplinary input; therefore, the initial documentation of functional status by the physician must be practical and complete. One such system is shown in Figure 1-54. Findings from both the history and the physical examination should be used to define functional status.

**SUMMARY AND PROBLEM LIST**

After obtaining the history, performing the physical examination, and recording the results, the physiatrist should summarize the findings, construct a problem list, and formulate a plan.

A summary of findings can be a useful component of the written record. In a few sentences, a summary can provide a succinct description of relevant findings in the history and examination.

---

**TABLE 1.6** Example of Summary, Problem List, and Plan

<table>
<thead>
<tr>
<th>Summary</th>
<th>Example of Summary, Problem List, and Plan</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A 55-year-old male carpenter with a left hearing deficit and poorly treated hypertension presents 4 d after sudden moderate left-side spastic hemiparesis with moderate sensory deficits, left neglect, nocturnal bladder incontinence, and dysarthria. He is alert, oriented, and normotensive; motor function in his left hip and knee is returning; and he has an elevated serum cholesterol level. He is divorced, lives alone, and has no close family. Computed tomography of the head shows moderate right subcortical infarction. Evidence of ischemia is not shown in the electrocardiogram.</strong></td>
<td></td>
</tr>
</tbody>
</table>
| **Medical problems and plans** | 1. Right hemisphere infarction with motor, sensory, perceptual, and speech deficits: monitor neuromuscular function, maintain ROM, control spasticity (air splint, possible medications), provide motor reeducation, and provide patient education and risk factors  
2. Hypertension: monitor systolic/diastolic blood pressure and treat with antihypertensive agents as appropriate  
3. Dyslipidemia: low-fat diet, patient education about diet and food preparation, and lipid-lowering agents  
4. Urinary incontinence: check residual urine volume and culture specimens; treat urosepsis. If residual volume is low, offer urinal frequently, with or without nocturnal condom catheterization. If residual volume is high, begin 1,800-mL fluid intake schedule with catheterization every 6 h, urodynamics, and bladder retraining |
| **Rehabilitation problems and plans** | 1. Communication deficits: speech pathologist for evaluation and therapy  
2. Left neglect: OT for perceptual testing, retraining, and compensation; verbal clues to scan left; RN and PT to reinforce OT  
3. Left sensory deficits: monitor skin, offer patient education on care of insensate skin  
4. Self-care deficits: OT for upper-extremity ROM, reeducation about strengthening, ADL retraining, adaptive aids  
5. Safety and judgment deficits: four bed rails, RN monitoring at night, verbal clues, physical spotting  
6. Transfer deficits: PT for retraining, left wheelchair brake extension  
7. Mobility deficits: PT for lower-extremity ROM, reeducation about strengthening, gait retraining, gait aids  
8. Driving dependency: retasting and retraining with improvement  
9. Community reentry and poor support system: assess home for architectural barriers, assess home health services, identify additional social support  
10. Reactive depression: refer for psychological support  
11. Vocational issues: consider prevocational counseling and testing |

ADL, activities of daily living; OT, occupational therapy; PT, physical therapy; RN, registered nurse; ROM, range of motion.
For the management of chronic diseases, physiatrists must commonly address myriad physical, psychological, social, and vocational problems. Weed’s (20) problem-oriented medical record has been applied to the management of patients undergoing rehabilitation (21–24). Although the use of the problem list itself is the essential factor, a consensus as to the organization and the use of the entire system in the rehabilitation setting has proved challenging. The recommendation of Grabois (23) that medical and rehabilitation problems be separately listed is beneficial. In addition, it may be helpful to delineate individual plans for each problem at the conclusion of the workup (Table 1-6).

REFERENCES

APPENDICES

APPENDIX A  Long thoracic nerve; thoracic anterior nerve.

APPENDIX B  Dorsal scapular nerve; suprascapular nerve.
APPENDIX C  Thoracodorsal nerve; subscapular nerve.

APPENDIX D  Axillary nerve.
APPENDIX E  Radial nerve.

APPENDIX F  Median nerve.
APPENDIX G  Ulnar nerve.

APPENDIX H  Femoral nerve.
APPENDIX I  Lateral cutaneous nerve.

APPENDIX J  Sciatic nerve; tibial nerve; posterior tibial nerve.
APPENDIX K  Deep peroneal nerve.

APPENDIX L  Superficial peroneal nerve.
The hallmark of medicine has always been the physical examination. Perhaps more than the actual diagnosis, the process by which the physician arrives at his or her conclusion has defined the “art” of medicine. Much has been written about the techniques by which this art is performed, and much more will continue to be written. Each generation will take from the past and apply these techniques to the future of medicine.

The physical examination is an extension of the history and extends the doctor-patient relationship initially established during the history. The skill with which the examination is performed instills a sense of confidence in the patient that the examiner knows what he or she is doing. This confidence in the physician has a positive outcome on the patient’s ability to recover. Finally, the physical examination serves to narrow the list of diagnostic possibilities.

In a specialty like physiatry, in which the whole person is evaluated in terms of function, there is no adjunct more important than the physical examination. The examination provides the foundation to formulate a plan to improve a person’s function. Importantly, though, in looking at function, each piece must be applied to the whole person. The examination of one joint must be applied to the whole picture of the patient, and an understanding of functional biomechanics will enable the physician to include in the physical examination other structures that may indirectly contribute to the impairment.

The focus on function and application to the whole person in physiatry can be best seen in understanding the concept of the kinetic chain. No one joint, bone, or muscle acts alone in the body. An ankle sprain can lead to low-back pain. Low-back pain can affect the serve of a tennis professional. Lateral epicondylitis can alter shoulder mechanics and lead to rotator cuff impingement. It is because of these relationships that the physiatrist must perform a thorough examination. It is this comprehensive manner that sets apart the physiatric approach from others. A thorough knowledge of the neuromuscular system and an understanding of functional biomechanics will narrow the focus of the examination so it can be done in a time-efficient manner. The relationship between the different joints and regions must be understood. In addition, a complete understanding of the muscles and their innervation is required.

An understanding of the muscle kinesiology and biomechanics is very important in the physical examination. Each muscle functions across one or more joints to provide motion or stabilization. One example would be the hamstrings. When the foot is planted, the hamstrings act as powerful hip extensors. However, with the foot off the ground, they can become knee flexors. With a patient prone and the knee bent at 90 degrees, the gluteus maximus acts as the primary extensor because of the shortened hamstrings. Place the knee in full extension, and the hamstrings will once again act as hip extensors. We will look further into these types of relationships in the physical examination.

In today’s medicine, there exists a tremendous amount of information to digest. The number of articles indexed in MEDLINE has grown in size from 1,098,000 citations in 1970 to 11,761,000 in 2000. The modern physician must have an understanding of the body down to a microcellular level. In addition, access to modern tests like magnetic resonance imaging (MRI) is achieved by a greater number of patients. Any test has its limitations, and in the example of the MRI, these can be multiple false-positive findings (1). The MRI should be used to confirm not make a diagnosis. Many physician referrals are generated from a radiologist’s interpretation of a study, often without physical examination findings consistent with the results of the study. It is at this point that the well-trained physiatrist can be the link using evidence-based medicine as it applies to diagnosis, history, and physical examination.

Whole texts are dedicated to the physical exam. Due to the limits of one chapter, this will be an introduction to the physical examination and kinesiology of the cervical spine, shoulder, lumbar spine, and knee. That said, the reader should be able to approach any joint in the manner laid out here to aid in his or her diagnosis. Examination of any joint should be performed in a systematic approach. As the examination begins, the clinician should make sure that the area to be examined is properly exposed for evaluation and the patient appropriately draped. We have focused on the major joints seen in our practice—the cervical and lumbar regions of the spine, the shoulder, and the knee. Other joints will be addressed in chapters in this text. We will now address the physical examination, and the kinesiology of the muscles and joints will be explained. For reference, the dermatomes, myotomes, and sclerotomes are illustrated in Chapter 21.

It is the task of the physiatrist to perform a thorough physical examination to confirm his or her diagnosis derived from the history and additional information. It even is more important today, because of the additional tests modern technology has advanced, to understand physical examination maneuvers and their diagnostic relevance.
**EXAMINATION OF THE CERVICAL SPINE**

**Inspection**
Inspection of the neck begins upon meeting the patient. Look to see if the patient moves the shoulders when he or she turns the neck, a sign of decreased range of motion, or if he or she winces with certain motions. Take note of the patient’s relaxed posture as changes to improve poor posture can be easily addressed in therapy. As the examination proceeds, the clinician should make sure that the neck is properly exposed for evaluation. Look at the muscle bulk and symmetry of the neck, upper back, and shoulders. Also look at the skin for scarring or discoloration. You will be surprised at the details left out by patients. It is not uncommon to learn about a patient’s previous surgery during the exam.

**Palpation**
The next step involves palpation of the neck and upper thoracic region. Begin in a systematic fashion, either starting from the front or back. From the back, the paraspinal muscles and the nuchal ligament can be palpated. Working down, the upper and middle trapezius muscles should also be palpated for tender or trigger points (2). Palpate for the spinous process of the seventh cervical vertebrae, which should be larger than the superior segments in a neutral position of the cervical spine.

Place the patient in the supine position with the patient’s head near the end of the table. Sit with your stool directly behind the patient’s head, and continue palpation. Rotate the patient’s neck 45 degrees, palpate each zygapophyseal joint, and note whether the patient feels discomfort at a joint that is greater than the opposite corresponding zygapophyseal joint. From this position, the anterior muscles, most notably the sternocleidomastoid and more laterally the scalenes, can be palpated. Palpate the sternocleidomastoid muscle from its origin at the sternoclavicular joint to the insertion on the mastoid process. Rotate the neck from side to side to make the muscle more prominent if it is initially difficult to find. Look for symmetry and bulk.

**Range of Motion**
Range of motion should be checked in flexion, extension, rotation, and lateral or side bending. Motion is not divided equally between the vertebrae. Approximately 50% of flexion and extension come from the atlanto-occipital joint. At the atlantoaxial joint, approximately 50% of the rotation takes place (3).

Guidelines for normal motion are as follows: Normal flexion allows the patient to touch his chin to his chest, and extension allows the patient to look up at the ceiling. In normal rotation, the patient should be able to bring her chin over the acromion. Side bending done toward the ipsilateral shoulder should be approximately 45 degrees. Always begin with active range of motion, particularly in the injured patient. The patient may guard, and this will reduce the range. Forcing motion may make the patient uncomfortable and can injure a patient with zygapophyseal joint dysfunction (4).

**Neurological**
Included in any examination of the neck is a full neurologic examination of the upper limbs. Radiculopathies can be very subtle, and all components of the examination, manual muscle testing, sensory examination, and reflexes must be addressed to find these subtle changes. The order to proceed is examiner dependent. Manual muscle testing should also be confirmed with additional muscles when subtleties exist, as the muscles of the upper limb have two or more levels of innervation. Table 2-1 shows what should be included in manual muscle testing.

Reflexes can be addressed next. Table 2-2 shows what should be included in reflex testing.

Finally, sensation can be tested for both pinprick (lateral spinothalamic tract) and light touch (dorsal columns). If there is a concern about carpal tunnel or double crush, two-point discrimination may be more sensitive (5). Table 2-3 shows what should be included in sensation testing.

Examination of the neck should also include a compression test or Spurling’s maneuver (6) (Fig. 2-1). The test assesses the

---

**TABLE 2.1 Manual Muscle Testing of the Cervical Spine**

<table>
<thead>
<tr>
<th>Root</th>
<th>Muscle</th>
</tr>
</thead>
<tbody>
<tr>
<td>C5/6</td>
<td>Biceps, deltoid, and rotator cuff muscles</td>
</tr>
<tr>
<td>C6/7</td>
<td>Wrist extensors and pronator teres</td>
</tr>
<tr>
<td>C7/8</td>
<td>Triceps and wrist flexors</td>
</tr>
<tr>
<td>C8/T1</td>
<td>Finger flexors and intrinsic muscles of the hand</td>
</tr>
<tr>
<td>C8/T1</td>
<td>Intrinsic muscles of the hand</td>
</tr>
</tbody>
</table>

*Primary innervating root and initial muscle tested are shown in bold.

**TABLE 2.2 Reflex Testing of the Cervical Spine**

<table>
<thead>
<tr>
<th>Root</th>
<th>Reflex</th>
</tr>
</thead>
<tbody>
<tr>
<td>C5</td>
<td>Biceps</td>
</tr>
<tr>
<td>C6</td>
<td>Pronator teres or brachioradialis</td>
</tr>
<tr>
<td>C7</td>
<td>Triceps</td>
</tr>
</tbody>
</table>

*Primary innervating root tested.

**TABLE 2.3 Sensory Testing of the Cervical Spine**

<table>
<thead>
<tr>
<th>Root</th>
<th>Sensory Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>C5</td>
<td>Lateral (radial) side of the antecubital fossa</td>
</tr>
<tr>
<td>C6</td>
<td>Thumb</td>
</tr>
<tr>
<td>C7</td>
<td>Middle finger</td>
</tr>
<tr>
<td>C8</td>
<td>Little finger</td>
</tr>
<tr>
<td>T1</td>
<td>Medial (ulnar) side of the antecubital fossa</td>
</tr>
</tbody>
</table>

*Primary innervating root tested.*
mechanical neuroforaminal narrowing of the C4-5, C5-6, and C6-7 with ipsilateral oblique extension (7). The objective of the test is to compress an irritated nerve with the following motion: The neck is brought into slight extension and side bending followed by an axial compression. A positive result reproduces pain along a dermatome below the shoulder. Finally, as with the joints, it is always important to examine the adjacent joints. In the case of the cervical spine, a full examination of the shoulder should be performed to rule out underlying or contributing shoulder pathology.

EXAMINATION OF THE SHOULDER

Inspection

Inspection of the shoulder requires that the shoulder be exposed and the patient appropriately draped. The shoulder “joint” actually consists of four different joints: sternoclavicular, acromioclavicular, glenohumeral, and scapulothoracic. Three of the joints are true joints, while the scapulothoracic joint is not a true articulating joint lined with cartilage. It is important to visualize each of the joints. Begin by inspecting the normal bony prominences and muscle bulk. The most obvious changes can be seen in the acromioclavicular joint. Comparison to the other shoulder is essential.

Palpation

Palpation can be done either from in front of the patient or from behind the patient. Begin by palpating from the sternoclavicular joint along the clavicle to the acromioclavicular joint. Palpate the coracoid process and the coracoclavicular ligament. Move laterally, and palpate the tendon of the long head of the biceps. Continue palpation medially to the lesser tuberosity and laterally to the greater tuberosity. Next, palpate the scapula along the acromion and medially along the spine of the scapula. Find the superior and inferior angles of the scapula.

The muscles should also be palpated for tender points and evaluation of their bulk. The supraspinatus, infraspinatus, and teres minor can be palpated by bringing the upper limb into extension at the shoulder and palpating anteriorly (3).

Range of Motion

The motion of the joints should be observed. Watch the different joints and their symmetry of motion. This should be done from in front of and behind the patient.

The shoulder has the greatest range of motion of any joint. Subtle changes must be assessed and asymmetries noted during the physical examination. Active and passive motions should be assessed. To begin, check the patient’s active range of motion. There are six directions of motion: abduction, adduction, extension, flexion, internal rotation, and external rotation. Active abduction should allow the patient to touch the dorsal surface of his hands with the arms straight above the head. Adduction will allow the patient to bring her arm into the plane of the torso. Each of these can be tested in conjunction with the testing for internal and external rotations or alone. Functional internal rotation can be demonstrated by having the patient touch his midback (Fig. 2-2). Record the level that the thumb touches, and repeat on the opposite side. Have the patient reach over the head and touch the upper back to test external rotation. As with internal rotation, record both sides. Finally, have the patient bring the straight upper limb forward to test flexion and backward to test extension.

The shoulder should then be checked for passive range of motion. The importance of checking the passive range can be seen in a patient with adhesive capsulitis. Although there may appear to be both internal and external rotations, the motion often comes from the scapular thoracic joint. By isolating the glenohumeral motion, both can be assessed, and there is increased reliability in the assessment of the glenohumeral motion (8).

Passive internal and external rotations can be tested by bringing the shoulder into 90 degrees of abduction while holding
the elbow to 90 degrees of flexion. Stabilizing the scapula with one hand to truly evaluate glenohumeral motion, internally and externally rotate the shoulder (Fig. 2-3). For some examiners, placing the patient in the supine position with a posteriorly directed force on the coracoid process might be easier and has been found to be reliable (9). Note the motion, compare it to the other side, and repeat with the scapula free to see the scapulothoracic motion. Table 2-4 shows the normal range of motion.

**Neurological**

Motor testing of the shoulder should follow the examination of the range of motion. Each motion should be tested for strength. The major muscles used to move the shoulder are the deltoid, pectoralis major, latissimus dorsi, biceps, and triceps. In addition, there are smaller stabilizing muscles, including the rotator cuff muscles. Additionally, the scapular position and control are coordinated by the trapezius, levator scapulae, rhomboids, and serratus anterior. Test the major movers with one hand stabilizing the shoulder and the other providing resistance.

After testing the larger movers of the shoulder, it is important that the smaller stabilizers are addressed, as these are often involved in the pathology of the shoulder. The supraspinatus is tested with the upper limb abducted 90 degrees, internally rotated with the thumb down and in the plane of the scapula. Apply steady pressure while asking the patient to abduct the limb (Fig. 2-4). Next, the external rotators can be assessed. Have the patient adduct the limb and flex the elbow at 90 degrees. The examiner stabilizes the elbow against the torso with one hand and places the other hand on the distal forearm. The patient then rotates the forearm away from the body against resistance.

Finally, the subscapularis should be assessed. This is the most difficult to check for subtle changes. The classically described maneuver is the “lift-off test.” This is done by the patient placing the dorsum of his hand on his back while the elbow is flexed at 90 degrees. The examiner then holds the hand off the back and instructs the patient to hold his hand in that position once the hand is released. If the patient is able to maintain the hand position, the subscapularis is intact. If the hand falls to the back, there is some deficiency in the muscle (Fig. 2-5).

<table>
<thead>
<tr>
<th>TABLE 2.4 Normal Range of Motion of the Shoulder</th>
</tr>
</thead>
<tbody>
<tr>
<td>Motion</td>
</tr>
<tr>
<td>Flexion</td>
</tr>
<tr>
<td>Extension</td>
</tr>
<tr>
<td>Abduction</td>
</tr>
<tr>
<td>Adduction</td>
</tr>
<tr>
<td>Internal rotation</td>
</tr>
<tr>
<td>External rotation</td>
</tr>
</tbody>
</table>

**FIGURE 2-3.** Stabilization of the scapula during shoulder range of motion evaluation.

**FIGURE 2-4.** Supraspinatus strength determination during shoulder evaluation.

**FIGURE 2-5.** Subscapularis strength determination during shoulder evaluation.
Sensory testing of the shoulder should be done in conjunction with the neck. Of importance for the shoulder, the dermatome for the axillary nerve should be tested. This is a silver dollar–sized area over the deltoid on the lateral upper arm. This is especially important after dislocations, as the axillary nerve can be injured.

**Additional Tests**

**Impingement Tests**

There are many tests for impingement of the rotator cuff muscles. We will address two of the more common tests. The first is the Hawkins’ maneuver (10). With the arm abducted to 90 degrees, elbow flexed at 90 degrees, and the humerus in the plane of the scapula, the examiner stabilizes the scapula and internally rotates the shoulder (Fig. 2-6). Pain with this maneuver is caused by impingement of the greater tuberosity on the coracoacromial ligament.

The Neer’s impingement sign is performed by stabilizing the scapula and slowly forward flexing the shoulder (11; Fig. 2-7). The elbow should be straight during the maneuver. The limb can be tested both internally rotated and neutral during testing.

**Labral Tests**

Another test of importance to the shoulder exam is the active compression test. The test is used to assess for anterior labral tears and acromioclavicular injuries. With the patient standing, the examiner stands on the affected side. The shoulder is brought into 90 degrees of abduction, 10 to 15 degrees of adduction, and internal rotation of the upper limb. The patient then resists a downward force by the examiner. At this point, the patient should either feel pain at the top of the shoulder (A–C joint pathology) or inside the shoulder (anterior labrum). The limb is then brought into full external rotation and the symptoms should be alleviated. Sensitivity and specificity are both excellent for the test (12).

**Stability Tests**

Shoulder instability can be diagnosed with a variety of maneuvers and most likely more accurately using the results of two or more tests. The examiner can begin with the apprehension test. The patient is placed in a supine position with the upper limb to be examined next to the edge of the table. The shoulder is then abducted to 90 degrees, and the elbow is flexed. The examiner then externally rotates the shoulder (Fig. 2-8). A patient with a positive “apprehension sign” has discomfort and a feeling of apprehension in the shoulder as it is externally rotated past 90 degrees that is relieved when the examiner stabilizes the shoulder with a posteriorly directed force to the shoulder with his free hand. The second part of the examination is named the “relocation sign” (Fig. 2-9). Both parts of the examination check for anterior instability of the shoulder, although the relocation test adds specificity to the diagnosis.
superior labrum anterior posterior (SLAP) lesion, or the flaws in a runner that lead to patellofemoral pain, requires an understanding muscle kinesiology and joint biomechanics. This helps determine not only the factors that may have been causative but also those that may increase the risk of an injury, and if so identified, allow prevention. This can be especially valuable during preparticipation examinations (15). The focus of sports medicine continues to be treatment rather than prevention. One reason for this is the limited scientific data that clearly demonstrate the effectiveness of prevention programs based on preinjury evaluations. Intuitively we believe that restrictions in motion or certain strength deficits may predispose an athlete to an injury, and if we address them, we can lower the risk. But do we have the research to back this up?

Much of our understanding of muscle kinesiology comes from work done in labs using electromyography (EMG) to look at muscle firing patterns. One must be extremely cautious interpreting these studies. Although very general muscle firing patterns can be determined, some important technical factors are often overlooked. The relative activity of one muscle cannot be compared with another for several reasons. One is that the amplitude of a muscle’s EMG signal varies widely based on whether a muscle is contracting concentrically or eccentrically. During certain sports activities, there will usually be muscles undergoing both (usually agonists and antagonists) types of contractions simultaneously, and it may not always be readily evident which is doing which. For example, during the acceleration phase of throwing, the shoulder flexes forward, but the exact point at which the shoulder internally rotates is important in determining which of the rotator cuff muscles are contracting concentrically and which are contracting eccentrically. To be certain, one must also perform a video kinematic analysis.

Several other factors play an important role in comparing the EMG signal of different muscles. The amplitude of the EMG signal will vary based on the location of the electrode (in relation to the muscle’s motor point), the type of electrode (surface vs. intramuscular), and the degree of muscle fatigue. Furthermore, because one compares the EMG activity to activity during maximal voluntary contraction (MVC) of the same muscle, activity determined during MVC must be reliable and statistically reproducible. This issue is frequently not fully and adequately addressed.
Nevertheless, the information collected on muscle kinesiology has allowed us to better understand basic muscle mechanics. Understanding the major technical limitations will help prevent us from drawing erroneous conclusions.

**Biomechanics of the Overhead Athlete**

The biomechanics of the overhead athlete have been extensively studied. The motion of throwing a baseball and serving a tennis ball overhead has similarly been broken into five phases: windup, early cocking, late cocking, acceleration, and follow-through. The stage of late cocking, during which the shoulder is abducted and externally rotated, may potentially be dangerous to the glenohumeral joint, where inherent instability may lead to anterior translation and load the labrum or capsule anteriorly. Symptoms of posterior impingement can also be elicited when there is excessive anterior/posterior translation compressing redundant scar tissue in the region of the posterior capsule. Kinesiological studies have demonstrated that all four muscles of the cuff are most active from the late cocking to acceleration phase (16–18). This is not surprising because the cuff is felt to be a dynamic stabilizer of the glenohumeral joint, and the position in late cocking puts the glenohumeral joint in a potentially unstable position. Studies have also demonstrated that the triceps begins to fire in late cocking and then in acceleration (16,17). This is probably to prevent hyperflexion of the elbow during late cocking and may also serve as a prestretch to create a plyometric type of contraction of the triceps during acceleration to propel the forearm, wrist, and hand along with either the racquet or the baseball. The biceps then fires during the deceleration phase to allow elbow extension to occur in a controlled fashion. If this occurs too rapidly because of inadequate biceps control, overload can occur to the biceps muscle or biceps tendon or lead to avulsion, in which the biceps tendon anchors itself along the superior labrum. Injury to the labrum at this level has been identified as a SLAP lesion (19). The muscle kinesiological data collected have supported the theoretical basis for the mechanisms of injury to these various structures. This information can then be used on physical examination so that the clinician can reproduce symptoms in the phase where injury occurs. One then combines some basic physical examination findings based on observation and palpation with functional tests, such as the apprehension sign or testing the biceps during an eccentric load. Not only can an anatomic diagnosis be made of the injured structure, but a functional diagnosis can be made as well. One must also be careful to not confuse strength with motor skill. Adequate strength on manual muscle testing does not guarantee proper muscle function. Poorly developed muscle skill, proprioception, the proper agonist/antagonist balance during contractions, and the lack of proper timing of muscle firing can all contribute to an overuse injury. Any of these should be considered at least a potentially contributing factor.

To further shed light on a more complete biomechanical picture, the kinetic chain must also be considered. This requires a sound understanding of the role each component of the chain plays during a skilled athletic maneuver. Any pathology at any point in the chain can alter the athlete’s mechanics and lead to overload elsewhere. This may sound somewhat vague and generalized, but it is part of the functional approach practitioners working with athletes should consider. Throwers who have lost trunk/spine flexion/extension or pelvic/hip rotation may lose power from the loss of torque normally created during late cocking into acceleration phase or may have difficulty slowing down elbow extension during deceleration phase (20). Other components of the kinetic chain essential to minimizing trauma to the shoulder and arm are adequate neck rotation and eccentric strength of quadriceps. Compensation for restrictions in motion and relative weakness may lead to greater demands on power generated by the rotator cuff. This can create greater torque in the glenohumeral joint or require a greater and excessive eccentric contraction of the biceps. Eccentric overload of the biceps may injure the bicipital tendon of the labrum at its point of origin creating a SLAP lesion (19).

**EXAMINATION OF THE LUMBAR SPINE**

**Inspection**

The examination of the low back, like the other areas of the body, should begin as the patient enters the office and examination room. Watch how the patient moves while walking and how he or she moves changing positions. The patient’s posture should be noted. The patient should be in a gown that opens in the back for full exposure. Look at the muscle bulk and symmetry of the low back. Also look at the skin for scarring or discoloration. Inspect the lumbar spine from behind and the side to assess for lordosis. Often, patients with stiffness may have hypolordosis because of spinal stenosis. Young athletes might have hyperlordosis because of an imbalance of paraspinal to abdominal strength.

**Palpation**

The next step involves palpation of the muscles of the back, spinous processes, and important landmarks of the pelvis. From the back, the paraspinal muscles and the interspinous ligaments can be palpated. Palpate for the spinous processes, and in an older patient, these should be percussed to help in the diagnosis of a compression fracture. Finally, palpate for the bilateral posterior superior iliac spines (PSIS) to determine pelvis alignment. The examiner should place her thumbs on the bilateral PSIS and index fingers on the iliac crests. The height of the pelvis can be checked for alignment by comparing the two sides. Look for symmetry of bulk.

**Range of Motion**

Range of motion should be tested both actively and actively assisted if possible. Both are important in the evaluation of the low back. Range of motion should be checked in flexion,
extension, rotation, and side bending. If there is posterior pain to one side, the examination should include extension to both the left and the right to stress the zygapophyseal joints and to narrow the foramen in a patient with foraminal stenosis or a foraminal disc protrusion.

It is important to watch the spine during motion. In forward flexion, ask the patient to touch his or her toes and watch to see whether the motion comes from the spine or hips. Watch for reversal of the lumbar lordosis by inspecting the prominence of the spinous processes. In extension, look for the motion in the lumbar spine versus the hip and knees in many patients. While assessing range, ask the patient whether the discomfort is greater in flexion or extension. Be aware of conditions that can lead to spinal inflexibility like ankylosing spondylitis or diffuse idiopathic spinal hyperostosis (DISH).

Rotation and side bending can be evaluated next. The patient should be able to rotate his or her shoulders perpendicular to the pelvis. It is often helpful to stabilize the pelvis while the patient is rotating. Have the patient side bend next, and compare it to the opposite side. With each maneuver, the examiner can follow the active motion with active assisted motion to see to what degree the active motion is limited.

Examination of the hip joint and the muscles crossing it is an important part of the lumbar spine examination because of the intimate association with the pelvis and lumbar spine. Limited hip rotation may lead to increased rotatory forces in the spine. A tight rectus femoris may tilt the pelvis anteriorly, increasing the lumbar lordosis, whereas hamstring tightness may tilt it posteriorly and decrease it.

Maybe no other joint in the young person has seen more change in approach over the past few years than the hip. In evaluating the spine, the examiner should have an idea of any suspected loss of range of motion. In the older patient, the loss of range of motion, particularly internal rotation, needs to be documented, and the practitioner needs to determine how much that pain contributes to the patient’s symptoms. In a younger patient, the loss of range of motion can be early osteoarthritis, but in the absence of joint space loss on plain film radiographs, it could be a soft tissue injury or a bony anatomy change. Studies have shown that labral tears can be seen in young patients with complaints of groin pain approximately 20% of the time (21). These lesions have a high association with bony abnormalities (22) and could be precursors for osteoarthritis (23).

The examination of the hip should consist of at least three elements. The first standing on one leg or walking to look for dynamic weakness in the form of a lurch to the opposite side or compensation to the same side due to weakness. This can be checked in the side lying position statically. Next, the patient should be supine and simple range of motion should be checked at 90 degrees of hip and knee flexion. Finally, the hip should be checked in flexion at 90 degrees, adduction, and internal rotation for the presence of groin pain. Table 2-6 shows normal range of motion of the hip.

### Neurological

The examination of the low back always includes a full neurologic examination of the lower limbs. Radiculopathies can be very subtle, and as with the cervical spine examination, manual muscle testing, sensory examination, and reflexes all must be addressed to find these subtle changes. The order to proceed is examiner dependent. Similar to the cervical spine examination, manual muscle testing should also be confirmed with additional muscles when subtleties exist because the muscles of the lower limbs have two or more levels of innervation. However, unlike the upper limbs, the lower limb muscles can generate greater force. The examiner needs to provide enough resistance to detect subtle muscle weakness. In addition, heel and toe walking can be added to the gait examination to test the tibialis anterior and gastrocnemius-soleus muscles. Table 2-7 lists what should be included in manual muscle testing (19).

Reflexes can be addressed next. Table 2-8 lists what should be included in reflex testing (19).
Palpation
Palpation of the knee should be done systematically. Begin either medially or laterally, and work across the knee. Address the skin, soft tissue, and bony aspects of the joint. Begin the examination by laterally palpating the overlying skin, which should move freely over the soft tissue and bones. The lateral collateral ligament can be palpated next. Palpate along the length of the ligament from the lateral femoral condyle to the insertion on the fibula. Have the patient cross the leg (FABER—Flexion of the knee to 90 degrees, Adduction and External Rotation of the hip) to better palpate the ligament. Moving more proximally, the biceps femoris tendon can be palpated as it comes down to its insertion on the fibular head.

Bony palpation should include the lateral tibial plateau, fibular head, and the lateral femoral condyle. All should be felt for tenderness or palpable osteophytes. In addition, the anterior portion of the lateral meniscus lies on the lateral tibial plateau and may be tender after an injury. This should be checked with the knee in 90 degrees of flexion. Moving across, the anterior portion of the knee should be palpated. Palpate over the prepatellar bursa (above the patella) and the superficial infrapatellar bursa (overlying the infrapatellar ligament). Next in the region is the patella. All four poles of the patella should be palpated, in addition to the undersurface of the medial and lateral aspects. Palpation of the medial and lateral facets of the patella can be performed with the patient lying supine and the knee completely relaxed. Tenderness or hypersensitivity is indicative of patellofemoral pathology. Furthermore, one should palpate the lateral retinacula for the presence of a synovial plica (Fig. 2-11). Proximally, the quadriceps muscle should be palpated for any discomfort of defects. Distally, the infrapatellar tendon should be palpated to its insertion on the tibia at the tibial tubercle.

The medial portion of the knee is addressed in a similar fashion. Palpate the skin, and palpate in the region of the pes anserine bursa (medial to the tibial tubercle and just above the insertion of the tendons of the sartorius, gracilis, and

Additional Tests
Examination of the low back should include special tests that are specific for certain pathologies. Every back examination should include a straight leg raise if there is concern about radiculopathy. The straight leg raise, also known as the Lasegue’s test, can be performed with the patient seated or in the supine position. With the patient supine, raise the affected lower limb with the knee in full extension. Starting at 30 degrees of leg elevation, patients with nerve root irritation will begin to have discomfort. Stretch on the nerve will be maximal at 65 degrees, and pelvic rotation will begin. A positive test is pain down the limb to the knee in the arc of 35 to 65 degrees. For more subtle cases, ankle dorsiflexion can be added to maximize the nerve stretch.

EXAMINATION OF THE KNEE
Inspection
Inspection is only possible with adequate exposure. Begin by placing the patient in shorts or tying the gown up above the knee. The patient’s gait should be observed first. Pay attention to the positioning of the knee on both the medial/lateral plane (valgus vs. varus) and the anterior/posterior plane (extension lag vs. knee recurvatum). Also, observe the joint above and below. Be sure to note any restrictions in the hip or ankle motion. Look at the foot for evidence of pes cavus (high arch) or pes planus (flat footed). The patient should then sit or lie on a table.

It is important to next look at the joint for any gross evidence of effusion or discoloration. Any changes can be evaluated further as the examination proceeds. Next, the examiner should assess for any muscle atrophy or fasciculations. If there is atrophy, the thigh or calf circumference should be measured and compared with the unaffected side. Finally, check the skin for any evidence of scarring from previous surgery of trauma.

Finally, sensation can be tested for both pinprick (lateral spinothalamic tract) and light touch (dorsal columns). Table 2-9 lists what should be included in sensation testing (19).

### Table 2.9: Sensation Testing of the Lumbar Spine

<table>
<thead>
<tr>
<th>Root&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Sensory Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>L2</td>
<td>Midanterior thigh</td>
</tr>
<tr>
<td>L3</td>
<td>Medial femoral condyle</td>
</tr>
<tr>
<td>L4</td>
<td>Medial heel</td>
</tr>
<tr>
<td>L5</td>
<td>Dorsum of the foot at the third metatarsal phalangeal joint</td>
</tr>
<tr>
<td>S1</td>
<td>Lateral heel</td>
</tr>
</tbody>
</table>

<sup>a</sup>Primary innervating root.

![Figure 2-11](image-url) Palpation of the lateral retinacula of the knee for synovial plica.
TABLE 2.10 Manual Muscle Testing of the Knee

<table>
<thead>
<tr>
<th>Root</th>
<th>Muscle</th>
</tr>
</thead>
<tbody>
<tr>
<td>L2,3,4</td>
<td>Quadriceps, iliopsoas, hip adductors</td>
</tr>
<tr>
<td>L3,4,5</td>
<td>Anterior tibialis, quadriceps, hamstrings</td>
</tr>
<tr>
<td>L4,5</td>
<td>Extensor hallucis longus, gluteus medius</td>
</tr>
</tbody>
</table>

*Primary innervating root and initial muscle tested are in bold.

Neurological

The neurologic examination should consist of manual muscle testing, sensation, and reflexes. The manual muscle testing is performed to test quadriceps strength by extending the knee. Hamstring testing should be performed with the patient flexing the knee while sitting. Another useful test is a step down test. Watch the patient step down from a foot stool in the room to assess his or her control descending and the amount of increase in the Q angle. Table 2-10 lists what should be included in manual muscle testing (24).

Reflexes can be addressed next. Table 2-11 lists what should be included in reflex testing (24).

Finally, sensation can be tested for both pinprick (lateral spinothalamic tract) and light touch (dorsal columns). Table 2-12 lists what should be included in sensation testing (24).

Ligament Stability

Stability of the ligaments should be tested with the patient relaxed and in a supine position. Beginning with the collateral ligaments, the examiner should firmly grasp the distal leg and provide a valgus (laterally applied) force to the knee. This will test the medial collateral ligament. The test should be completed with the knee in 20 to 30 degrees of flexion and also with the knee in full extension to test medial capsular integrity (Fig. 2-12). Remember to apply three points of pressure, one being distal lateral leg, the next lateral knee, and finally distal medial knee to maintain control of the leg. If possible, palpate around the knee, and palpate the ligament for a defect during the application of a valgus force.

In a similar fashion, apply a varus (medially applied) force to the knee to check the lateral collateral ligament. Again, it is helpful to place a finger on the ligament during the maneuver. It is also important to apply three points of pressure. As with the medial side, check in full extension and in 20 to 30 degrees of flexion.

The anterior and posterior cruciate ligaments should be examined next. The Lachman’s maneuver is the most sensitive...
test for injury to the anterior cruciate ligament (ACL). The test is performed by firmly grasping the distal lateral thigh with the outside hand in a supine patient. The knee is then placed in slight flexion, approximately 30%. Next, the proximal medial leg is grasped by the examiner’s inside hand and slightly laterally rotated. A quick upward force is then applied to the tibia by the inside hand while the thigh remains stabilized by the outside hand. The examiner is feeling for a sharp end point of the ACL. This examination maneuver is difficult and must be practiced many times before it can be done correctly (Fig. 2-13), but this is the most accurate method of judging the integrity of the ACL (25).

With the patient in a supine position and the hip flexed at 45 degrees while the knee is in 90 degrees of flexion, the examiner can test both the posterior and the anterior cruciate ligaments. The foot is stabilized when the examiner sits on the patient’s foot. To test the ACL, the examiner grasps around the proximal tibia and places the thumbs on the medial and lateral tibial plateaus. The tibia is then pulled anteriorly with respect to the femur. The amount of anterior movement should be minimal and equal to the opposite side. The movement is compared with the opposite knee.

Testing of the posterior cruciate ligament is completed just after the ACL. With the patient supine, the hip is flexed to 45 degrees and the knee flexed to 90 degrees. The foot is immobilized by the examiner sitting on the foot. The examiner then gives a posteriorly directed force to the tibia with the thumbs on the tibiofemoral junction. As with the anterior drawer test, the laxity is compared with the opposite side. Another indication of a posterior cruciate tear is hyperextension of the knee joint. This can be observed with the patient supine and the hip and knee flexed at 90 degrees. The examiner elevates the leg by lifting the heel with all muscles relaxed. Again, both sides should be tested for comparison.

The posterolateral complex of the knee includes the posterolateral capsule, the popliteus muscle, and the lateral collateral ligament. When one or more of these structures are injured, particularly in the setting of a posterior cruciate ligament deficiency, the knee becomes susceptible to rotatory instability. Posterolateral complex laxity can be demonstrated by examining the tibial external rotation with the knee flexed at 90 degrees and comparing it with the contralateral side.

**Medial and Lateral Menisci**

The medial and lateral menisci may account for the second most commonly injured structures in the knee, second only to the patellofemoral joint (PFJ) as a source of knee pain in the younger patient groups. Rotatory motion, particularly when combined with compression, is felt to be the common biomechanical factor leading to injury. An aging and degenerative meniscus is probably more susceptible to this type of trauma. Commonly, injury to the meniscus will result in an effusion, making the detection of an effusion an important clinical test when looking for meniscal injury. Joint line tenderness is sensitive for meniscal injury but not specific. The posterior horns are loaded during flexion so that simultaneous knee flexion and rotation will be sensitive for pain secondary to a posterior horn meniscal tear. Pain associated with the internal tibial rotation tends to be more indicative of injury of the lateral meniscus, whereas external rotation may be more suggestive of the medial meniscus.

A test for meniscal injury would be the McMurray’s test (26). McMurray’s test is performed with the patient supine.

---

**TABLE 2.12 Sensation Testing of the Knee**

<table>
<thead>
<tr>
<th>Root&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>L3</td>
<td>Superior and medial to the patella</td>
</tr>
<tr>
<td>L4</td>
<td>Medial side of the knee and anterior medial shin</td>
</tr>
<tr>
<td>L5</td>
<td>Anterolateral shin and the dorsum of the foot between the second and the third web space</td>
</tr>
<tr>
<td>S2</td>
<td>Popliteal fossa</td>
</tr>
</tbody>
</table>

<sup>a</sup>Primary innervating root.
FIGURE 2-14. Evaluation for meniscal injury or McMurray’s test of the lateral meniscus for the knee.

The knee is brought into full flexion, and the tibia is internally rotated and then extended to 90 degrees while being held internally rotated. An audible pop, click, or locking is considered a positive McMurray’s test and felt to be specific for posterior horn bucket handle lateral meniscal tear. Externally rotating the tibia and performing the same motion will detect injury to the posterior horn of the lateral meniscus (Fig. 2-14).

**Biomechanics of the Knee**

The knee appears to function primarily as a hinge joint, but with closer observation, its biomechanics are more complex. Rotatory motion also occurs and, although very limited, may play an important role for many of the acute traumatic and chronic overuse injuries. The primary static stabilizers include the anterior cruciate and posterior cruciate ligaments, the posterolateral complex, the remaining capsular structures, and, to a lesser extent, the medial and lateral menisci. The role of the dynamic stabilizers of the knee in controlling rotatory motion has not been well studied. However, it does appear that the medial hamstrings, lateral hamstrings, and popliteus muscles play a role here in dynamic rotatory stabilization. Although knee muscle kinesiology has been extensively studied, the great majority of work has been looking at the biomechanics of the PFJ (27–33). This is not surprising, considering that patellofemoral syndrome is the most common knee disorder causing pain and limiting function.

There is no other musculoskeletal disorder in which the kinetic chain plays a greater role or requires a more thorough analysis than with patellofemoral-related pain. It is widely believed that the relative position of the patella in the PFJ, how it sits at rest, and how it travels during dynamic activities can contribute to patellofemoral syndrome and be a risk factor for patellofemoral subluxation/dislocation (26,34). The quadriceps muscles are the primary knee extensors, with a small contribution coming from some fibers of the adductor magnus (35). Three muscles of the PFJ—the vastus lateralis (VL), the vastus medialis, and the vastus intermedius—cross only the knee joint and are relatively fixed in their line of pull. Tightness in the lateral or medial retinacular structures can somewhat alter this. The hip joint is the primary rotator of the lower limb, and the degree of rotation may play an important role in patella tracking disorders. The fourth quadriceps muscle, the rectus femoris, is a two-joint muscle that crosses the hip in addition to the knee joint. It originates from the anterior superior iliac spine (ASIS), and calculating the Q angle reflects its line of pull. The Q angle is measured by extending a line from the ASIS to the midpoint of the patella. One measures the angle created by the intersection of the second line that connects the midpoint of the patella to the tibial tubercle. The normal Q angle is 10 to 14 degrees, and any significant deviation from this may lead to improper patella tracking and subsequent PFJ pain. External rotation of the hip decreases the Q angle, whereas internal rotation increases it. During normal gait mechanics, ankle pronation occurs simultaneously with hip internal rotation; conversely, supination occurs with hip external rotation. Therefore, hyperpronation can increase the Q angle, whereas hypersupination can decrease it.

EMG has been used to study knee muscle function, primarily looking at the balance and relationship among the VL, vastus medius (VM), and vastus medialis oblique (VMO), and to better understand patellofemoral maltracking syndromes (27,36–39). Szepanski et al. (38) compared VMO and VL EMG activity during concentric and eccentric isokinetic exercises in asymptomatic individuals and found a greater VMO/VL ratio only during concentric contractions at 120 degrees per second. Reynolds et al. (37) studied asymptomatic women and found no difference in the VMO/VL relationship through full range of motion. In a study that looked at the effect of Q angles, Boucher et al. (27) found no significant differences in VMO/VL EMG ratio between asymptomatic volunteers and patients with patellofemoral maltracking syndromes. They did find a decrease in the VML/VL ratio in a subset of patellofemoral syndrome (PFS) patients with Q angles greater than 22 degrees at 15 degrees of knee extension. Voight and Wieder (39) compared the reflex response times of the VMO and VL EMG following a tendon tap. There was an increase in the VL response times in patellofemoral maltracking syndrome patients. These studies are far from conclusive, and the debate about the relationship between the VMO and the VL as contributing factors for patellofemoral disorders continues, while conventional clinical management remains based on these principles.

Kinesiological work has also been done to better understand muscle mechanics as it pertains to patients who have torn their ACLs and to help determine the most effective methods of managing these patients both nonsurgically and postoperatively. The ACL restrains anteromedial rotation of the tibia. An EMG study by Limbard et al. (40) on ACL-deficient patients found an increase in biceps femoris activity with a simultaneous decrease in quadriceps activity during swing-to-stance transition at normal walking speeds. At this point in gait, the hamstring may have been firing to prevent anteromedial tibial rotation. The hamstrings were less active...
in these patients from midstance to terminal stance. Branch et al. (41) found an increase in EMG activity of the lateral hamstrings in ACL-deficient patients during swing phase and an increase in medial hamstring and a decrease in quadriceps activity during stance phase. Tibone et al. (42) had reported similar findings. Solomonow et al. (43) stressed that an intact ACL led to excitement of the hamstrings and inhibition of the quadriceps. Baratta et al. (44) studied coactivation patterns. Hypertrophy of the quadriceps impaired hamstring coactivation, and strengthening of the hamstrings reduced this. Lutz et al. (45) demonstrated a greater ability to perform cocontractions of the hamstrings and quadriceps during closed kinetic chain exercises, thus conferring more stability to the knee. Weresh et al. (46) studied the popliteus muscle and found no difference in activation between ACL-deficient patients and controls.

Based on these EMG studies, one can now look for some of the muscle imbalances and other anatomic factors for patellofemoral maltracking syndromes such as hyperpronation or excessive hip internal rotation during the physical examination. Once these findings have been identified, they can then be more specifically addressed with physical therapy or some other form of a structured exercise program.

**SUMMARY**

The physical examination is one of the most basic and valuable clinical tools of a physician. A trend to spend less time examining patients and more money on tests is the result of the increasing demands on physicians’ time and the technological advances in medicine. We believe the manual skills of the physiatrist and the comprehensive physical examination performed are basic and unique to the clinical care that physiatrists offer. This chapter has attempted to review some of the basic traditional clinical examination skills in conjunction with the entire kinetic chain. We have also incorporated some of the relevant biomechanical and kinesiological components that can help the physician approach the physical examination in a more functional manner.

**REFERENCES**

To move is all mankind can do, and for such the sole executant is muscle.

Sir Charles Sherrington, Edinburgh, 1937–1938

The relevance of skeletal muscle to the performance of all types of physical activities (i.e., therapeutic, recreational, occupational, and others) and the successful participation in daily life and societal obligations should be well appreciated by all those who work in rehabilitation. Skeletal muscle plays an important role, primary and/or secondary, in the pathophysiology of many diseases, and skeletal muscle function is key to defining the nature and extent of impairments and activity limitations. Thus, understanding how to measure skeletal muscle function and how to interpret the results of various physiologic and functional tests is a necessary component of the education of all physiatrists and rehabilitation professionals. It is worth noting that this understanding is of special value to the advancement of research in the rehabilitation sciences because many biological and functional outcome variables used in scientific studies are directly associated with the function and structure of skeletal muscle. Finally, this chapter discusses this topic in the context of what is known about human skeletal muscle in health and disease. We will not review the extensive literature on muscle function based on studies in various animal models. With very few exceptions, the references will be those from human studies.

WHY MUSCLE?

The main function of the approximately 600 muscles in the human body is to convert chemical energy (i.e., fat and carbohydrates) into mechanical energy and thereby generate force. This force is transmitted from the active muscle fibers to the tendons with the help of the sarcolemma, special extracellular protein complexes, and connective tissue elements. The action of the tendons on bony structures results in the conversion of the force into joint and limb movement and displacement of individual body parts or the body as a unit. In principle, force generation can occur during brief moments, resulting in what is generally referred to as muscle strength, or force generation can be maintained over a period of time referred to as muscle endurance. In the clinical setting, the failure to generate force during a brief moment is what we generally call muscle weakness, as opposed to the inability to maintain force, which we refer to as muscle fatigue.

Skeletal muscle comprises 40% to 45% of the total body mass (1–3), and 55% of total muscle mass is distributed in the lower limbs. Muscle contains approximately 50% of the total body protein (4), and protein turnover in muscle represents 25% of the total body protein turnover (3,5). More than half of the protein in muscle is found in the thick (myosin) and thin (actin) contractile filaments that generate and regulate force production (4,6). Actin and myosin account for more than 80% of the protein in the myofibrillar complex. In addition to force generation, skeletal muscles contribute to basal metabolism; produce heat to maintain core temperature; regulate blood glucose; serve as storage for carbohydrates, lipids, and amino acids; contribute to energy production during exercise; and protect internal organs (7).

During illness, nitrogen must be mobilized from muscle to provide amino acids to the immune system, liver, and other organs. Thus, if sufficient nitrogen is not available due to muscle wasting associated with aging, immobilization, or severe illness, the body’s capacity to withstand an acute insult declines. The relationship between muscle function, illness, morbidity, and mortality is more obvious if one considers that morbidity becomes demonstrable at a 5% loss of lean mass and that the loss of 40% of lean body mass (LBM) is fatal (8). Finally, the extensive and considerable plasticity demonstrated by skeletal muscle under various conditions and in response to environmental influences such as bed rest, exercise training, and electrical stimulation makes it an ideal target for therapeutic and rehabilitative interventions.

From a functional point of view, skeletal muscle strength has been associated with comfortable and maximal walking speed (9,10), the incidence and prevalence of disability (11–13), balance (12), time to rise from a chair (14,15), ability to climb stairs (16), incidence of falls (17), and survival rate (18,19). Muscle power, a related but distinct property of skeletal muscle, also shows a positive and significant association with functional status (20). This evidence provides strong support to the conclusion that enhancing and maintaining muscle strength and muscle endurance throughout the life span, whether it is through prevention or rehabilitation, may reduce the prevalence of limitations in recreational, household, daily, and personal care activities, both in health and disease (21).
**MUSCLE ACTIONS AND UNITS OF MEASUREMENT**

The need to use consistent terminology, definitions, and measurement units across research studies, in educational programs, and in clinical rehabilitation is of great importance. Furthermore, measuring devices must be reliable and valid indicators of muscle function. Muscle function and structure can be quantified using the International System of Units, a refinement of the metric system (22–24) (Table 3-1).

All types of muscle actions result in the production of force or torque (tendency of a force to produce rotation about an axis). When the force is applied against an immovable object and there is no joint angular movement, the action is called static (isometric). Work is defined as the product of force × distance, and power as the ratio of work over time. Therefore, by definition, because during a static muscle action the distance is zero, no work is performed and no power is produced.

When the muscle action results in the displacement of a given mass or body part at the same time that the origin and insertion of a muscle move closer together, the action is called dynamic or concentric action. When the action results in the displacement of a given mass, and the origin and insertion of a muscle move further apart, the action is called dynamic eccentric or lengthening action. Both concentric and eccentric muscle actions result in work (force × distance), positive in the first case and negative in the latter. During many natural activities such as walking and running, concentric muscle actions occur in immediate combination with eccentric actions and are referred to as the **stretch-shortening cycle** (25, 26).

In these activities such as jumping and running upstairs, the rate at which force is developed is more important than generating maximal force. Thus, power ([work/time] or [force × velocity]) and not strength becomes the limiting factor. Some of the most important physical and biomechanical concepts in the study and measurement of muscle function are illustrated in Figure 3-1 (27).

Isokinetic muscle actions are dynamic and could be concentric or eccentric. This kind of muscle action is characterized by a combination of constant angular velocity and variable resistance. The resistance generated by the isokinetic device varies throughout the range of motion in order to match the torque generated by the muscle at each angle of the range of motion. It should be recognized that isokinetic actions represent an artificial situation that does not usually occur in nature outside of a laboratory. Many devices have been developed to measure muscle torque, work, power, and endurance based on the isokinetic concept. Isokinetic dynamometers, although expensive, are found in many research laboratories as well as in rehabilitation clinics (Fig. 3-2). Advantages of these devices

---

**TABLE 3.1** Examples of Units for Assessing Muscle Structure and Function

<table>
<thead>
<tr>
<th>Physical Quantity</th>
<th>SI Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mass</td>
<td>Kilograms (kg)</td>
</tr>
<tr>
<td>Distance</td>
<td>Meter (m)</td>
</tr>
<tr>
<td>Time</td>
<td>Second (s)</td>
</tr>
<tr>
<td>Force (mass × acceleration)</td>
<td>Newton (N)</td>
</tr>
<tr>
<td>Work (force × distance)</td>
<td>Joule (J)</td>
</tr>
<tr>
<td>Power (force × velocity)</td>
<td>Watt (W)</td>
</tr>
<tr>
<td>Velocity</td>
<td>Meters per second (m/s)</td>
</tr>
<tr>
<td>Torque</td>
<td>Newton-meter (Nm)</td>
</tr>
<tr>
<td>Angle</td>
<td>Radian (rad)</td>
</tr>
<tr>
<td>Angular velocity</td>
<td>Radians per second (rad/s)</td>
</tr>
<tr>
<td>Volume</td>
<td>Liter (L)</td>
</tr>
</tbody>
</table>

---

**FIGURE 3-1.** Physical and biomechanical concepts relevant to the assessment of human skeletal muscle function. Note the relevance of time on several variables, such as displacement, velocity, and work. Also, note the sequence of formulas leading from mass to power (27).

**FIGURE 3-2.** Picture of an isokinetic dynamometer. Here, the torque of the ankle dorsiflexor muscles is measured.
include the objective quantification of muscle function, the immediate availability of reports, the provision of feedback to the patient, their high reproducibility, and the opportunity to standardize sequential testing for follow-up purposes during the rehabilitation process. Disadvantages include the cost, lack of portability, and limited specificity in relation to muscle actions typical of daily activities.

The terms open kinetic chain and closed kinetic chain are used to describe two forms of muscle contractions and movements. The kinetic chain is a concept that describes a body segment as a series of mobile segments and linkages (28–30). In the case of the lower extremities, this chain allows forward propulsion during gait. When the foot is in contact with the ground, the kinetic chain is considered to be closed. When the foot is off the ground, the chain is said to be open. Examples of open kinetic chain exercises used in rehabilitation programs are leg extension, leg curls, arm curls, and bench press exercises. Exercises such as leg press, squats, and push-ups are examples of closed kinetic chain exercises. Closed kinetic chain exercises tend to activate agonist and antagonist muscle groups simultaneously (e.g., the knee extensors and flexors during squat exercises) and tend to be more functional (31).

Both types of exercises could result in significant functional improvements after reconstruction of the anterior cruciate ligament (30).

Energetics of Muscle Actions
The energy needed for the muscle to perform its mechanical functions is supplied by three different energy-producing biochemical pathways (Fig. 3-3). The relative contribution of each pathway is determined by the duration and intensity of the muscle actions. Performance of a particular task is determined not only by the integrity and capacity of the sarcomeric proteins but also by the ability of these pathways to supply adenosine triphosphate (ATP). Thus, the results of the functional tests discussed later could be used as indicators of the status of the biochemical pathways. The low strength and/or endurance performance scores in patients with various neuromuscular diseases may relate to abnormalities in these pathways.

In general, short-duration tasks lasting up to 10 seconds depend on existing stores of ATP and creatine phosphate (CP) (32,33). These two stores are readily available and therefore could be used instantaneously. However, from a quantitative point of view, these ATP and CP stores are very small and have a limited ability to sustain muscle performance over time. Activities lasting between 10 seconds and 2 minutes are driven by the process of anaerobic glycolysis fueled by the transport of glucose into the muscle cell or the breakdown (glycogenolysis) of intramuscular carbohydrates (34–38). Finally, the energy for activities lasting more than 2 minutes is supplied mainly by the oxidative pathways in the mitochondria (Fig. 3-4). The fuel for these pathways can be derived from the end product of anaerobic glycolysis, circulating fatty acids, or intramuscular lipid stores (39,40).

In real life, these biochemical processes combine in various proportions to provide ATP during physical activity and exercise. Activities can be classified as “predominantly” dependent on a particular pathway, since very few activities can be considered purely dependent on any given pathway.
In other words, a given activity may require a combination of all three processes, depending on fluctuations in the intensity of the exercise. For example, when a person is walking on a level at a comfortable speed, ATP supply may depend predominantly on oxidative pathways. Confronted with an incline or hill, the contribution of the glycolytic pathways increases. Another example in a different type of activity is the sprint at the end of a marathon race requiring the activation of the glycolytic pathway in a predominantly oxidative event.

**Functional Characteristics of Skeletal Muscle**

**Muscle Strength**

Muscle strength can be defined as the maximal force (or torque) generated by a muscle or muscle group at a specified velocity. Because strength depends on force production, it is generally measured in newtons (N) or newton-meters (Nm) in the case of torque. When reporting measurements of muscle strength, the type of muscle action must be stated (24). In other words, strength can be static (at different joint angles), dynamic (concentric or eccentric), or isokinetic (at different angular velocities).

It should be clear that there is no single strength measurement and that different kinds of strength can be expressed. Furthermore, under static conditions, force is influenced by fiber (and sarcomere) length (41) and mechanical leverage. Moreover, under dynamic conditions, the level of force is influenced by the velocity of the movement (41). These relationships are two of the most fundamental biologic properties of skeletal muscle and must be understood to appreciate the meaning of the results of functional tests. For example, the patient’s performance during a manual muscle test will not be reliable unless the test is always done at the same joint angle.

The force-length relationship illustrates how the sarcomere length, which defines the degree of overlap between actin and myosin and the formation of cross-bridges, determines force (Fig. 3-5). The optimal sarcomere length varies with the type of activity. For example, the optimal sarcomere length has been reported to be in the region around the plateau for ankle bending (42), walking (43), and jumping (44) activities and in the descending limb for slow pedaling (45). On the other hand, the force-velocity curve demonstrates a gradient of strength that ranges from the highest level during fast eccentric actions to the lowest level during fast concentric actions (Fig. 3-6). Static actions generate more than dynamic concentric actions but less force than dynamic eccentric actions, independent of the velocity.

**Testing of Muscle Strength**

Various methods and devices are used to measure the different types of muscle strength (46). These methods used in clinical practice and research require a maximal voluntary action on the part of the patient or volunteer. This is dependent on the ability of higher central nervous centers to recruit and modulate the frequency of discharge of the appropriate spinal motoneuron pool. The implication is that all those factors that influence the activation of the neuromuscular system such as age, various disorders in the central and peripheral nervous system, presence of pain, joint swelling, medications, fear and anxiety over the test, lack of motivation, time of the day, and environmental conditions such as noise may have a significant effect on strength measurements and should be controlled for...
or factored into the analysis of the results (47–49). Needless to say, testing conditions must be standardized as much as possible, the same device and/or method and testing protocol must be used when repeated measurements are required, the subject must be encouraged to make a maximal effort during the test, and the presence of symptoms such as pain should be considered when interpreting the measurements. Furthermore, for comparisons among groups, it may be necessary to adjust for differences in muscle/body size using statistical techniques (50). Even under optimal conditions, a valid and reliable level may require that the strength test be repeated more than once (51). Over the last decade, there has been a growing interest in, and a gradual development of, the statistical methods for the analysis of reliability. Today, there is a general agreement that a comprehensive set of several statistical methods are required to fully address the reliability of a measurement method (52).

**Manual Muscle Testing**

The method of muscle strength measurement used most frequently in the busy clinical setting is manual muscle testing. This technique uses a subjective scale (Table 3-2) that ranges from zero (complete paralysis) to normal strength (erroneously called normal muscle power by some authors) and is generally known as the “Classification of the Medical Research Council of Great Britain” (53,54). The tester’s perception of the strength of a given muscle is influenced by the duration of the tester’s effort and the force applied during the test (55). The manual muscle testing scale is characterized by a fairly high level of intra- and interrater variability that limits its usefulness for research studies and clinical follow-up (56). To distinguish among the various degrees of muscle strength within a given level, this scale has been modified with the addition of intermediate levels (e.g., 4+ and 4−). Although clinically useful, there is no evidence that this modification increases the validity or the reliability of the method. It should be understood that this method represents an estimate of static strength at the tested joint angle. Extrapolation of the results to other joint angles and especially to dynamic actions must be done with great caution.

**Static (Isometric) Maximal Voluntary Contraction**

A static maximal voluntary contraction (MVC) refers to a condition in which a person attempts to recruit as many muscle fibers in a muscle as possible for the purpose of developing force (24). Although the need for a maximal voluntary effort applies to all forms of strength testing, the term maximal voluntary contraction is frequently associated with static strength testing (57). Devices such as hand-held dynamometers, cable tensiometers, force transducers, and isokinetic (angular velocity set at zero) dynamometers can be used to measure static muscle strength. Several studies (58–60) have shown good intra- and interrater reliability for the hand-held dynamometer in various patient populations. The simplicity and portability of these devices make it an attractive clinical instrument.

The validity of the strength measurement depends on the activation level of the nervous system. Merton (61) introduced the use of electrical stimulation superimposed on a static maximal effort in an attempt to activate directly those motor units and muscle fibers not stimulated by the voluntary effort of the subject. The stimulus is applied to the motor nerve during the MVC, and if the force increases, it indicates a suboptimal activation of motor units by the central nervous system. This is often referred to as central activation failure (CAF) (62). In many studies, single impulse stimulation has been used to detect CAF. It has been shown that high-frequency maximal train stimulation may improve the detection of CAF during static (isometric) knee extensions (57,62). This may be important in the clinical assessment of weakness, as it may distinguish weakness caused by CAF from that due to muscle wasting. This, in turn, could have very important implications for the design and evaluation of effective muscle-strengthening exercise therapy. The simultaneous assessment of CAF and muscle mass is advantageous and will facilitate the identification of the mechanism underlying muscle weakness in a particular patient. It is generally considered that healthy men and women even above the age of 70 have the ability to fully activate their muscles during an MVC (63,64).

**Repetition Maximum**

During the course of physical rehabilitation and research studies involving exercise training, muscle strength is frequently measured using the one repetition maximum (1 RM) method. In the case of the extensors of the knee, DeLorme (65) defined the 1 RM as the maximum weight that can be lifted with one repetition with the knee going into complete

<table>
<thead>
<tr>
<th>TABLE 3.2</th>
<th>Scale for Manual Muscle Testing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Numeric Scale</td>
<td>Descriptor</td>
</tr>
<tr>
<td>0</td>
<td>Zero</td>
</tr>
<tr>
<td>1</td>
<td>Trace</td>
</tr>
<tr>
<td>2</td>
<td>Poor</td>
</tr>
<tr>
<td>3</td>
<td>Fair</td>
</tr>
<tr>
<td>4</td>
<td>Good</td>
</tr>
<tr>
<td>5</td>
<td>Normal</td>
</tr>
</tbody>
</table>
Second), reliability is often considered poor. The comparison of strength values among the different brands of isokinetic dynamometers appears to be less valid. The spectrum of indicators of muscle function provided includes measurements of muscle work (integrated area under the curve), endurance (see later), agonists/antagonists ratios, and differences between sides that can be used to assess for muscle imbalances and asymmetries. On the other hand, the devices are expensive, the isokinetic nature of the muscle action does not allow for direct comparisons with daily activities, and the range of angular velocities does not extend to the speed of many sports and daily actions.

Finally, measurements of isokinetic strength at various angular velocities have been used to study, in vivo, the whole muscle torque-velocity curve and the effects of strength training on muscle contractile behavior (73–75). These studies require the use of torque at multiple given joint angles rather than peak torque over the continuum of available angles in order to control for the effect of muscle length on force development.

**Isokinetic Strength**

A number of isokinetic devices have been used to measure muscle strength in research laboratories and rehabilitation clinics (see Fig. 3-2). Although several muscle groups have been studied with these devices, most of the information concerns the knee, ankle (Fig. 3-8), and shoulder muscle groups. The lever arm is aligned with the axis of rotation of the joint to be tested, and proximal and distal segments are stabilized using Velcro straps limiting the contribution of agonists. The tester sets the angular velocity (available range 0 to 450 degrees per second or 0 to 7.9 rad per second), and the research subjects or patients are usually asked to complete three to five maximal repetitions, with the maximal torque used as a measurement of strength.

Individual brands of isokinetic dynamometers have been shown to be reliable (48,49,51,68–70), including measurements of isokinetic eccentric strength (49,69) and strength in patients with Duchenne or Becker muscular dystrophies (71) or stroke (72). Reliability is usually best at low angular velocities and gradually decreases with increased angular velocity. At very high velocities (e.g., above 180 degrees per second), reliability is often considered poor. The comparison of strength values among the different brands of isokinetic dynamometers appears to be less valid. The spectrum of indicators of muscle function provided includes measurements of muscle work (integrated area under the curve), endurance (see later), agonists/antagonists ratios, and differences between sides that can be used to assess for muscle imbalances and asymmetries. On the other hand, the devices are expensive, the isokinetic nature of the muscle action does not allow for direct comparisons with daily activities, and the range of angular velocities does not extend to the speed of many sports and daily actions.

Finally, measurements of isokinetic strength at various angular velocities have been used to study, in vivo, the whole muscle torque-velocity curve and the effects of strength training on muscle contractile behavior (73–75). These studies require the use of torque at multiple given joint angles rather than peak torque over the continuum of available angles in order to control for the effect of muscle length on force development.

**Muscle Power**

Leg extensor power could be a very important measurement in sports as well as in rehabilitation because power is more relevant for time-dependent/time-critical activities than strength. Power has been shown to correlate with gait speed, time to rise from a chair, stair-climb time, and self-reported disability (76,77). In addition, power declines with aging at a faster rate than strength (78). The maximum leg extension power output produced by both legs and each leg separately can be estimated using a power rig bench or resistance training machines. The power rig bench measures force and velocity...
of leg movement, whereas the test using resistance training machines is based on the performance of one repetition at a percentage of the 1 RM. Peak power has been reported to occur usually at 70% of the 1 RM using the resistance training machines (20). Several test trials are usually performed with a 45- to 60-second rest between trials and the maximum value recorded. The coefficient of variation of this method is 6% (79).

Other investigators have designed power tests that combine an eccentric muscle action followed by a concentric action of the same muscle group with a very brief static action in between (80). This is a test of the stretch-shortening cycle (81). A mat or platform that is connected to a timer is used to measure flight time (time spent off the ground) during a vertical jump from which jump height can be calculated. The jump is preceded by a countermovement or semisquat position. The output is fed into a computer for analysis and calculation of muscle power.

**Muscle Endurance (and Fatigue)**

The definitions of **endurance** and **fatigue** vary with the source. For the purpose of this chapter, endurance will be defined as the "time limit of a person's ability to maintain either an isometric force or a power level involving combinations of concentric and/or eccentric muscular actions" (24). Endurance can be measured in seconds. Fatigue, on the other hand, is the inability to maintain a given level of force output (82). Alternatively, **neuromuscular fatigue** can be defined as "any reduction in the force-generating capacity of the total neuromuscular system" and can be due to factors that affect muscle fibers, the neuromuscular junction, and/or the nervous system (83). In humans, fatigue results in the loss of voluntary force and power production (84), a reduction in electrically stimulated maximal force and the maximum rate of force development (85), altered recruitment pattern of muscles (86), impaired neuromuscular performance in activities such as jumping (87), and even a reduction in sensory function such as the acuity of the joint movement sense (88). Methods used to measure these effects could be good indicators of the presence and degree of fatigue in younger and older populations (89).

In the context of the study of muscle performance, endurance could be divided into central (cardiovascular) and peripheral (local muscular) types. Several physiologic systems are involved in the expression of endurance and the avoidance of fatigue, including the nervous (central and peripheral), cardiovascular, hormonal, and metabolic systems. It is likely that the contribution of each system to endurance and fatigue depends on factors such as the type of activity, environmental conditions, nutritional status, and fitness level. Measurements of muscle endurance represent an integrated evaluation of several of these systems.

The basic mechanisms underlying muscle fatigue have not been precisely defined, although our understanding of this phenomenon has been substantially enhanced by recent research. Several excellent reviews summarize the available literature and discuss the contribution to fatigue of many of the components of the neuromuscular system (90–92). Clearly, each and every step along the chain of transmission of the electrical and chemical signals from the brain to the actin–myosin cross-bridges and the biochemical pathways that supply fuel for muscle action is a potential site of failure. The development of fatigue is significantly influenced by the fiber type composition of skeletal muscle as the effector organ (93). Other relevant factors include central and peripheral mechanisms (94), such as changes in the excitability of the motor cortex or inadequate neural drive upstream of the motor cortex (95,96). This type of fatigue is of particular significance in neurologic rehabilitation, as it is considered prevalent in various nervous system disorders such as multiple sclerosis and stroke (97).

At the cellular level, the role of changes in the ADP/ATP ratio and the concentration of inorganic phosphate have received strong support (98,99) challenging the traditional view that accumulation of lactic acid and hydrogen ions is the main biochemical correlate of muscle fatigue. Others have highlighted the importance of cation pumps such as the Na+/K+ ATPase (100) and the Ca2+ ATPase (101) in maintaining muscle excitability and preventing muscle fatigue. Cellular factors may also play a role in muscle fatigue in disorders of the central nervous system such as multiple sclerosis and amyotrophic lateral sclerosis (102,103). It is likely that multiple mechanisms interact under various conditions, resulting in the clinical phenomenon called muscle fatigue.

**Absolute and Relative Endurance**

One simple test of dynamic endurance is to count the number of repetitions that a subject/patient can do with a given load. Similarly, a test of static endurance could measure the length of time that a given level of force can be maintained. Since both tests use an absolute load or level of force, stronger subjects/patients will be tested at a relatively lower percentage of their strength than weaker subjects. For example, a dynamic test conducted using a 40-kg load represents 40% and 80% of the strength of individuals with a 1 RM test of 100 and 50 kg, respectively. Tested at a lower percentage of his or her strength, the stronger individual will demonstrate better endurance.

To control for the effects of muscle and body size, a similar testing paradigm can be used, but based on a percentage of the individual's strength. This relative endurance test requires the performance of as many repetitions as possible using, for example, 50% of the dynamic strength of the individual. It can also be done by asking the individual to maintain a relative level of static force for as long as possible. The relative endurance test limits the influence of body or muscle size on the measurement and may allow for a better way to compare different groups of patients.

**Fatigue Indices**

Isokinetic devices have been used to quantify muscle fatigue and endurance by measuring relative losses in peak torque and/or work during a standardized test. For example, the
ratio between peak torque or work performed during the first and last three or five repetitions of a series of repetitions, usually 25 to 100, depending on the muscle group under study, has been used as an index of muscle fatigue. Since the calculation of work performed is based on the integration of the area under the curve, it is important to make sure that, during the test, the joint moves through the entire predetermined range of motion. Studies that have measured both fatigue and strength reproducibility using isokinetic dynamometers have found that fatigue tests are less reliable than strength tests (104–106). In a study of the reliability of an isokinetic fatigue test for concentric ankle dorsiflexion, it was shown that fatigue indices using decreases in peak torque were more reliable than an index based on relative decreases in work, but that all indices showed acceptable to good reliability (107).

A similar approach has been used by other investigators, but during static muscle actions. A combination of the percent reduction in force over a given period of time and the average augmentation with electrical stimulation superimposed during contractions has been proposed as a clinically useful method of quantifying muscle fatigue (108).

Spectral Analysis of the Surface Electromyographic Signal
Spectral analysis of the surface electromyographic signal is based on the idea that, during the performance of static muscle actions, a reduction of the median frequency correlates with the drop in force and therefore is an indirect indicator of muscle fatigue (109). The median frequency reflects muscle fiber conduction velocity and motor unit recruitment and therefore is influenced by the fiber type composition of the muscle. Some authors have proposed that a shift in the mean power frequency of the EMG is actually a selective indicator of fatigue of the fast-twitch motor units (109) and that subjects with higher concentrations of fibers expressing type II myosin heavy chain isoform displayed greater reductions in median frequency of EMG and muscle fiber conduction velocity during exercise (110).

Another approach based on the study of changes in the EMG signal during voluntary contractions has been proposed by Merletti and Roy (111). They studied the rate of spectral compression of the surface EMG signal during the early phases of static voluntary muscle actions of the tibialis anterior. The slope of the median frequency during the first 30 seconds correlated with endurance time. The authors proposed the use of this technique in the clinical setting in part because the test does not require the use of muscle actions to failure.

Nuclear Magnetic Resonance
Nuclear magnetic resonance (NMR) is used to study the metabolic correlates of muscle fatigue at rest and during exercise (Fig. 3-9). Changes in intracellular pH, concentrations of inorganic phosphate, phosphocreatine (PCr), ADP, and ATP resulting from the metabolic demands imposed on active muscles can be monitored in real time at rest, during exercise and fatiguing conditions, and during the recovery period (112). The speed at which a muscle recovers from these metabolic alterations is a good indicator of the fitness level of the individual. An interesting point is that the accumulation of these metabolites alters the conduction of the action potential and explains the changes in the spectral content of the EMG signal mentioned earlier. This technique has the advantage of being noninvasive. However, it is expensive and not readily available in the clinical setting.

Other Tests of Muscle Properties Related to Fatigue and Endurance
Near Infrared Spectroscopy and Microdialysis
Near infrared spectroscopy (NIRS) is a relatively new, noninvasive research instrument to measure oxygen (O$_2$) through tissues up to 10 cm in depth (113). NIRS can be used to investigate the delivery and utilization of oxygen in response to exercise and in the clinical setting to assess circulatory and metabolic abnormalities (114). Quantitative measures of blood flow are also possible using NIRS.

Microdialysis is invasive research technique that has allowed mechanistic investigations to be performed in human skeletal muscle. The microdialysis catheter has been equated to an artificial blood vessel, which is introduced into the tissue. By means of this “vessel,” the concentrations of compounds in the interstitial space of skeletal muscle (as well as other tissues) can be monitored. A number of important observations on the in vivo regulation of lipolysis, carbohydrate metabolism, and blood flow in human skeletal muscle and adipose tissue have been made using microdialysis (115).
STRUCTURAL CHARACTERISTICS OF SKELETAL MUSCLE

The Relationship Between Function and Structure

The relationship between the structure and the function of biological systems has received significant scientific attention. Ewald Weibel, the great Swiss morphologist, concluded his book on the structure and function of the mammalian respiratory system that, “I have become convinced not only that structure determines function, but that functional demand also determines structural design, be it through evolution or by modulation of design features” (116). The importance of function as a determinant of form has also been emphasized by Russell et al. (117). In their review, they demonstrated how the functional demands imposed on muscle cells result in adaptations in muscle cell shape, size, and force production.

Based on the previous paragraph, a discussion about muscle function cannot be comprehensive unless some of the structural and architectural features of skeletal muscle are analyzed. The level of force and power generated by an active muscle is influenced by muscle size, fiber angle of insertion, and sarcomere length (118,119). Muscle length contributes to shortening velocity, and muscle fiber type composition is an important determinant of speed and endurance. We will now discuss some of the measurements of muscle structure that are used in rehabilitation clinics and research laboratories.

Body Composition: Estimating and Measuring Muscle Mass in Humans

The human body is composed of various types of tissues (including skeletal muscle) with different chemical consistency. The study of body composition allows a determination of the amount and anatomic quality of these components. Body composition is associated with health status and functional capacity, among other outcomes. The study of body composition can be accomplished using different compartment models (Fig. 3-10) (120). Skeletal muscle and its constituents, although sometimes measured indirectly, are important components in each of these models. The chemical composition of muscle has been described as approximately 75% water and 20% protein, with the remaining 5% made up of inorganic salts and other substances that include high-energy phosphates, urea, lactic acid, calcium, magnesium, phosphorus, enzymes, potassium, sodium, chloride, amino acids, fats, and carbohydrates (121).

The traditional gold standard in the study of human body composition, hydrodensitometry or underwater weighing, and the more recent air displacement (plethysmography) are based on a two-compartment model. This model divides the body into fat mass (FM) and fat-free mass (FFM). In this context, the use of the term fat-free mass is thought not to be very accurate, since the nonfat component includes a small percentage of essential fat stores within the central nervous system, bone marrow, and internal organs. Thus, LBM is a more appropriate designation. LBM includes essential fat, muscle, and bone. Hydrodensitometry estimates body density (122), which is then used to calculate FM with the Siri equation, assuming a constant density for FM and LBM (123):

\[
\text{Percent body fat} = \frac{495}{\text{density}} - 450
\]

where density = mass (in air)/loss of mass in water (or mass in air mass in water). Accurate determination of density requires correction for water temperature. Finally, LBM is obtained by subtracting FM from the total body mass.

Elements at the atomic level can be quantified using neutron activation techniques (124,125). This is relevant to our topic of discussion because, as mentioned earlier, muscle is an important source of nitrogen. In addition, most of the potassium present in the body is intracellular, and a large amount is present in myofibers. The use of whole-body counters, special instruments capable of detecting radiation from the decay of the naturally occurring radioactive isotope of potassium, allows quantification of cell body mass and, indirectly, lean body and muscle mass (124,125).

Techniques such as dual-energy x-ray absorptiometry (DEXA) are used to quantify bone, fat (both subcutaneous and visceral) mass, and skeletal muscle mass (126,127). Total body and appendicular estimates can be obtained (126) at a lower cost and radiation exposure than with computerized tomography (CT). The most accurate in vivo methods of measuring muscle mass and segmental body composition are imaging techniques such as CT and magnetic resonance imaging (MRI) (see later).
**Muscle Size**

A reasonable, but not always accurate, estimate of muscle size at a given limb level can be obtained in the clinic by measuring limb circumference with a metric tape. However, these measurements include fat, bone, and other noncontractile tissues. Furthermore, such measurements cannot be corrected for changes in limb circumference related to fluid shifts or tissue edema. In pathologic conditions where muscle is replaced by connective tissue and/or fat, such as in certain myopathic conditions, measurements of limb circumference can be very misleading.

Reliable estimates of muscle cross-sectional area can be obtained with modern techniques such as ultrasound (128,129), CT (130), or MRI (130–132). Ultrasound has been used extensively because of its simplicity, its low cost, and the fact that it does not involve exposure to radiation. However, it does not provide an optimal distinction of the borders between the various tissues. Another application for ultrasound techniques has been the in vivo study of the architecture of the muscle tendon complex (133). Measurements of tendon length, muscle fascicle length, pennation angle of muscle fibers, tendon elongation during muscle contraction, tendon strain, and related properties have been done in vivo in humans. These measurements can help us understand muscle function. For example, because muscle fibers consist of sarcomeres in series, the muscle fiber length indicates the velocity potential of a muscle during contraction. The longer fascicles in the vastus lateralis muscle result in faster velocity than that occurs with the medial gastrocnemius. However, the tendon of the medial gastrocnemius has a higher compliance and therefore a greater potential to store elastic energy (134).

Both CT and MRI have been widely used to obtain valid (using cadaver measurements as the standard) and reliable measurements of total thigh, muscle, subcutaneous fat, and bone cross-sectional areas in health and disease (130,135). In both CT- and MRI-based methodologies, anatomic cross-sectional area refers to the area measured from a single cross section of the muscle. The measurement is made perpendicular to the long axis of the muscle. Because in pennate muscles like the quadriceps femoris fibers do not run parallel to the longitudinal axis of the muscle, these measurements underestimate the true muscle area (136; Fig. 3-11). On the other hand, if measurements of muscle volume obtained by MRI are combined with estimates of muscle fiber length from cadavers, physiologic cross-sectional area can be calculated with reasonable accuracy (137,138).

Both CT and MRI have been shown to provide valid estimates of muscle cross-sectional area (130). The higher definition MRI provides better area estimates than CT images. The latter tends to systematically overestimate anatomic cross-sectional area by 10% to 20% (139). Both MRI and CT have been shown to have high intrarater and interrater reliability and are capable of detecting exercise-training induced adaptations in skeletal muscle (74,140,141).

**FIGURE 3-11.** The influence of pennation angle on the thickness of muscle and estimations of muscle size. CSA, anatomic cross-sectional area; PCSA, physiologic cross-sectional area. These two muscles have the same number of fibers of the same thickness. By increasing the pennation angle $\alpha$, thickness of the muscle (CSA) belly decreases. The pennation angle allows more of the muscle mass to be closer to the joint, reducing the resistance to movement (136).

T1-weighed MRI allows muscle tissue and noncontractile tissue components (connective tissue, subcutaneous and interstitial adipose tissue) to be separated (130,140). Cross-sectional images of the human thigh and leg obtained using MRI are presented in Figure 3-12 (132). Computer-based image analysis systems can be used to quantify the composition of muscle, fat, and connective tissue by assigning different density values to the different tissue components. Muscle attenuation, a function of tissue density and chemical composition, has been shown to relate to its lipid content (141). Furthermore, since disease and exercise training may alter both the contractile and noncontractile components of the limb, it is very useful to be able to quantify changes in these two compartments (140).

MRI has also been used to assess the level of activation of different muscles or parts of muscles during volitional exercise and neuromuscular electrical stimulation (142,143). MRI can provide a noninvasive way to quantitatively evaluate and localize muscle activity. The underlying mechanism that explains this phenomenon is the accumulation of osmolites (phosphate, lactate, sodium) in the cytoplasm during exercise, resulting in the influx of fluid and an increase in muscle T2 relaxation. A practical application of this technique is the use of the T2 relaxation times to evaluate muscle activation during specific exercises. (For a detailed discussion of radiologic techniques, see Chapter 6.) For example, a recent study demonstrated that both the “empty can” and the “full can” exercises activate the
supraspinatus muscle to a larger extent than the horizontal abduction exercise (144). This information could be used to design the optimal combination of exercises for strengthening rehabilitation programs. Thus, only those exercises that activate a large proportion of the injured or weak muscles are included in the rehabilitation program.

When comparing different patient populations, subjects with different levels of physical fitness, or measurements done before and after exercise training, we must take into account the fact that force varies with size (Fig. 3-13) and that larger individuals will be stronger if absolute strength is used as an outcome. These comparisons should only be made after strength is adjusted for differences in muscle physiologic cross-sectional area or volume or for the amount of noncontractile tissue components. The ratio between force generated and measurements of muscle area is known as specific force and has been used as an index of "muscle quality" (138,145–147). A simple ratio, however, may not be the best way of adjusting muscle strength, and more sophisticated statistical treatment may be needed (148–150).

Muscle Biopsy Techniques
Despite recent advances in noninvasive muscle-imaging techniques, analysis of small samples of human skeletal muscle is still the only way to determine various changes in the fiber population as a result of an intervention or to investigate the underlying pathophysiologic mechanisms of muscle

**FIGURE 3-12.** Cross-sectional images of the human thigh (A) and leg (B) obtained with MRI and a schematic drawing of the muscles of the leg. Note the clear difference in density of the muscle (contractile) and fat and connective (noncontractile) tissues (132).

**FIGURE 3-13.** Relationship between muscle strength and size in humans. Note that the correlation is not 1, suggesting that factors other than muscle size (e.g., motor unit activation, fiber type distribution, muscle fiber quality) explain a significant percentage of the variability in muscle strength among subjects or patients.
dysfunction. Muscle biopsies for various morphologic and biochemical studies can be obtained with the open and percutaneous (needle or semiopen conchotome) biopsy techniques (Fig. 3-14). The needle biopsy technique was introduced by Duchenne in 1868 (151) and has been successfully adapted or modified by others (152–156) for the study of neuromuscular and systemic diseases affecting skeletal muscle (157–159), adaptations to exercise training (160–162), and aging (163). An alternative technique is the semiopen conchotome technique (155,164), which usually yields larger samples than the needle technique. With either technique, the procedure is done under local anesthesia and is usually well tolerated by research subjects and patients. Ultrasound (165,166) and MRI (167) have been used to guide the placement of the biopsy needle to the area of interest in the muscle under study. The skeletal muscles most frequently biopsied are the vastus lateralis, lateral gastrocnemius, triceps surae, tibialis anterior, and biceps brachii.

It must be noted that a biopsy sample represents a very small fraction of all the fibers in a large muscle such as the vastus lateralis. A biopsy sample from this muscle may contain several hundred to a few thousand fibers, compared with hundreds of thousands in the whole muscle. Thus, because there is significant intramuscular variability in the distribution of fiber types, more than one sample may be needed to obtain a good estimate of the fiber type proportion (168). The determination of fiber size and capillarity, however, appears to be less problematic. It has been suggested that one biopsy sample may suffice (169–171), although several small samples are advantageous when muscle fiber areas are determined (172).

**Muscle Fiber Size and Type Distribution**

Histologic and histochemical techniques have been used for more than half a century to study muscle cross sections under light microscopy (173,174). Over the last decades, immunocytochemical techniques have been introduced. These techniques use chemical reagents (stains) that react with components of the muscle fiber, allowing the visualization and assessment of various structural and functional muscle characteristics such as fiber size and shape, fiber type distribution, capillary supply, enzyme activity, nuclei distribution, satellite cells, intracellular carbohydrates and lipid stores, amount of intercellular connective tissue, degree of fat infiltration, presence of intracellular structures such as inclusion bodies, and others.

Table 3-3 summarizes the most common staining procedures of human muscle biopsies. Many of these techniques are used as part of the clinical investigation of muscle biopsies (e.g., to diagnose a specific neuromuscular disorder). Other procedures are used to investigate pathologic changes in general, and yet others are only used for research.

Muscle biopsy cross sections can be stained with hematoxylin and eosin for the evaluation of the general morphology of muscle fibers, including its size, shape, and location of nuclei. The identification of capillaries in the muscle biopsy cross section can be accomplished with the periodic acid-Schiff amylase stain or immunocytochemical procedures with antibodies against laminin or ulex europaeus agglutinin I (UEA) lectin (175). This is an important property of skeletal muscle, given the fact that capillarity changes significantly with exercise training, disease (e.g., peripheral vascular disease), and immobilization. A qualitative assessment of various oxidative and glycolytic enzymes can be done with specific enzymatic stains such as succinate dehydrogenase (SDH), NADH tetrazolium reductase (NADH-TR), $\alpha$-glycerophosphate dehydrogenase (GPDH), cytochrome C oxidase (COX), and phosphofructokinase (PFK). Finally, qualitative and quantitative estimates of lipid and carbohydrate content are possible via histochemical and biochemical methods. The glycogen depletion technique has been used for several decades to identify the pattern of activation of motor units in various types of exercises; the reader is referred to other publications on this subject for more details (176).

The mATPase stain is one of the most widely used to visualize the fiber population and evaluate changes in the size and proportion of the different fiber types following an intervention such as detraining (Fig. 3-15). Based on the mATPase activity, the human muscle fiber population is commonly classified into types I, IIA, IIB, and IIC (177). This original mATPase-based semiqualitative classification of muscle fibers has been expanded, and seven different human muscle fiber types are now being identified: I, IC, IIC, IIA, IIB, IIAC, and IIC.
TABLE 3.3  Examples of Staining Procedures of Human Muscle Biopsies

<table>
<thead>
<tr>
<th>Histologic stains</th>
<th>H&amp;E Stains nuclei and cytoplasm, used to assess general structure, location of nuclei, inflammation, regeneration</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Periodic acid-Shiff (PAS) Glycogen, shows abnormal accumulations of glycogen and related metabolic disorders</td>
</tr>
<tr>
<td></td>
<td>Oil red O or Sudan black Lipid droplets</td>
</tr>
<tr>
<td>Histochemical stains</td>
<td>Modified Gomori trichrome Stains mitochondria, nuclei, used to visualize ragged red fibers with excess mitochondria</td>
</tr>
<tr>
<td>Histochemical stains</td>
<td>Myosin ATPase at pH 4.3, 4.6, and 9.4                         Used to visualize and differentiate muscle fiber types</td>
</tr>
<tr>
<td>Oxidative enzymes</td>
<td>NADH Stains intermyofibrillar material such as mitochondria, sarcoplasmic reticulum, t tubules, target fibers</td>
</tr>
<tr>
<td></td>
<td>SDH Stains mitochondria (complex II, encoded entirely in nuclear DNA), visualizes subsarcolemmal and intermyofibrillar distribution, indicates excess mitochondrial activity</td>
</tr>
<tr>
<td></td>
<td>COX Stains mitochondria (complex IV), used to detect mitochondrial myopathies that lack the enzyme</td>
</tr>
<tr>
<td>Other enzymes</td>
<td>This enzyme is active in the glycolysis and absent in type VII glycosgenosis</td>
</tr>
<tr>
<td></td>
<td>PFK Indicates excess of glycolytic activity or absence, as in McArdle’s disease (type V glycosgenosis)</td>
</tr>
<tr>
<td>Immunocytochemical stains</td>
<td>UEA Stains capillaries</td>
</tr>
<tr>
<td></td>
<td>Dystrophin Duchenne and Becker muscular dystrophy</td>
</tr>
<tr>
<td></td>
<td>Desmin and vimentin Cytoskeletal proteins, used to visualize congenital myopathies and myopathies with cytoplasmic bodies</td>
</tr>
<tr>
<td></td>
<td>Lymphocyte markers Used to visualize mononuclear inflammatory cells</td>
</tr>
<tr>
<td></td>
<td>N-CAM Used to visualize regenerating and denervated fibers</td>
</tr>
</tbody>
</table>

and IIB (178,179). Perrie and Bumford (180) refined the electrophoretic techniques and separated the human myosin heavy chains (MHC) into three isoforms: I, IIa, and IIx, which have been shown to correlate with the three major mATPase-based fiber types (181–183). It is now known that the mATPase-based type IIB fibers correspond to MHC IIx identified by gel electrophoresis (184,185). Studies of single muscle fibers have also shown that fibers can contain a mixture of MHC isoforms. The mATPase types IC, IIC, and IIAC coexpress the MHC I and IIa to a varying degree, whereas type IIAB fibers coexpress MHC IIa and IIb/x (179). As the enzyme histochemical and electrophoretic techniques can yield different information, it is recommended that both techniques are used when the muscle fiber type population is quantified (186).

STUDIES OF SINGLE MUSCLE FIBERS

The percutaneous biopsy technique has been used to obtain muscle fiber segments for the study of the morphologic, contractile, and biochemical properties of single human skeletal muscle fibers (187,188). These single fibers are permeabilized with high glycerol concentrations and a detergent solution, resulting in breaks in the sarcolemma and the sarcoplasmic reticulum. The absence of these barriers permits the rapid diffusion of calcium into the cell and its interaction with troponin C, triggering the cascade of events that lead to force generation. Before these experiments are done, the fibers are attached to a force transducer that measures force generation during activation and a servomotor that controls the sarcomere length (Fig. 3-16). An image analysis system is used to measure sarcomere length, muscle fiber length, diameter, and depth; the latter two are used to calculate fiber cross-sectional area.

Measurements of maximal static force, peak power, and shortening velocity are obtained by activating the fiber with calcium (150,189). Using a different experimental protocol, the force-velocity and force-power relationships of the single fibers
PART I | PRINCIPLES OF ASSESSMENT AND EVALUATION

FIGURE 3-16. Light microscopy image of a human single muscle fiber. The fiber is attached to a force transducer and a lever system to control its length. Notice the well-preserved striation pattern.

can also be studied. When contractility studies are completed, the fiber segment is submitted to protein gel electrophoresis to determine the expression of myosin heavy chain isoforms and thereby the fiber type (Fig. 3-17). This electrophoretic technique has the advantage over the APTase stain of allowing the identification of hybrid fibers, fibers expressing simultaneously two or more myosin heavy chain isoforms.

The single muscle fiber method complements the techniques used to measure in vivo muscle function by making it possible to study the contractile behavior of single fibers in the absence of confounding variables such as the influence of the nervous system, the presence of intercellular connective tissue, the heterogeneity of muscle fiber type distribution, and the connectivity to the tendon-bone interface. Adjusting measurements of contractility for differences in muscle fiber size has been used as an index of muscle fiber quality.

FUNCTIONAL TESTS OF MUSCLE PERFORMANCE

In a clinical setting, it is sometimes not practical or efficient to evaluate muscle function with expensive devices and time-consuming laboratory-based tests. Furthermore, in research and patient care, there is a need to measure performance in tasks that are more similar to activities of daily living (ADLs) and instrumental activities of daily living (IADLs) than many sophisticated laboratory tests. These two considerations have resulted in the development of simple, inexpensive, and easy-to-administer tests of neuromuscular function or performance. These measurements do not represent direct evaluations of skeletal muscle or any other isolated physiologic characteristic but rather an integrated response of the human body to the demands of a particular task. However, muscle strength and power correlate well with test performance scores. These tests are described in various publications together with normative data (20,190,191), and some of them will be described later.

Chair Stand Time

In general, the chair stand time test (190,191) requires a chair with arms and a seat placed against a wall for support and safety. Subjects should sit all the way back in the chair with their back against the back of the chair. The feet should be placed so that the knee is in a 90-degree angle. The patient or research subject is instructed to stand up and sit back down as fast as possible for five or ten repetitions with the arms crossed at the chest. Time is measured with a stopwatch to the nearest 0.01 second and stopped after the last stand. The subjects should come to a complete stand each time and completely sit down with their backs against the back of the chair each time. This test can be considered an indirect index of knee extensor function, since this muscle group produces 72% of the power needed to rise from a chair (15).

Stair-Climbing Power

A standard stair flight with handrails on each side is used for the stair-climbing test (192). Subjects are instructed to ascend stairs as quickly as possible using the handrail and/or an assistive device if necessary. The test is started with both of the subject’s feet on the bottom step. The subject is instructed to climb up the steps as quickly as possible. A stopwatch is used and stopped when both feet are planted at the top step. Time is recorded to the nearest 0.01 second, and the best of two trials (separated by a 2-minute rest period) is taken. Power, in watts, is calculated using the following equation:

\[
\text{Power (watts)} = \frac{\text{body weight (kg) \times vertical distance (m)}}{\text{time (s) \times 60}} \times 6.12
\]

(Note that 6.12 kg·m/min = 1 watt.)
Habitual and Maximal Walking Speed

Walking speed correlates with strength of the ankle dorsiflexors and knee extensors (10,193–195) and leg extension power (195). This relationship has been shown to be nonlinear (10,196). In other words, the decline in gait time with increases in strength is curvilinear, and there is a threshold value below which the relationship is lost. In longitudinal studies, baseline strength of the knee extensors has been shown to be a good predictor of 2- and 4-year decline in walking velocity (193). Walking speed is also correlated with isokinetic knee extensor and flexor torque in chronic stroke patients; strength for the paretic limb explains up to 50% of the variance in gait performance (197).

Walking velocities can be measured to the nearest 0.01 second as the mean of two trials using a stopwatch and a set distance or a gait speed monitor. The subject or patient will start the test with feet parallel and both heels and toes on the floor. For the habitual walking speed test, patients or subjects are instructed to walk at their normal or comfortable speed. During the maximal walking speed test, they are required to walk as quickly as possible. Velocity is recorded after subjects walked a given distance to control for individual differences in acceleration. Measurements of walking speed have been shown to be highly reliable in a number of patient groups (198).

The 6-Minute Walk Test

The 6-minute walk test requires that subjects walk for 6 minutes covering as much ground as possible at a pace that would allow them to talk without becoming short of breath (199). The test may be conducted in a circular track or on a level course. The subject is allowed to rest but instructed to continue walking if possible. The tester times the walk with a stopwatch. After 6 minutes, the distance walked is measured.

Although sometimes considered a test of aerobic capacity, endurance, and survival in patients with heart disease (200), it has been shown that performance in the 6-minute walk test correlates with lower-extremity muscle strength and power (197, 201). Performance in this test depends on multiple physiologic, psychologic, and health factors (202). Muscle strength, balance, medication use, and age explain the largest proportion of the variance in the performance of this test.

When chronic illness has resulted in severe functional losses, exercise training may be used as a form of recreation. In the steep part of the curve, a relatively small loss of muscle function could result in disproportionate reduction in performance. Thus, even minor adaptations to exercise training could preserve functional capacity and independence. The minimal level of muscle function needed to achieve this goal remains to be determined.

CONCLUSION

Understanding the role of skeletal muscle in health and disease is of paramount importance for those interested in rehabilitation medicine. A wide range of techniques and methods have been developed to assess muscle function and structure. Becoming proficient in their applicability and use is relevant to patient care, education, and research in physical medicine and rehabilitation.

REFERENCES

PART I


CHAPTER 3 | ASSESSMENT OF HUMAN MUSCLE FUNCTION


INTRODUCTION

The ability to perform high quality electrodiagnostic studies is a critical skill for the practicing physiatrist. Electrodiagnosis is an important and helpful extension of the physical examination and can detect minor abnormalities when physical examination cannot. Moreover, the electrodiagnostic examination has several other key functions in the assessment of peripheral nervous system injuries. These techniques provide insight into the pathophysiology of a nerve lesion (such as demyelination or axon loss). By quantifying the degree of axon loss, one can estimate prognosis for recovery of lower motor neuron lesions. Finally, electrodiagnostic methods enable one to quantitatively track disease processes and complement the qualitative estimates afforded by the clinical examination.

PREPARING FOR THE ELECTRODIAGNOSTIC EXAMINATION

The electrodiagnostic examination needs to be interpreted in the context of the clinical presentation. Simply performing nerve conduction studies or needle electromyography (EMG) without understanding the patient’s presentation or developing a differential diagnosis does a disservice to the patient, as the diagnosis will often be incomplete or incorrect.

For these reasons, there are several key steps to prepare for the electrodiagnostic examination, including review of referral materials, eliciting the patient’s history, performing a directed physical examination, developing a differential diagnosis, and putting together a plan for the electrodiagnostic evaluation.

Review of referral materials: It is critical that the electrodiagnostic consultant starts by reviewing the question from the referring physician so that he or she knows what is being asked for. Without reviewing this information first, it is easy to be mislead into another diagnostic pathway and never answer the referring physician’s question. If there is any uncertainty about what is being asked for, it is often helpful to talk to the referring physician. If it is not possible to address the question being asked by the electrodiagnostic examination, one should have that discussion with the referring physician before starting the electrodiagnostic study.

Patient history: A focused patient history is very important to allow one to establish a differential diagnosis before heading into the physical examination and then the electrodiagnostic evaluation. The history should start off with the chief complaint, but should also include exacerbating or relieving factors and past medical history. It is often useful to ask what medications the patient is taking so that one is not later surprised by easy bleeding (in the setting of warfarin usage) or by a polyneuropathy (e.g., diabetes). Parts of the past medical history may be especially contributory to the differential diagnosis such as diabetes, history of cancer, or toxic exposures. When one starts to consider diseases that may be hereditary, it is always wise to ask about any family history of neuromuscular problems.

The physical examination needs to be focused and adapted to the individual presentation of the patient. Typically, it is useful to assess muscle strength not only in the involved body part but also in bilateral upper and lower limbs. One will want to detect even subtle weakness in the patient before starting the study. Since most patients coming to the EMG laboratory will have at least 4/5 strength, it is usually not possible to fully test the ankle plantarflexors or dorsiflexors by manual muscle testing. It is much more effective to have the patient walk on his or her heals and toes and do other functional tests as appropriate (e.g., squatting).

One should get a rough sense of the distribution of sensory loss. Two techniques are particularly helpful in this regard. First, in mild sensory loss, it is often helpful to compare with an area of normal sensation and ask, “If this is 100 (testing normal area), how much is this (testing abnormal area)?” Generally, losses of small percentages are not important. Secondly, if an area of abnormal or absent sensation is clearly established, it is useful to map this out, putting small pen marks where sensation changes from abnormal to normal. When discussing the distribution of sensory loss, it is better to describe it in anatomical terms (e.g., small finger) rather than in root or nerve distribution terms, since one often does not know whether sensory losses are in a peripheral nerve, plexus, root, or other distribution (e.g., ulnar nerve, medial cord, lower trunk, C8 root or spinal cord).

Muscle stretch reflexes are one of the few objective indicators of nerve or muscle disease. Hyperreflexia usually suggests an upper motor neuron etiology, while diminished reflexes or areflexia are more consistent with a peripheral nerve or lower motor neuron process.
Provocative signs can be somewhat helpful but have limited diagnostic accuracy. One may consider Spurling’s sign for cervical radiculopathy, straight leg raise sign for lumbar radiculopathy, Tinel’s, Phalen’s, or flick sign for carpal tunnel syndrome, and other maneuvers. Generally, these provocative signs have a limited specificity (1).

In many settings, where an electronic medical record is present, it is possible to briefly review radiologic findings and other laboratory data before establishing the differential diagnosis. This can be helpful to either expand or appropriately contract the differential diagnosis as needed. One should keep in mind, however, that there will not always be a good correlation between imaging studies and electrodiagnostic findings.

After completing the clinical assessment above (which should usually take no more than 15 to 20 minutes), one should establish the differential diagnosis. This important step is often skipped and can result in overlooking important possible diagnoses or missing significant tests. The examining physician should ideally write down his or her differential diagnosis. One can consider the differential diagnosis starting from most likely to least likely or starting centrally in the nervous system and work peripherally so that important possibilities are not overlooked. It is generally wise to include the diagnosis the referring physician has suggested on your list, since you want to be able to answer his or her question in your final report.

After the differential diagnoses have been articulated, one can then establish a plan for diagnostic testing. One should look at the differential diagnoses and decide what elements of electrodiagnostic study are needed to assess for those possible diagnoses. One may elect to modify this plan as initial information is attained, and in fact, this is often an iterative process. However, without going through the rigor of developing (and writing down) a plan based on the differential diagnosis, one will miss important findings (Fig. 4-1).

**TYPES OF ELECTRODIAGNOSTIC TESTING**

Although there are many techniques that can be employed in electrodiagnostic assessment, this chapter will cover the primary techniques including nerve action potentials (sensory and mixed), motor nerve conduction studies, late responses (F waves, H waves, and A waves), needle EMG, and repetitive stimulation. For each technique, this chapter will discuss the physiology, normal findings, applications, abnormalities, limitations, and the likely encountered sources of error.

**Sensory Nerve Action Potentials and Compound Nerve Action Potentials**

Sensory nerve action potentials (SNAPs) and compound nerve action potentials (CNAPs) both involve electrical stimulation of nerves and recording of the synchronized discharge of axons within the nerve at some distance from the point of stimulation. SNAPs are responses that only involve sensory axons, whereas CNAPs may involve sensory axons, motor axons, or a mixture of both.

The nerve is stimulated supramaximally, meaning that all the axons are electrically stimulated. In practice, this means that stimulator intensity (mA) is gradually increased while monitoring the size of the nerve response. When the response from the nerve gets no larger as the intensity continues to increase, this indicates that all the available axons have been activated and further increases in stimulation will activate no more axons. One should be aware, however, that increasing the stimulator substantially above this level will sometimes activate other nearby nerves, and responses from these other nerves can be recorded from the same recording electrodes (Fig. 4-2) through volume conduction; this is to be avoided.

---

**FIGURE 4-1.** Steps in the evaluation of a patient referred for an electrodiagnostic medicine consultation.

**FIGURE 4-2.** Amplitude of CMAP or SNAP as current is increased. Note that with very high levels of stimulation, amplitude will increase above supramaximal due to co-stimulation of nearby nerves.
One records from the same nerve at a specified or measured distance away from the stimulation and records the synchronized discharge of all the axons that have been activated (Fig. 4-3).

Several items are then measured from the nerve action potential. Most commonly, these include the latency and amplitude of the response. The latency can either be measured to the onset of the response or the peak. The advantage of measuring to the onset is that this represents the fastest conducted fibers, and it also depends less on the distance between active and reference electrodes. The advantage of measuring latency to the peak is that the peak is more reliably determined and less influenced by the sensitivity of the display or the amplifier. In either case, one should have reference (normal) data, which have been derived using the same technique. The amplitude or size of the response can be measured from baseline to peak or from peak to peak, again depending upon how the reference data were collected. Conduction velocity is in units of distance over time, in this case meters per second (m/s), which is the same as millimeters per millisecond (mm/ms). In the case of nerve action potentials, one can measure the distance between stimulation and recording and divide that by onset latency to obtain the velocity in meters per second.

Conceptually, one should consider what factors will influence the latency (or conduction velocity) and amplitude of sensory responses. The latency can be affected by a number of both physiologic and pathologic processes. Cold slows conduction velocity. This effect varies from nerve to nerve but generally is in the range of about 5% per °C. Conduction velocity also changes with age. Initially, children are born with conduction velocities about half of that of adults. As their nerves become more myelinated, these values approach normal values around age 3. Later, in the fifth and sixth decades, velocities start to slow again but not to the degree of that seen in infancy. For this reason, reference values are often adjusted for age.

With respect to pathological changes, loss of myelin will cause slowing of conduction velocity and increase in latencies. When demyelination becomes severe enough, however, conduction block will occur. One might think that such fibers could conduct slowly, as unmyelinated fibers do. However, since myelinated fibers do not typically have sodium channels underneath the areas of myelin, in the setting of intense demyelination, they cannot conduct as unmyelinated fibers do and so conduction is blocked.

The amplitude of the response is also affected by temperature. Cold tends to increase the amplitude of the response if the recording electrodes are over a cold area of the limb. This increase in size is because cold slows both the opening and, even more, the closing of sodium channels. As a result, more charge is exchanged across the axon plasma membrane and a greater size axon potential is elicited. This increased potential is then summated across many axons and the nerve action potential becomes enlarged. Amplitude is also dependent upon the distance of the electrodes from the generator source. The electrical potential amplitude falls off rapidly with distance as one moves further from the nerve. Thus, nerve conduction studies in which the response is close to the skin (such as recording from the fingers) usually generate a bigger amplitude response than when one is recording from proximal areas of the limb (where the nerves are located deeper).

Amplitude of the SNAP or CNAP can be thought of as representing the number of axons that are participating in the response. As one loses axons, the amplitude falls roughly in proportion to the degree of axon loss. In some cases, a reduction in amplitude can result from axon loss occurring between the stimulating and recording electrodes, in which case changes in the potentials would be seen immediately. However, SNAP amplitudes will also be reduced if sensory axons are affected anywhere from their cell bodies (which are located in the dorsal root ganglion or DRG) distally. In such cases, the axons undergo Wallerian degeneration, and by 10 days after a proximal loss of sensory neurons or axons, the distal response will disappear due to axonal degeneration. On the other hand, if lesions are proximal to the DRG, then the distal axons will remain connected to their cell bodies in the DRG and will conduct normally, even if the patient has complete anesthesia.

Sensory nerve conduction studies can be most useful in diagnosis of lesions that are distal to the DRG, including plexopathies, entrapment neuropathies, and polyneuropathies. Sensory nerve conduction studies should usually be normal in the setting of radiculopathies (when injury is typically proximal to the DRG), central nervous system disorders, neuromuscular junction (NMJ) diseases, and myopathies (the latter two because recordings are obtained from nerve and not muscle).

One should be aware of several potential limitations in utilizing SNAP or CNAP studies. First, these can be technically challenging to obtain when responses are small, such as in...
people with polyneuropathies, people with thick or calloused skin, or in older individuals. It can be challenging as well in unusual nerves that are not commonly studied. Hence, in less frequently studied nerves, one should interpret the absence of a response with caution and consider studying the contralateral limb as well as other sensory nerves. Electronic averaging of responses can be useful to elicit small amplitude responses that would otherwise be unnoticeable in the baseline electronic noise.

As mentioned above, temperature will have significant impact on sensory nerve conduction studies. One should consider that the limb might be cold if latencies are prolonged but amplitudes are large. In contrast, with disease states, latencies would be prolonged but amplitudes would usually be small. If one is unsure about whether temperature is having an influential effect, it is best to warm the limb. This author puts the limb into a plastic bag and then submerges it in a “cooler” or sink filled with warm water for 5 minutes. There are many other techniques, including warm air (e.g., from a hair dryer), infrared lamps, and other methods.

**Compound Muscle Action Potentials and Motor Nerve Conduction Studies**

Compound muscle action potentials (CMAPs) involve electrical stimulation of nerves and recording of the synchronized discharge of muscle fibers in the muscle supplied by the stimulated nerve.

As above, for sensory nerves, the nerve is stimulated supramaximally. As opposed to SNAP or CNAP, one now records not from axons but from muscle fibers at a specified or measured distance away from the stimulation, and records the synchronized discharge of all the muscle fibers that have been activated (Fig. 4-4).

As with SNAPs, several measurements are taken from the CMAP. The amplitude or size of the response is usually measured from baseline to peak. The latency is measured to the onset of the response. Conduction velocity is more complicated to calculate than for SNAPs. Because the latency includes not only conduction time along the nerve but also time for NMJ transmission (about 1 ms), one cannot simply divide the distance between stimulation and recording by the proximal latency. Instead, we stimulate at two points along the nerve and take the difference in distance divided by the difference in latency to derive a conduction velocity (see Fig. 4-4).

The latency can be affected by the same factors as for SNAPs. Physiologically, cold and advancing age slow conduction velocity. With respect to pathological changes, demyelination will cause slowing of conduction velocity or increase in latencies. When demyelination becomes severe, conduction slowing and block will become evident.

CMAP amplitude is dependent not only upon the number of axons that are participating in the response, but also on the integrity of the NMJ and the number and integrity of the muscle fibers supplied by the axons. So, while NMJ defects and myopathies will leave CNAPs and SNAPs unaffected, these lesions can produce a reduction in CMAP amplitude. The CMAP will fall roughly in proportion to the degree of motor axon loss (or, in the case of myopathies, muscle fiber loss). As opposed to the SNAP (where lesions anywhere from the DRG distally cause an amplitude reduction), a reduction in CMAP amplitude can result from lesions from the anterior horn cell (the motor neuron cell body) distally. By 7 days after a proximal loss of motor cell bodies or axons, the distal response will disappear due to axonal degeneration; this occurs earlier than for the SNAP due to earlier failure of the NMJ.

Motor nerve conduction studies can be most useful in diagnosis of lesions that affect the lower motor neuron or surrounding myelin. This includes significant spinal root lesions, plexopathies, entrapment neuropathies, and polyneuropathies.

**Late Responses**

While the typical motor and sensory conduction studies mentioned above rely on conduction distally in the limbs, there are several responses, known as late responses, which depend upon conduction proximally in the peripheral nervous system and thus occur with longer latency.

The F wave was so named because it was originally discovered in the foot (2). Physiologically, F waves depend upon conduction proximally to the motor neuron cell body. Since axons conduct in both directions, when motor axons are activated distally in the limb, they conduct both proximally and distally. For most axons, the impulse travels proximally, past the axon hillock, into the motor neuron and ends. But for a small number (about 3% to 5%), the depolarization traverses through the dendritic tree and comes back through the axon hillock after it is no longer refractory (about 1 ms). As a result, a small late response will be recorded from the muscle after enough time for a round trip from the stimulus site, back to the spinal cord and then back to muscle (Fig. 4-5). The reference (normal) values in humans are about 32 ms in the upper limbs and 55 ms in the lower limbs (3).
one must mark all the F-wave responses obtained. But mean values are less dependent upon sample size, and the measure will be more stable, making it a more reliable parameter.

There are other latency measures that are also sometimes used. Maximal latency reflects the speed of conduction of the slower motor axons. Chronodispersion is measured as the difference between the shortest and longest F-wave latencies. Greater chronodispersion may indicate selective slowing of some motor axons more than others and can be expected in acquired demyelinating neuropathies (2).

Not every motor nerve stimulation results in an F wave. Persistence or penetrance is measured as the percentage of stimulations that produce an F wave. Persistence is altered by many factors (including activity level), and generally, low persistence is not considered a diagnostic finding (7). Absence of F waves, however, is considered abnormal.

Amplitude of the F wave can be measured, usually as a percentage of the corresponding CMAP. This, however, is also variable and is usually not used as a diagnostic finding.

As above, there are many measures one can obtain from F waves. A problem is that each measure has a false-positive rate which, if the reference (normal) values were analyzed optimally, is about 2.5%. As one performs more comparisons with reference (normal) data, each comparison roughly adds 2.5% to the false-positive rate. Thus, if one were to analyze minimal latency, mean latency, chronodispersion, persistence, and amplitude, there is roughly a 12.5% chance that at least one of these five measures would be “abnormal” in the healthy population. Consequently, it is this author’s preference to only measure the mean or minimal latency or indicate if the response is absent.

The technique is similar to motor conduction studies but has several important differences (4). The sweep speed needs to be slow enough to capture the F waves (50 ms analysis time in the upper limb and 100 ms in the lower limb). Since F waves are smaller than CMAPs, sensitivity needs to be sufficient to observe these responses (200 to 500 \( \mu \text{V/} \text{div} \)); some EMG instruments have split screens with the first part of the sweep (e.g., first 20 ms) at 2 to 5 \( \mu \text{V/} \text{div} \) to observe the CMAP and the later part of the sweep at 200 to 500 \( \mu \text{V/} \text{div} \). Conventional teaching is that the cathode should be placed proximally along the nerve with the anode distally to avoid proximal anodal block, but this probably makes little or no difference (5) as long as the cathode is placed at the same location. Conventional teaching is also that the stimulation should be supramaximal, but latencies are similar even with submaximal stimulation (6).

Since F waves are derived from different populations of motor neurons for each response, the latency and appearance vary with each stimulation. Thus, measuring latency is more complicated than for a CMAP or SNAP, where the latency is the same with each stimulation. Because of this variability, one must acquire multiple F waves (typically 10 to 20) and take measurements that consider this variability. Most commonly, one obtains either the minimal latency (the shortest latency of multiple responses) or the mean latency (the average latency). The minimal latency is easiest to measure and is most commonly used. However, the minimal latency will depend upon the number of responses obtained. As more responses are recorded, the chances of finding a shorter latency response increase and the minimal latency will drop. The mean latency, on the other hand, is more time-consuming to measure since

**FIGURE 4-5.** Examples of F waves from an upper limb (left) and a lower limb (right). Note the variability of the response. (Note also sensitivity is different on the left of the sweep than on the right.)
Although F waves will be abnormal in many clinical settings, they provide unique information in relatively few diagnoses. Probably the most useful setting for F waves is when evaluating for acquired demyelinating polyneuropathies, such as acute inflammatory demyelinating polyradiculoneuropathy (AIDP or Guillain-Barre syndrome) (8). In these disorders, the most proximal and distal ends of the peripheral nervous system (i.e., roots and distal axons) are affected first. As a result, one of the earliest findings on nerve conduction studies is absence of F waves or prolongation of F-wave latencies, presumably reflecting slowing or conduction block at the root level. This may be the only finding in some cases.

F waves also provide unique information regarding proximal conduction in early traumatic nerve lesions and will demonstrate changes before distal axon wallerian degeneration occurs and before changes appear on needle EMG. There is some evidence that F waves may change after walking in lumbar spinal stenosis (9), though it is unclear how F waves should contribute to the diagnosis of lumbar spinal stenosis (10).

There are a number of limitations to be aware of when employing F waves for diagnostic evaluation. First, since they involve only a few of the total pool of motor axons, lesions affecting only some of the axons can be missed. For instance, F waves are generally not helpful for detecting radiculopathies (11), although some authors have reported frequent abnormalities (12). Since most muscles have more than one root supplying them, normal F waves can be obtained via the other root supply, thereby bypassing slowing via the involved root. Another limitation is that F waves measure conduction over a long distance. This limits their ability to detect focal pathology that causes only focal slowing across a short segment of nerve, which is then diluted by much longer segments of normal nerve conduction. In addition to the above, F waves are generally applicable to only distal limb muscles. In proximal muscles, the F-wave latency is so short (it does not have far to go to the spinal cord and back) that the F wave is buried in the CMAP response.

H waves (named after Dr. Hoffman) utilize a pathway similar to a muscle stretch reflex (2). In a muscle stretch reflex, the receptors in the muscle spindle are activated by sudden stretch of the muscle. This sends a wave of depolarization proximally up the large-diameter Ia sensory fibers to the spinal cord. At the spinal cord, there is a monosynaptic reflex to the α-motor neuron, and the descending motor axons are activated. H waves use a similar pathway, but the wave of depolarization starts at the Ia afferent fiber instead of the muscle spindle (13).

In normal healthy adults, H waves are easily elicited primarily in the soleus and the flexor carpi radialis muscles. Young children have H waves in many more muscles (before descending inhibitory motor pathways are fully developed). Patients with upper motor neuron lesions will also have easily elicited H waves in many other muscles (14).

When recording H waves, the sweep speed needs to be slow enough to capture the response (50 ms analysis times in the upper limb and 100 ms in the lower limb).
Sensitivity needs to be sufficient to observe these responses although H waves are typically larger than F waves. Similar to F waves, conventional teaching is that the cathode should be placed proximally along the nerve with the anode distally. However, one often uses less current if one stimulates through the knee, with a small 1-cm disc electrode over the tibial nerve (midway between the hamstring tendons at the popliteal fossa) and a 3-cm electrode (the type usually used as a ground electrode) anteriorly over the patella (15). Stimulation is best accomplished with long duration pulses (0.5 to 1.0 ms), since these preferentially activate the large Ia afferent nerve fibers that initiate the reflex arc (16). Stimulation should be at a level that produces the largest H wave, which is usually submaximal for the M wave (Fig. 4-6). Higher levels of stimulation are thought to inhibit H waves at the spinal level (2).

Latency is measured to the onset of the response—that is, the first departure from the baseline. Since the perceived onset will depend upon the display sensitivity (17), it is usually preferable to use a consistent display gain for measurement—ideally, the same used for one’s reference values (e.g., 500 μV per division). Latencies will depend upon height or leg length and most reference values take this into account. Side-to-side latency differences are more sensitive to detecting abnormalities than comparing a measured latency to reference values. A side-to-side latency difference exceeding 1.2 ms is likely abnormal for the tibial H wave (18). Amplitude of the H wave is variable and is also best compared to the other side. When the smaller response is less than 40% of the larger response, this is likely abnormal (18).

Although H waves will be abnormal in many clinical settings, they provide unique information in relatively few diagnoses. Probably the most useful setting for H waves is when evaluating for S1 radiculopathies. H waves are likely more sensitive for detecting S1 radiculopathies than is needle EMG (19). It is postulated that H waves can detect demyelination or sensory axon loss, while needle EMG only detects motor axon loss. H waves, however, provide less information about acuity or chronicity than needle EMG; they usually disappear early (or immediately) after onset of a radiculopathy and often never return. Obtaining and measuring H waves in the upper limbs are more challenging. They are somewhat more difficult to elicit than in the lower limb, and the onset is often obscured by the preceding M wave.

The A wave is another response that is sometimes noted during studies in which one is trying to record late responses. The A wave is a stable, small response that is the same with every stimulation (Fig. 4-7). A waves can appear with either submaximal or supramaximal stimulation. The proposed physiology is activation of motor axons due to either axon branching (usually seen with submaximal stimulation) or ephaptic transmission (seen with supramaximal stimulation) (Fig. 4-8). These are usually considered abnormal, though they are sometimes seen in the tibial nerve of healthy individuals (20). When seen with only submaximal stimulation, they may represent chronic nerve lesions with subsequent reinnervation and axon branching (21). A waves elicited with supramaximal stimulation are likely due to ephaptic transmission. These are sometimes seen in acute demyelinating polyneuropathies (e.g., AIDP) since demyelination creates an ideal setting for ephaptic transmission (see Fig. 4-8).

**Needle EMG**

Needle EMG is a very sensitive and useful technique for detecting pathology in the lower motor neuron, NMJ, and muscle. Key to an understanding of needle EMG is a familiarity with the concept of the motor unit (Fig. 4-9). The motor unit consists of the motor neuron cell body (which, for the limbs, resides in the anterior horns of the spinal cord), the axon, and all the muscle fibers the axon supplies. The number of muscle fibers innervated per axon varies widely. For small muscles that control very fine movements (e.g., extraocular muscles or laryngeal muscles), there may be only 4 to 6 muscle fibers per axon. In large muscles for which force is more important than fine control, such as the quadriceps, there are 500 or more muscle fibers per axon.
In healthy individuals, if one puts a recording needle into a limb at rest, it is usually electrically quiet. When the individual gives a small voluntary contraction and activates the motor unit, a motor unit action potential (MUAP) can be recorded from the muscle (see Fig. 4-9). The MUAP represents the synchronous discharge of all the muscle fibers supplied by the axon. As more force is produced, these MUAPs fire more rapidly, and additional motor units are recruited until there are many MUAPs firing rapidly and none can be individually distinguished.

There are a number of potentials that can be recorded at rest (spontaneous potentials), some in healthy individuals, but most representing disease processes. When moving the needle electrode through the muscle, there is normally a brief electrical discharge from the muscle, known as insertional activity. This is usually high frequency and brief, usually lasting less than 300 ms (depending to some extent on the examiner).

Insertional activity can be prolonged, with discharges persisting longer after needle movement than normal. Prolonged or increased insertional activity can be seen in early denervation, in some myopathies, and in some inherited syndromes (22) but is usually not considered diagnostic when seen in isolation. Reduced insertional activity is measured along a different dimension than increased insertional activity (which is measured in time). Reduced insertional activity refers to reduced amplitude or absence of the burst of muscle fiber discharges when moving the needle through muscle. Reduced insertional activity is seen when there are no resting muscle membrane potentials in the area of study, such as muscle fibrosis (long-term denervation), muscle necrosis (e.g., compartment syndrome) (23), or if it is not in muscle (but, for instance, in adipose tissue).

In healthy individuals, one can easily enter the motor point of the muscle where there are abundant NMJs. When this occurs, one records endplate noise and endplate spikes. Endplate noise consists of small irregular potentials that sound (when played over a loudspeaker) like the sound of a seashell held up to one’s ear. They result from spontaneously released vesicles of acetylcholine (miniature endplate potentials or MEPPs). Endplate spikes are very brief (<5 ms duration) spikes with initially negative (upward) deflections, and they create a sputtering irregular sound similar to fat in a frying pan. They result from MEPPs summatiting to form localized endplate potentials (EPPs). These are normal findings, but one should pull back the needle electrode slightly and move to a different area for two reasons. These are typically painful areas to explore, and the endplate spikes, when viewed from a distance, can take on an initial positivity and look like fibrillation potentials.

Fibrillation potentials and positive sharp waves represent abnormal spontaneous single muscle fiber discharges and are frequently seen in lesions of the lower motor neuron or muscle
Complex repetitive discharges (CRDs) can be seen in a variety of chronic neuropathic and myopathic conditions. It is believed that CRDs result from abnormal muscle membranes that allow ephaptic transmission (i.e., a depolarizing muscle fiber activates an adjacent muscle fiber via local currents, not via any synapse). A pacer muscle cell fires regularly, thereby activating adjacent muscle fibers, which then fire in a sustained rhythmic pattern. CRDs are constant iterative discharges that appear and disappear suddenly without much variation in firing rate or amplitude. They sound like steadily operating machinery. They are not usually considered normal findings but suggest chronic axon or muscle fiber loss. CRDs are most commonly seen in chronic neuropathic and myopathic conditions but are not specific to either one. Some authors report that they can be found normally in select muscles such as the iliopsoas (30). If they are found in a myotomal distribution, they can be considered diagnostic of radiculopathy (11).

Fasciculation potentials represent spontaneous discharges of an entire motor unit or a large part of a motor unit. They often originate in the anterior horn cell but can also originate distally along the axon or a branch of the axon (31). Fasciculation potentials are recognized by their appearance and rhythm of firing. Individually, they look just like a single motor unit. However, they can be distinguished from a voluntary motor unit by their random discharge rate. These discharges are also spontaneous and not under voluntary control. To observe fasciculation potentials, one must often wait with the needle quietly in the muscles without moving the needle for a minute or more. Altering the sweep speed to a slower sweep (e.g., 100 ms per division) is often helpful. Fasciculations otherwise could be easily missed when one is rapidly looking through a muscle on a faster sweep speed.

Fasciculation potentials can be seen in a variety of disorders. Most commonly, one sees benign fasciculations. When asked, approximately 50% of people report fasciculations in their calves, worse with activity and sometimes increased with the use of caffeine (32). Fasciculation potentials can also be seen in thyrotoxicosis and exposure to anticholinesterase medications. They can be seen in radiculopathies and chronic neuropathies as well. Perhaps the most worrisome disease in which fasciculations can be seen is motor neuron disease. Amyotrophic lateral sclerosis (ALS), progressive bulbar palsy (PBP), and other variants of motor neuron disease can all present with fasciculation potentials. The primary way to observe fasciculation potentials is to look at the sweep speed. If one is looking at a faster sweep speed, fasciculations are more difficult to see. If one is looking at a slower sweep speed, fasciculations are easier to see.

Both fibrillation potentials and positive sharp waves are seen in a variety of disorders and neither is specific to denervation (hence one should not use the term denervation potentials). They are commonly seen in disorders of motor neurons, motor axons, the NMJ (presynaptic more commonly than postsynaptic), and muscle (inflammatory myopathies and dystrophies more commonly). Upper motor neuron lesions (e.g., spinal cord injury, stroke, and traumatic brain injury) can also produce fibrillation potentials and positive sharp waves in weak muscles, though less commonly than lower motor neuron lesions do.

When fibrillation potentials and positive sharp waves are recorded, one usually grades them on a 1+ to 4+ scale (see table). The examiner should remember that this is an ordinal scale but not a ratio or interval scale. Thus, we know that a grading of 2+ represents more axon loss than 1+, but not necessarily twice as much. Moreover, a finding of 4+—fibrillations does not mean that all axons have been lost—this can be seen in very partial axon loss (25). Some authors have reported that positive sharp waves and fibrillations can be seen in normal paraspinal and foot muscles (26–28), but other authors have not found this to be common (29).

Complex repetitive discharges (CRDs) can be seen in a variety of chronic neuropathic and myopathic conditions. It is believed that CRDs result from abnormal muscle membranes that allow ephaptic transmission (i.e., a depolarizing muscle fiber activates an adjacent muscle fiber via local currents, not via any synapse). A pacer muscle cell fires regularly, thereby activating adjacent muscle fibers, which then fire in a sustained rhythmic pattern. CRDs are constant iterative discharges that appear and disappear suddenly without much variation in firing rate or amplitude. They sound like steadily operating machinery. They are not usually considered normal findings but suggest chronic axon or muscle fiber loss. CRDs are most commonly seen in chronic neuropathic and myopathic conditions but are not specific to either one. Some authors report that they can be found normally in select muscles such as the iliopsoas (30). If they are found in a myotomal distribution, they can be considered diagnostic of radiculopathy (11).

Fasciculation potentials represent spontaneous discharges of an entire motor unit or a large part of a motor unit. They often originate in the anterior horn cell but can also originate distally along the axon or a branch of the axon (31). Fasciculation potentials are recognized by their appearance and rhythm of firing. Individually, they look just like a single motor unit. However, they can be distinguished from a voluntary motor unit by their random discharge rate. These discharges are also spontaneous and not under voluntary control. To observe fasciculation potentials, one must often wait with the needle quietly in the muscles without moving the needle for a minute or more. Altering the sweep speed to a slower sweep (e.g., 100 ms per division) is often helpful. Fasciculations otherwise could be easily missed when one is rapidly looking through a muscle on a faster sweep speed.

Fasciculation potentials can be seen in a variety of disorders. Most commonly, one sees benign fasciculations. When asked, approximately 50% of people report fasciculations in their calves, worse with activity and sometimes increased with the use of caffeine (32). Fasciculation potentials are also seen in thyrotoxicosis and exposure to anticholinesterase medications. They can be seen in radiculopathies and chronic neuropathies as well. Perhaps the most worrisome disease in which fasciculations can be seen is motor neuron disease. Amyotrophic lateral sclerosis (ALS), progressive bulbar palsy (PBP), and other variants of motor neuron disease can all present with fasciculation potentials. The primary way to observe fasciculation potentials is to look at the sweep speed. If one is looking at a faster sweep speed, fasciculations are more difficult to see. If one is looking at a slower sweep speed, fasciculations are easier to see.
in which benign fasciculations can be distinguished from those associated with disease is by the “company they keep.” In motor neuron and other progressive disease processes, one will often see fibrillations potentials, positive sharp waves, and large amplitude long duration motor unit potentials, along with the fasciculations potentials. Benign fasciculations do not present with these other findings. When comparing fasciculations associated with motor neuron disease to those in benign fasciculations, there are some other minor differences, but these are not sufficient to be diagnostic by themselves. Fasciculations associated with motor neuron disease tend to be less frequent in firing rate, and they tend to be larger in amplitude, polyphasic, and longer duration (reflecting the concurrent ongoing reinnervation).

Myokymia is often thought of as a grouped fasciculation. Myokymic discharges represent groups of motor units firing in a burst pattern (usually with a regular burst rate) (33). These discharges, when played over the EMG instrument’s loudspeaker, are often described as sounding like marching soldiers. However, the bursts can be quite different in both their duration and discharge rate making recognition sometimes difficult. Myokymia is distinguished from CRDs because of the bursting nature rather than a single on and off with constant firing as seen in CRDs. Myokymic discharges are also distinguished from myotonic discharges because the former do not change in firing frequency or amplitude. Myokymia can be seen in a variety of conditions and is usually thought of separately for facial myokymia and limb myokymia. In facial myokymia, one often thinks of multiple sclerosis, pontine gliomas, and other brainstem disorders. In the limb, myokymia can be seen in some chronic compressions such as radiculopathies and entrapment neuropathies.

It is probably most important to remember that limb myokymia can occur in the presence of radiation plexopathy (34). This is useful when patients have had radiation for breast cancer, Hodgkin’s lymphoma, or other malignant tumors. Occasionally, these patients will present with a later-onset plexopathy, and the diagnostic question is whether the plexopathy represents tumor invasion or delayed-onset radiation plexopathy. Tumor invasion in the brachial plexus tends to present with pain, lower trunk lesions, and a Horner’s syndrome. Radiation plexitis presents with upper trunk lesions, paresthesias, and myokymia. Thus, myokymia argues for a diagnosis of radiation plexopathy in these cases.

Myotonia is a discharge that originates in single muscle fibers. It is believed to be due to abnormal chloride conductance across the muscle cell membrane. These discharges are noted by their unique waxing and waning quality, both in frequency and in amplitude. Because the firing frequency changes, the pitch produced over the loudspeakers changes as well, and one hears sounds often described as a “dive bomber” or “revving motorcycle.” The amplitude of the response also changes over time making the sound louder and softer during the discharge pattern. This can be distinguished from CRD because of its changing nature (CRDs are noticeably constant in firing frequency and amplitude).

Myotonia can be seen in a variety of muscle diseases including myotonic dystrophy, myotonia congenita, paramyotonia, and other disorders (35). Myotonic dystrophy is an unusual myopathy in that distal muscles are affected more than proximal muscles, and the same is true for where one might record these discharges. Distal hand muscles (e.g., first dorsal interosseous [FDI]) are often most affected. Myotonia is often worsened by cold and can be brought out by icing the distal part of the limb before recording. Myotonic discharges are also reduced immediately after exercise. Since myotonia originates from muscle, it is not expected in patients with neuropathic disease.

A possible variant of myotonia is the syndrome of diffusely abnormal increased insertional activity, also known as “EMG disease.” This disorder is an autosomal dominant inherited syndrome and presents with increased insertional activity and persistent positive sharp waves in essentially all muscles tested in the body (22). One can tell, however, that the patient does not have a true neuropathic disease because there are no fibrillation potentials, motor unit potentials are normal in morphology and size, and recruitment is normal. Because of the diffuse positive sharp waves, these patients can often be erroneously diagnosed with motor neuron disease or other serious disorders. It is difficult to know whether this syndrome produces symptoms or not, since those coming to the EMG laboratory are usually selected for some symptoms to start with. However, relatives of the patient with this disorder are often asymptomatic.

**Motor Unit Analysis**

Analysis of MUAPs yields important information about the integrity of the motor unit and whether there are changes in axons, muscle fibers, or the NMJ. The MUAP, as mentioned earlier, represents the synchronous discharge of all the muscle fibers supplied by a single motor neuron. Duration of the MUAP is largely influenced by the size of the motor unit territory, that is, the number of muscle fibers supplied by the single axon that are within the recording area of the needle electrode. Duration of the MUAP is not markedly affected by proximity of the depolarizing muscle fibers to the needle electrode.

Amplitude of the MUAP is also related to motor unit territory but is more influenced by the proximity of the discharging muscle fibers to the recording electrode. When the needle is close to the depolarizing muscle fibers, the amplitude is much larger than when it is some distance away. Thus, motor unit duration is a more reliable indicator of motor unit territory than is amplitude. There are reference values established for concentric electrodes for most of the major muscles in the body (36). These vary according to muscle and age. More proximal muscles, and especially bulbar muscles, have shorter duration MUAPs. MUAP duration tends to increase with age. There are not well-established reference values for monopolar electrodes. Thus, when performing quantitative MUAP analysis, one should use a concentric electrode.

Polyphasicity is noted when there are more than five phases of an MUAP. A phase represents a change in direction that crosses the baseline. Thus, one can calculate the number of...
phases as the number of baseline crossings plus one. Changes in the direction that do not extend across the baseline are not counted as a phase. The presence of polyphasicity alone is usually not diagnostic, since most normal muscles have a small percentage (approximately 20%) of polyphasic MUAPs. However, polyphasic, long duration, large amplitude MUAPs are usually seen in neuropathies, and polyphasic, short duration, small amplitude MUAPs are typically seen in myopathic conditions.

In neuropathic conditions, two types of reinnervation may be noted that change the MUAP morphology. When incomplete nerve injuries are present, some axons are spared, and others undergo wallerian degeneration. In this case, the remaining axons send distal sprouts to reinnervate the denervated muscle fibers. These sprouts are initially poorly myelinated and immature, which results in less synchronous discharge of muscle fibers than seen normally. When this distal axonal sprouting happens, one will observe polyphasic long-duration large-amplitude MUAPs (Fig. 4-11). As the new sprouts mature, the polyphasicity is reduced because the new sprouts become better myelinated and the muscle fibers fire synchronously again. Thus, late after reinnervation by axon sprouting, the examiner will see large amplitude long duration but not polyphasic MUAPs.

In cases of complete nerve transactions or complete axon loss, the picture is different. In these cases, there is no distal axonal sprouting because there are no viable distal axons. Rather, one is dependent upon axons regrowing from the site of injury down to the muscle. When these new axons first reach the denervated muscle, they innervate just a few muscle fibers. These new MUAPs (which used to be called nascent potentials) are typically short in duration, polyphasic, and small in amplitude because the axons have supplied only a few muscle fibers. As these axons continue to sprout, they become larger and longer duration and retain polyphasicity until the new sprouts mature several months after reinnervation.

Whether reinnervation occurs by distal axon sprouting or axon regrowth from the site of injury, recruitment will be reduced or discrete. Because there are fewer MUAPs, the existing motor units will fire more rapidly than normal, and there will be fewer MUAPs firing even with maximum contraction.

The model for myopathic changes in MUAPs can be thought of as random loss of muscle fibers within the motor unit (Fig. 4-12). In this case, since there are reduced numbers of muscle fibers, the MUAP duration and amplitude will be reduced. These MUAPs are polyphasic possibly due to less muscle fibers in the motor unit or possibly due to temporal dispersion along muscle fibers that conduct at different speeds in the setting of disease. In either case, in myopathies, one sees small-amplitude short-duration polyphasic MUAPs. These MUAPs are recruited early with many MUAPs appearing on the EMG screen with even small amounts of force generation. It is difficult to isolate a single MUAP in patients with the most severe myopathies.

NMJ disease can look, in many ways, just like a myopathy on needle EMG. Since the NMJs in a muscle do not all fire in these cases, there are less functional muscle fibers per motor unit, and the duration and amplitude of the MUAPs will be reduced. A feature commonly found in NMJ disease that is not seen in myopathy, however, is that of motor unit instability. Since the NMJs will fire variably and unreliably, the entire MUAP will vary in its size and morphology as muscle fibers fire or do not fire with each successive MUAP discharge. This motor unit instability can be seen in NMJ disease as well as in recent reinnervation, since in the latter case the NMJs are also not mature and reliable. To observe motor unit instability, it is most helpful to look at MUAPs with a trigger and delay line.
so that the MUAP fires repetitively at the same location on the screen.

While one typically evaluates MUAPs in a semiqualitative fashion, there are situations in which it is important to evaluate MUAPs more quantitatively. In the setting of possible myopathy, it is often very important to quantitatively measure MUAP duration to see whether it is shorter than reference values. Myopathies are sometimes more difficult to diagnose without quantitative analysis than neuropathic conditions, since one more frequently lacks the abnormal spontaneous activity or abnormal nerve conduction that are present in a neuropathic lesion.

To perform quantitative needle EMG, one must change the technical aspects of data collection in several ways. First, most reference values have been collected using concentric needle electrodes rather than monopolar needle electrodes. The concentric electrode has the advantage of a smaller recording volume and better reliability in duration measurements. Thus, one must use a concentric electrode. Normally, filters for needle EMG are set roughly from 20 Hz to 10 kHz. The low-frequency filter setting of 20 Hz is sufficient for capturing most electrophysiologic muscle potentials and reduces baseline wander, but it also distorts the low-frequency initial and terminal phases of the MUAP where departure from and return to baseline are gradual. For these reasons quantitative needle EMG requires that the low-frequency filter be reduced to 2 Hz, and multiple firings of individual MUAPs are averaged to present a clean single MUAP.

Multiple studies have been reported dating from the 1950s looking at how many MUAPs need be collected to obtain a reproducible mean MUAP duration for a particular muscle. These studies have shown that minimally 20 different MUAPs should be collected to obtain a reproducible value for mean MUAP duration, unless there are extremely abnormal MUAPs (37). Twenty different areas of the muscle should be studied, and this can usually be accomplished with only three to five different needle insertions through the skin. While quantitative needle EMG is commonly used in the setting of myopathy, generally only one or two muscles are studied this way. It is less common to use quantitative needle EMG in multiple muscles or in many neuropathic conditions where there are other hints of neuropathies.

After MUAPs are assessed, the electrodiagnostic medical consultant should then evaluate MUAP recruitment. Normally, when one provides a very small force, one will recruit a single motor unit firing slowly, typically about 5 Hz. As more force is exerted, this MUAP will speed up to 10 to 12 Hz and produce greater levels of force. However, as the firing rate reaches about 12 to 15 Hz (depending on the muscle), a second MUAP is then recruited to allow additional force production. In this way, muscle forces are generated both by increasing rates of MUAP firing and by increasing the number of MUAPs firing. Normally, at full contraction, there are many MUAPs firing rapidly and the baseline cannot be distinguished. There are generally three types of abnormalities of recruitment. Sometimes, patients have insufficient central drive from the upper motor neurons to generate high levels of MUAP firing. This could be due to pain, poor voluntary effort, or upper motor neuron lesions. In these cases, one sees less than the full number of MUAPs, but the ones that are present fire slowly due to reduced upper motor neuron drive. This is often termed central recruitment.

In other cases, there is sufficient upper motor neuron drive, but there are reduced numbers of motor units available to participate in force production. In these cases, the initial MUAPs fire more rapidly before a second MUAP is recruited because there are fewer motor units available to recruit. In extreme cases, there may only be one or two MUAPs firing very rapidly (up to 30 to 40 Hz) without any additional MUAPs present. This is typically termed reduced recruitment (or in cases where there are only very few MUAPs, discrete recruitment). It is distinguished from central recruitment in that the MUAP firing rate is quite fast.

A third abnormality one can see in MUAP recruitment is early recruitment. This is typically seen in myopathies or NMJ defects. With early recruitment, since each MUAP in a myopathy produces less force than normal, more motor units are recruited sooner or earlier than normal for a given level of force generation. The examiner asks the patient to produce a small amount of force and sees many MUAPs firing, more than expected for that level of force. In many cases of early recruitment, it is difficult to have a patient fire just one MUAP; it is usually either many or none. Thus, early recruitment represents many MUAPs firing for a low level of force production. Assessing early recruitment requires both measurement of force (whether qualitative or quantitative) and observation of MUAP firing patterns.

Repetitive Stimulation Studies
Repetitive stimulation studies are primarily useful for detection of NMJ abnormalities. To understand the underlying reasoning of how one performs NMJ studies, one needs to first review the physiology of NMJ, which is well described elsewhere (38).

Normally, when a motor axon is depolarized, the wave of polarization travels toward the axon terminal as a result of the propagation of opening and closing of sodium channels. When it reaches the terminal end of the motor axon near the end plate, voltage-gated calcium channels are activated. The opening of these channels allows an influx of calcium ions (Ca\(^{2+}\)) into the presynaptic terminal. The opening of the voltage-gated and dependent calcium channels and the resulting calcium ion influx result in an increase of intracellular Ca\(^{2+}\) at the presynaptic terminal for about 100 to 200 ms before the Ca\(^{2+}\) is pumped out and the intra-axonal Ca\(^{2+}\) is returned to the resting concentration. This influx of calcium, however, allows the acetylcholine vesicles to fuse with the presynaptic membrane and release their quanta of acetylcholine into the extracellular space at the synapse.

At rest, the presynaptic terminal has occasional spontaneous release of quanta of acetylcholine that occur irregularly (38). When these quanta reach the postsynaptic terminal, they create MEPPs. However, when the presynaptic axon is excited or depolarized, the release of quanta is much greater. A high
concentration of acetylcholine is generated, which in turn activates the muscle fiber. In healthy individuals, the presynaptic terminal releases three to five times as much acetylcholine as is required to activate the postsynaptic membrane of the muscle fiber. This extra amount of acetylcholine release is known as the safety factor. While there is normally a large safety factor for the first discharge of a motorneuron, the amount of acetylcholine released drops with successive discharges when the nerve is activated at 2 to 3 Hz. However, because the safety factor is so large, there is still more than enough acetylcholine to activate the muscle fiber (Fig. 4-13).

Once the acetylcholine is released into the synaptic cleft, it encounters acetylcholinesterase, the enzyme so named because it hydrolyzes acetylcholine molecules. This enzyme does not "know" whether the acetylcholine it is digesting has recently come from the presynaptic terminal or has already been used to activate the postsynaptic receptors. It digests the acetylcholine it encounters either way. Due to this enzymatic activity, it is estimated that only about half of the acetylcholine released from the presynaptic terminal makes it to the postsynaptic receptors. Once the acetylcholine reaches the postsynaptic receptors, it creates localized areas of membrane depolarization known as EPPs. These EPPs then summate to create a muscle membrane depolarization, and the muscle fiber discharges.

Exercise also has significant influences on NMJ physiology. Immediately after exercise, one sees postexercise potentiation or facilitation. In part, this is due to the fact that motor units fire at rates approaching 30 Hz with full voluntary contraction. At these rates, there is only roughly 33 ms between each axon depolarization, which is less than the time needed for the Ca^{2+} to be pumped out of the presynaptic terminal. Thus, with exercise, Ca^{2+} builds up in the presynaptic terminal, and this further facilitates release of acetylcholine. When the safety factor is intact, this is not noticed because there is already more than enough acetylcholine to fully depolarize the muscle fiber. However, when the safety factor is reduced in disease states, some NMJs do not fire, and postexercise potentiation will create a larger amplitude CMAP than preexercise. This is termed postexercise facilitation.

There is also a phenomenon of postexercise exhaustion, which occurs 2 to 4 minutes after sustained exercise. In this period of postexercise exhaustion, there is less acetylcholine released than in the preexercise condition. Again, in healthy individuals, this would not be noticed because of the large safety factor. However, in individuals with marginal or reduced safety factors, some NMJs will not fire, and the CMAP will be reduced in amplitude.

There are several elements of the electrodiagnostic examination to keep in mind when evaluating an individual with possible NMJ disease (39). First, one should perform a standard motor nerve conduction study of the muscle and nerve to be studied with repetitive stimulation. This is to both assess the resting amplitude of the response as well as be sure that the nerve under study does not have a prior or subclinical injury that could alter the results of NMJ testing.

The second step is to record the distal CMAP at rest and then again immediately after 10 seconds of maximum voluntary isometric exercise (Table 4-1). This procedure is used to look for any postexercise facilitation. When presynaptic diseases of the NMJ are present, the initial resting CMAP will be low in amplitude but will increase markedly after 10 seconds of exercise, since Ca^{2+} buildup within the presynaptic terminal occurs and acetylcholine release is greatly facilitated.

The third step is to perform repetitive stimulation at slow rates of 2 or 3 Hz. Generally, trains of five stimulations are sufficient, although some laboratories use ten to look for consistency of the response. When stimulating at 3 Hz, there are 333 ms between each stimulation. This means that between each stimulation there is essentially a full return of baseline or resting state of Ca^{2+} concentration. Thus, there is no significant facilitation of acetylcholine release. As discussed above, there is depletion of the immediately available stores of acetylcholine vesicles and resultant diminution of acetylcholine release with each successive stimulation. In healthy individuals, while the acetylcholine release is reduced with each successive stimulation, the CMAP remains stable and unchanged because of the large safety factor. Most laboratories consider amplitude decrements of less than 10% between the first stimulation and the...
fourth or fifth to be normal. However, in the setting of NMJ diseases, when the safety factor is diminished, a reduction in CMAP with progressive stimulations will exceed the normal range. To be considered abnormal, the reduction should be more than 10% decrement compared to the baseline value.

One must be mindful of several technical points when performing repetitive stimulation to avoid erroneous observation. First, temperature is critical. When the temperature is below 34°C, enzymatic activity of acetylcholine esterase is reduced. As a result, cold causes an increase in acetylcholine levels within the NMJ, which compensates for any pathological reduction of the safety factor. This reduces the likelihood of detecting decrements that might be present otherwise. Second, a stable recording baseline is important. The movement of the limb has to be minimized so it does not create signal artifacts in the recordings. Many laboratories will attempt to immobilize the part of the limb being studied, although this is more difficult with proximal or bulbar muscles. The choice of the nerve muscle pair is also crucial. NMJ diseases, especially myasthenia gravis (MG), tend to affect proximal muscles more than distal muscles. However, the proximal studies are technically more difficult than the more distal limb nerve muscle pairs. Thus, many laboratories will start with the ulnar nerve recording from the abductor digiti minimi (ADM). If this is normal, then most laboratories will proceed to a proximal limb muscle such as the trapezius (stimulate spinal accessory nerve) or the deltoid (stimulate at Erb’s point). Finally, if these two studies are normal, then the study of the facial nerve recording from nasalis can yield important information. While the latter technique is more sensitive, it is also more vulnerable to movement artifact.

After the initial series of five stimulations at 3 Hz, one should then repeat studies both immediately after exercise (to look for postexercise potentiation) and several minutes after exercise (to look for postexercise exhaustion). After the pre-exercise testing, one asks the patient to maximally contract the muscle for 30 or 60 seconds. Immediately postexercise, one then gives another series of five stimulations as before the exercise. This series is repeated again at 1-minute intervals until 4 minutes postexercise. By performing this procedure one assesses for postexercise facilitation (or potentiation) immediately after exercise and for postexercise exhaustion, which typically occurs at 2 to 4 minutes postexercise.

In some cases, high-frequency repetitive stimulation is used. This is particularly helpful when the patient is not able to voluntarily exercise, and when one is looking for a presynaptic disease such as Lambert-Eaton myasthenic syndrome (LEMS) or botulism. Stimulation rates are set in the 30- to 50-Hz range which means there is only 20 to 33 ms between each successive stimulation. At these rates, the subsequent stimuli occur too frequently for the presynaptic nerve terminal to pump out the extra influx of Ca²⁺. Thus, calcium concentration builds up in the presynaptic terminal, and consequently, acetylcholine release is markedly increased. High-frequency repetitive stimulation is primarily used to look for postexercise facilitation and is mostly useful in presynaptic defects. Healthy patients have a moderate increase in CMAP size with these rates of stimulation, known as pseudofacilitation. The phenomenon is not likely due to mechanical artifact. While some have postulated that hypersynchronization of the muscle fiber action potentials is the cause, a more convincing explanation, for which there is now experimental evidence, is that the muscle fibers undergo hyperpolarization, due to the intramuscular release of norepinephrine and consequent stimulation of the electrogenic Na⁺/K⁺ pump (40).

One generally groups NMJ defects into presynaptic and postsynaptic types. MG is the most common postsynaptic NMJ disease. The pathophysiology has been postulated to result from acetylcholine receptor antibodies likely generated from an autoimmune process. These antibodies act as antagonists against the acetylcholine receptors at the postsynaptic cleft of the NMJ. In addition, there are marked distortions of the NMJ with a widened cleft and fewer infoldings postsynaptically well
visualized on morphological studies by electronmicroscopy (41,42). This is important for electromyographers to keep in mind since the widened cleft means that there is a greater chance that acetylcholinesterase will digest the acetylcholine before it reaches the postsynaptic cleft. Moreover, for the acetylcholine molecules that do reach the postsynaptic membrane, there are fewer available receptors to activate, hence there is a much reduced safety factor.

In MG, one usually sees relatively normal CMAPs at rest and little change immediately after exercise. However, with slow repetitive stimulation (2 to 3 Hz), one will often observe a reduction in the CMAP maximal at the fourth or fifth stimulation in the series. As mentioned above, the decrement needs to be greater than 10% to be considered normal. Repetitive stimulation in these patients should also display a repair of the decrement immediately after exercise due to postexercise potentiation. At 2 to 4 minutes after exercise, when postexercise exhaustion occurs, a more marked decrement is seen than in the preexercise phase. At times, patients with more subacute disease will demonstrate no decrement preexercise but will demonstrate a significant decrement 2 to 4 minutes postexercise. Thus, when evaluating for potential MG, one should always complete the series of postexercise runs of repetitive stimulation as described above. With high-frequency repetitive stimulation, patients with MG show only small increments as do normal individuals due to pseudofacilitation (described above).

LEMS is due to an abnormality in the voltage-gated calcium channels located in the postsynaptic terminal (43). It is most commonly associated with small-cell lung cancer but can also be seen in other cancers, some autoimmune disorders, and it is rarely idiopathic.

Patients with LEMS demonstrate very small CMAPs at rest (44). However, immediately after 10 seconds of exercise, the CMAP increases in size dramatically due to the enhanced release of acetylcholine precipitated by brief periods of exercise and consequent increase in calcium ion concentration within the presynaptic terminal. The results are quite striking, and few other diseases cause CMAP amplitude to increase so impressively. To be diagnostic, the increase in CMAP amplitude needs to be at least 100%, that is, a doubling of the initial CMAP. Remember that the CMAP does not increase in size above normal values, but simply goes from a very small initial amplitude to one that is closer to normal.

With repetitive stimulation at slow rates, patients with LEMS will demonstrate a decrement between the first and fourth potential that is not unlike that seen in MG. This is related to the abnormal safety factor that is seen in LEMS as well as MG. However, what will be noted is that in the series of five stimulations immediately postexercise, all of the CMAPs will be much larger than they were in the preexercise condition. At 2 to 4 minutes postexercise, the CMAPs fall to their baseline levels, and the results become similar to what was seen at rest.

It is in the setting of LEMS that high-frequency repetitive stimulation is most useful. With stimulation rates of 30 Hz or more, there is a buildup of Ca$_{2+}$ in the presynaptic terminal, which tends to overcome the initial defect. As a result, the CMAP becomes more than double its resting amplitude, often approaching a near-normal amplitude.

Botulism is another example of presynaptic NMJ disease. In this case, it is not the voltage-gated Ca$_{2+}$ channels that are impaired but rather the ability of the presynaptic vesicle membrane to fuse with the axon terminal plasma membrane and release acetylcholine (45). The effect of botulism on electrodiagnostic findings is conceptually similar to that seen in LEMS. However, in the real world, the presentation of botulism tends to be more variable and not all cases will have marked increments with fast repetitive stimulation or marked postexercise facilitation.

There are a number of limitations in NMJ testing. First, the sensitivity of repetitive stimulation studies is probably only about 60% to 70% in cases of MG, even with technically well-performed studies utilizing exercise (46,47); specificity, however, is very good. The reasons for only moderate sensitivity are likely multifold and are beyond the scope of discussion here. Single fiber EMG, which is not being reviewed in this chapter, has a greater sensitivity than repetitive stimulation. Similarly, repetitive stimulation is not as sensitive for presynaptic lesions as is single fiber EMG. Another limitation is that cold can have a significant influence on these studies and cause false-negative results. Patients who are on anticholinesterase medication such as pyridostigmine will not demonstrate the same degree of abnormalities as those who are without the medication. Finally, good examiners always keep in mind the technical aspects, with the most concerning one being errors introduced by limb movement. With repetitive stimulation, there is often considerable limb movement unless one is vigilant. Limb movement alters the position of the recording and stimulating electrodes, and their contact to the skin. In this way, movement can bring the stimulation to less than supramaximal, which could mimic a decrement.

**DISEASE PROCESSES**

**Entrapment Neuropathies**

There are a variety of entrapment neuropathies that commonly present to the electrodiagnostic laboratory for evaluation. The most common ones, which will be discussed here, include median neuropathy at the wrist (carpal tunnel syndrome), ulnar neuropathy at the elbow, radial neuropathy at the spiral groove, fibular neuropathy at the fibular head, and tarsal tunnel syndrome.

There are a few general concepts regarding entrapment neuropathies that can be useful to keep in mind as one evaluates these various neuropathies. First, entrapment of a nerve typically first produces demyelination and consequent conduction slowing in more mild cases. As the demyelination progresses, the electrophysiologic changes progress from slowing to conduction block. With more severe entrapments, axon loss can occur. There is some variation in this general principle across nerves since some nerves, such as tibial nerve at the...
Median Neuropathy at the Wrist

Median neuropathy at the wrist, which is usually responsible for carpal tunnel syndrome, is the most common entrapment neuropathy to be referred to electrodiagnostic laboratories in the United States (50). Symptoms commonly include hand numbness and weakness (51). The patient often does not localize the numbness to simply the median distribution but rather indicates that the whole hand becomes numb (52). A complaint of dropping things is frequent. Symptoms are usually worse at night, and patients may occasionally report they “flick” their wrist to relieve symptoms.

On examination, one can find weakness of the thenar muscles and possibly some mildly reduced sensation. There are a number of physical signs such as Tinel’s sign, Phalen’s sign, and the flick sign that are suggestive of carpal tunnel syndrome. However, the sensitivity and specificity of these tests are not high (1) and they should not be used to make or rule out a diagnosis.

There are number of risk factors for carpal tunnel syndrome that have been well documented in the literature (53). In polyneuropathies, nerves are more susceptible to superimposed entrapment such as in diabetes mellitus. Diseases in which there is more synovial tissue at the wrist such as rheumatoid arthritis also increase the risk of carpal tunnel syndrome by about threefold. Individuals who have manual work that involves high repetition and high are at significantly increased risk of carpal tunnel syndrome. Obesity is also a risk factor (54).

Since detection of slowing of median nerve conduction across the wrist is the most useful way to localize the entrapment neuropathy, this should be the focus of one’s electrodiagnostic assessment. There have been many approaches described for diagnosing carpal tunnel syndrome with nerve conduction studies. For a more in-depth coverage, readers are encouraged to review other articles (55,56). One’s general approach should be to measure sensory and motor conduction across the wrist and to compare the latencies with nearby nerves in the hand, such as the radial or ulnar nerve, that do not traverse the carpal tunnel. This helps to exclude the effects of temperature, age, and other factors such as polyneuropathy, which may influence nerve conduction. As is the case in most entrapment neuropathies, sensory fibers are usually affected first. Rarely, motor axons are preferentially affected, possibly because of focal compression of the recurrent branch of the median nerve or selective effects on fascicles within the median nerve at the wrist (57).

There are many approaches for evaluating median sensory conductions across the wrist, and it is critical to think through these alternative approaches before even seeing the patient. In particular, as mentioned above, one should not adopt the methodology of performing one test and, upon finding a normal result, performing another test until one finds an abnormality. Although this might seem tempting intuitively, it is risky since each additional test performed carries a 2.5% false-positive rate, which is roughly additive as each new test is performed.

When selecting sensory nerve conduction studies, one should select studies that are (in descending order of importance)

- Specific (few false positives)
- Sensitive (few false negatives)
- Reliable (obtain the same results today and tomorrow)
- Not influenced by covariates such as temperature and age

There are three sensory conduction studies that have been shown to be reasonably good in terms of the criteria mentioned above (58). These are demonstrated in Figure 4-14. Comparison of median and ulnar conduction to the ring finger allows the detection of slowing of median conduction in comparison to the ulnar nerve, which does not traverse the carpal tunnel. A median-ulnar difference exceeding 0.4 ms is likely abnormal. Similarly, comparison of the median and radial nerve to the thumb has similar advantages. Here, a difference exceeding 0.5 ms is probably abnormal. The third test is the median and ulnar comparison across the palm over an 8-cm distance. This study should demonstrate no more than a 0.3 ms difference in healthy individuals.

This author has published a method to summarize these three tests into one result known as the combined sensory index
or CSI (since the television show Crime Scene Investigation has become popular, this is now being called the Robinson Index). To calculate the CSI, one performs all three of the studies mentioned above and adds the latency differences (median minus ulnar or median minus radial) together (when these are negative, i.e., the median is faster, a negative number is used). The CSI, because it summarizes three different tests, has been shown to be highly specific and more sensitive than the individual tests (59). It is also more reliable than single tests when one studies the same patient on two different occasions (60). A CSI exceeding 0.9 ms is considered abnormal (59).

Motor nerve conduction studies are also, as mentioned above, an essential component of the electrodiagnostic evaluation of carpal tunnel syndrome. These should be performed even if sensory conduction studies are normal. Most commonly, studies are performed with stimulation of the median

![Figure 4-14](image-url)
nerves at the wrist and recording over the abductor pollicis brevis (APB). Generally, latencies exceeding 4.5 ms are considered abnormal. It is not particularly useful to compare one median nerve with the other side because of the frequency of bilateral CTS. However, some electromyographers compare the median motor latency with the ulnar motor latency; a difference exceeding 1.5 ms is considered abnormal. While some authors do advocate stimulating both at the wrist and at the palm (61), it is difficult to stimulate only the median nerve in the palm, and one can be easily mislead into a false diagnosis if the ulnar nerve is inadvertently stimulated in the palm (62).

Needle EMG is sometimes useful in evaluating patients with carpal tunnel syndrome (63). There is not a consensus about when thenar muscle EMG should be performed. It is this author's practice to perform needle EMG of the thenar muscles in three settings:

- Patients in whom the motor response is abnormal (this group has a higher yield)
- Patients with a history of trauma (in which axon loss is more likely)
- Patients with a clinical presentation that suggests another possible diagnosis (such as radiculopathy orplexopathy)

### Ulnar Neuropathy at the Elbow

Ulnar neuropathy at the elbow is another common entrapment neuropathy presenting to the electrodiagnostic medical consultant. The etiology of ulnar neuropathy at the elbow varies but can be due to acute injury, entrapment in the cubital tunnel (under the aponeurosis between the two heads of the flexor carpi ulnaris), or prolonged stretching of the nerve in the ulnar groove when the elbow is held in the flexed position (64). Tardy ulnar palsy is a result of prior elbow injury causing an elbow deformity and slowly progressive injury to the ulnar nerve.

Symptoms of ulnar neuropathy typically include numbness over the small finger and the ulnar half of the ring finger. Generally, an ulnar neuropathy at the elbow also affects sensation over the dorsum of the hand on the ulnar side, an area supplied by the dorsal ulnar cutaneous nerve, which branches from the ulnar nerve proximal to the wrist. By contrast, ulnar nerve lesions at the wrist spare the dorsal ulnar cutaneous territory because they are distal to this branch point. Ulnar neuropathy at the elbow should spare sensation over the medial forearm. This area is supplied by the medial antebrachial cutaneous nerve originating from the medial cord of the brachial plexus and should be spared in ulnar neuropathy at the elbow.

Patients often also present with weakness of ulnar hand muscles complaining that they have difficulty holding small objects and difficulty with grip strength. They may sometimes notice atrophy of the FDI muscle. At times, they will report that when they put their hand into their pocket, the small finger does not make it in. This is known as Wartenburg’s sign and reflects weakness of the interosseous muscles, specifically the adductor of the small finger.

On physical examination, one will often note weakness of interosseous muscles, atrophy of the FDI, and reduced sensation in the ulnar nerve territory in the hand. One may also find a Froment’s sign indicating weakness of the adductor pollicis and the FDI (65). A Tinel’s sign can often be noted at the elbow, but this is nonspecific and can be seen in a number of normal healthy individuals.

Because sensory conduction is difficult to reliably record across the elbow, most electromyographers will rely upon motor conduction studies of the ulnar nerve (56). There are a number of technical elements to keep in mind when performing these studies. First, it is advisable to record from both the ADM and the FDI at the same time, utilizing two channels of the EMG instrument. Although each muscle has similar sensitivity for detecting ulnar neuropathy at the elbow, there is not a complete overlap and sometimes one muscle will demonstrate conduction block when the other one does not (66). Stimulation usually occurs at the wrist, below the elbow, and above the elbow. When stimulating across the elbow, one should have the elbow in a flexed position with a roughly 70- to 90-degree angle. This is important because it stretches the nerve through the ulnar groove. If the elbow is not bent, the nerve is still long enough to accommodate elbow flexion but is redundant upon itself; therefore, surface measurement across the skin will underestimate the true distance and the calculated conduction velocity will be too slow.

There has been considerable discussion in the literature about the appropriate distances to use between the above and below elbow stimulation sites. Earlier literature suggested that in general one should have at least 10 cm of distance between stimulation sites (67). However, this was based upon measurements of error in the 1970s when measuring latencies on equipment using much older technology. Similar studies have now been repeated utilizing modern digital equipment (68), and this has demonstrated that a 6-cm distance should usually be sufficient and would have error similar to the 10-cm distance 30 years ago.

When performing ulnar motor conduction studies, one must be aware of the potential of Martin-Gruber anastomosis. This anastomosis is said to be present in 15% to 20% of individuals and typically involves fibers crossing from the median nerve to the ulnar nerve in the proximal forearm (69). At times, the fibers can originate from the anterior interosseus nerve rather than from the main branch of the median nerve. In the presence of Martin-Gruber anastomosis, one will record a normal large amplitude response from the ADM and FDI when stimulating the ulnar nerve at the wrist. However, while stimulating the ulnar nerve at the elbow, one will note a decreased amplitude response because one is stimulating only the ulnar nerve fibers and not those that cross in the proximal forearm. To the inexperienced electromyographer, this can masquerade as conduction block in the proximal forearm and can result in an erroneous diagnosis. The hint of a Martin-Gruber anastomosis rather than ulnar neuropathy in the forearm or elbow is that this drop in amplitude occurs between wrist and below elbow and not across
the elbow. The presence of this anomalous innervation can be proven by stimulating the median nerve at the elbow and recording from the ADM and FDI muscles; when a crossover exists, a sizable response can be recorded from these usually ulnar innervated muscles (70).

After recording ulnar conduction motor studies across the elbow, one will want to decide whether these are normal or not. There have generally been two ways to do this. Many authors advocate comparing ulnar conduction across the elbow to that recorded in the forearm. However, this comparison is flawed in that it assumes that ulnar conduction in the forearm is unaffected by a neuropathy proximally at the elbow (66). Unfortunately, this is not the case since with motor axon loss there is distal slowing due to preferential loss of the faster conducting fibers. As a result, comparison between the two segments is not valid. The other method for determining whether the conduction is normal is to compare the velocity to reference values. This has been shown to be preferable in terms of sensitivity and specificity (66). Our laboratory uses a reference value of 48 m/s as a lower limit of normal.

When there is concern for ulnar neuropathy at the elbow, it is frequently useful to perform ulnar “inching” studies. These studies involve stimulation of the ulnar nerve at 2-cm increments across the elbow looking for any focal slowing or conduction block. Latency differences exceeding 0.7 ms or amplitude differences exceeding 10% are suggestive of a focal lesion (71). It is preferable to see both latency and amplitude changes as well as changes in morphology to be certain of a focal lesion. Because the distances are small, and the error in measurement is large as a percentage, one should not consider the conduction velocity of inching studies in m/s but rather look at the established reference values (≤0.7 ms) for latency differences across 2 cm.

Ulnar sensory conduction studies can be useful at times. When stimulating at the wrist and recording at the small finger, responses are usually small in amplitude or absent. It is difficult to reliably and consistently record ulnar sensory conduction across the elbow recording at the small finger. The response from the dorsal ulnar cutaneous nerve should be affected to a similar degree as the ulnar sensory response to the small finger.

Needle EMG should generally be performed in patients referred for ulnar neuropathy at the elbow including the ADM, FDI, and the flexor digitorum profundus. Remember, however, that the FDP is often spared in ulnar neuropathy at the elbow. When there are abnormalities in the ulnar innervated hand muscles, it is important to check non-ulnar innervated C8/T1 muscles to look for root or plexus lesions that might mimic an ulnar neuropathy. Generally, it is useful to check APB and extensor indicis proprius (EIP) in these situations.

### Radial Nerve at the Humerus

Radial nerve entrapments in the absence of trauma are much less common than entrapments of the median or ulnar nerves in the upper limb. This is in contrast to traumatic neuropathies where radial nerve injuries are more common due to the proximity of the nerve to the humerus as it traverses along the spiral groove. The radial nerve is most typically affected along the spiral groove of the humerus, after the nerve has given branches to the triceps and the anconeus muscles but before it gives branches to the brachioradialis and the finger and wrist extensors. The site of the lesion is also proximal to the division of the nerve into superficial and deep branches (the deep radial nerve was formerly known as the posterior interosseous nerve) (72).

Patients typically present with symptoms of weakness when opening their hand and extending their wrist. They also report weak grip strength since grip is weaker with the wrist in flexion than it is with an extension. A strong grip requires wrist extension to produce a tenodesis effect on the long finger flexors. Patients will also report numbness in the radial distribution on the dorsum of the hand with a proximal radial neuropathy.

On physical examination, the primary finding on strength testing will be weakness of wrist and finger extension (73). There are, however, several special points to keep in mind when examining the patient with possible radial neuropathy. First, it is difficult to isolate the brachioradialis muscle, which is the first muscle supplied after the spiral groove, on manual muscle testing. It is best to test this muscle with the forearm in the neutral position and to flex the elbow while palpating and observing the brachioradialis muscle visually. Although the biceps and other elbow flexors are strong enough to substitute for the brachioradialis, one can usually palpate or visually appreciate a side-to-side difference when this muscle is weak. Secondly, one should take special care in examining finger extension. The extensor digitorum and other finger extensors supplied by the radial nerve primarily produce extension at the MCP (metacarpal phalangeal) joints. It is the ulnar innervated lumbricals and interossei that contribute to extension at the interphalangeal (IP) joints. As a result, extension of the fingers at the proximal interphalangeal joint (PIP) and distal interphalangeal joint (DIP) may be intact even with a complete radial neuropathy. Finally, one should be aware that when a patient has a radial neuropathy and weak finger extension that testing finger abduction (i.e., the interosseous muscles) will produce apparent weakness. This is not because of true muscle weakness but rather because testing finger abduction while the MCP joints are in a flexed position produces much less force than when they are fully extended. For this reason, it is important to have the patient lay his or her hand on a table or a book to complete testing of interosseous muscle function.

Physical examination will also usually show reduced sensation in the radial distribution of the hand and an absent brachioradialis muscle stretch reflex. Since the lesion is commonly distal to the branch supplying the triceps, the reflex at the triceps is typically intact.

In cases of radial neuropathy, needle EMG is often more useful than nerve conduction studies. Both recording and stimulation of the radial nerve for motor conduction studies can be problematic. Recording of the radial motor response
is typically achieved with surface electrodes placed over the EIP muscle. This is sometimes satisfactory but does suffer from the problem that other forearm muscles in the posterior compartment also contribute to this response via volume conduction. At times, some electromyographers will use needle recording to avoid volume conduction. While this does allow greater accuracy of measurement of latencies, it is less helpful for determining the CMAP amplitudes. With needle electrode recording, the amplitudes of motor responses are much more variable and dependent upon precise needle placement. Thus, needle recording should not be used for estimating how many axons are available in the nerve.

Stimulation, while relatively easily achieved in the distal forearm, is more difficult to achieve at the elbow and arm. It is difficult to obtain reliable consistent supramaximal responses at the spiral groove or proximally at the axilla. Probably the most useful setting for motor conduction studies in the assessment of radial neuropathy, is to evaluate the surface amplitude for the EIP with stimulation of the nerve in the distal forearm. This amplitude has been shown to correlate with the prognosis in radial neuropathy (74). Radial sensory responses are useful to evaluate for evidence of sensory axon loss and to evaluate whether the lesion is proximal to the branching of the radial nerve into superficial and deep branches. However, one generally does not stimulate the radial nerve proximally while recording sensory responses distally in the hand.

Needle EMG is usually the most useful electrodiagnostic assessment. One should generally consider studying the triceps muscle to look for evidence of denervation. If there is denervation in this muscle, then it is usually wise to proceed proximally to the deltoid muscle, which is also innervated by the posterior cord of the brachial plexus. If this is abnormal as well, then a broader evaluation of the limb is clearly indicated. Distally, it is useful to study the brachioradialis muscle since this is the first muscle to be innervated after the spiral groove (the anconeus muscle is innervated by the same branch that supplies the medial head of the triceps). The other muscles in the forearm that are useful to study include the extensor carpi radialis, extensor digitorum, and the extensor indicus proprius. It has been demonstrated that two of the more useful prognostic signs in radial neuropathy are presence or absence of a radial motor response to the EIP and the degree of recruitment noted in the brachioradialis muscle (74).

**Fibular Neuropathy**

The fibular nerve (the preferred name for the nerve formerly known as peroneal) (72) is the most commonly affected nerve in the lower limb. It is particularly vulnerable to pressure as it crosses just behind the fibular head near the knee. It can also be affected by prolonged knee flexion or squatting as in prolonged labor and in those with occupations requiring prolonged squatting, such as strawberry pickers. This is the most common lower limb nerve injury in athletes (75).

The most common symptom of fibular neuropathy is weakness of dorsiflexion, presenting with foot slap or dragging of the foot when walking (76). Patients also have sensory loss over the dorsum of the foot but rarely is this a prominent or presenting complaint.

Electrodiagnostic assessment generally includes motor and sensory conduction studies as well as needle EMG in the lower limb (77). Motor nerve conduction studies are often performed with recording over the extensor digitorum brevis (EDB) muscle. The EDB muscle is useful when studying the fibular nerve in the leg region, and it provides a conduction velocity between the fibular head and the ankle. On the other hand, the EDB is not a functionally useful muscle, and in moderately severe fibular neuropathies, or in polyneuropathies, the response may be absent. Thus, it is wise to record from both the EDB and the tibialis anterior (TA) muscles at the same time. The TA has the advantage of being a functionally more important muscle, and it may at times be present when the EDB is absent (78).

Sensory responses can be recorded from the superficial fibular nerve. This is helpful to distinguish between a proximal L5 root lesion (in which sensory conduction should be normal) and a more distal fibular nerve lesion. However, it is usually not otherwise helpful for localization. This response can also at times be misleading since a selective deep fibular nerve lesion will spare the superficial fibular sensory response.

Needle EMG can be very helpful in localization of fibular nerve lesions. It is usually helpful to study the TA and fibularis longus (formerly peroneous longus) muscles to evaluate both the deep and superficial branches. In order to exclude a more proximal lesion, one studies the short head of the biceps femoris since this muscle is supplied by the fibular division of the sciatic nerve proximal to fibular head. If this muscle is abnormal, then a wider examination of the lower limb is indicated. The EDB muscle can be studied with needle EMG but has a high frequency of false-positive results (28). It is also often useful to study tibial innervated muscles in the leg to exclude a more proximal sciatic lesion or a proximal root lesion.

Prognosis of fibular nerve lesions in large part is determined by both the amplitude of the motor responses and by the degree of motor unit recruitment in the TA muscle (78). Those with good recruitment in TA and present CMAP responses in TA and EDB generally do quite well.

**Tibial Nerve**

Neuropathy affecting the tibial nerve is relatively rare compared to the other nerve lesions discussed above. This is likely related to the relatively protected position of the tibial nerve within the leg and foot. Moreover, the fascicular anatomy of the tibial nerve is such that it has many small fascicles that are resistant to injury (79). In comparison, the fibular nerve is made up of a few larger fascicles that are more susceptible to entrapment and trauma as well as stretching (79). The tibial nerve can occasionally be injured at the ankle. It is relatively common to have posttraumatic tibial neuropathy at the ankle after calcaneal fractures or other foot injuries. Much
more controversial, however, is the possible diagnosis of tarsal tunnel syndrome. Patients with tibial neuropathy at the ankle can present with numbness or paresthesias in the sole of the foot (80). They may have predominant symptoms in either the lateral or medial plantar nerve distribution or occasionally in the calcaneal nerve distribution. As noted above, this can occur after trauma to the ankle or foot but may also be occasionally seen without history of injury. The etiology of nontraumatic tarsal tunnel syndrome is unclear. Some studies have shown that varicose veins, accessory muscles, or other space occupying lesions within the tarsal tunnel may be noted in these patients based on MRI examinations (81,82). Other investigators hypothesize that the hyperpronated foot, especially in runners, can predispose the tarsal tunnel syndrome, but this is more controversial.

There are several nerve conduction studies that can be performed to evaluate for possible tarsal tunnel syndrome (80,83). Motor conduction studies can be performed to muscles supplied by the medial and the lateral plantar nerves. One stimulates the tibial nerve at the ankle and records from the abductor hallucis and the abductor digiti quinti pedis at the same time. There are reference values for these two muscles in the literature, but one must be especially careful about temperature control since cold feet will markedly prolong the distal latencies. These motor studies are reported to be less sensitive than CNAP studies in the diagnosis of tarsal tunnel syndrome (84).

One can also perform CNAPs of the medial and lateral plantar nerves in the foot (85). This involves stimulation of the medial and lateral plantar nerves on the sole of the foot, recording proximally just above the medial malleolus. This is a CNAP rather than an SNAP since one is in part activating motor axons antidromically in addition to the orthodromic sensory potentials. Reportedly, these studies are more sensitive for demonstrating abnormalities than motor nerve conduction studies. However, these CNAP responses are usually quite small and often difficult to obtain. Thus, absence of a response is not diagnostic by itself. An asymmetrically delayed latency is more suggestive of a diagnosis of tarsal tunnel syndrome.

Needle EMG of intrinsic foot muscles can also be useful to assess for tibial neuropathy at the ankle. Some authors believe that tarsal tunnel syndrome can be a primarily axonal injury and thus EMG may be more useful at detecting this axon loss than the nerve conduction studies (86). This point, however, is not universally accepted. Needle EMGs are usually performed on a medial planar innervated muscle such as the abductor hallucis and a lateral planar innervated muscle such as one of the interossei. The FDI muscle can be used, but there is a high incidence of this muscle being jointly innervated by the deep fibular nerve (87). Thus, it is preferable to study the second dorsal interosseous muscle, between the second and third metatarsal bones. One must be cautious in interpreting findings on needle EMG in the intrinsic foot muscles since some studies have demonstrated a high incidence of false-positive results in otherwise healthy individuals (28). The abductor hallucis and interosseous muscles are less likely to show abnormalities than the EDB. However, one should still be cautious in interpreting mild or subtle changes.

The author believes that nontraumatic tarsal tunnel syndrome is a very rare condition and has personally seen less than five convincing nontraumatic cases after having done more than 10,000 electrodiagnostic studies!

**Traumatic Neuropathies**

Traumatic injury to peripheral nerves results in considerable disability across the world. In peacetime, peripheral nerve injuries commonly result from trauma due to motor vehicle accidents and less commonly from penetrating trauma, falls, and industrial accidents. Of all patients admitted to level I trauma centers, it is estimated that roughly 2% to 3% have peripheral nerve injuries (88,89). If plexus and root injuries are also included, the incidence is about 5% (88).

In the upper limb, the nerve most commonly reported injured is the radial nerve, followed by ulnar and median nerves (88,89). Lower limb peripheral nerve injuries are less common, with the sciatic most frequently injured, followed by fibular, and rarely tibial or femoral nerves.

Seddon has used the terms “neurapraxia,” “axonotmesis,” and “neurotmesis” to describe peripheral nerve injuries (90). Neurapraxia is a comparatively mild injury with motor and sensory loss but with no evidence of wallerian degeneration. The nerve conducts normally distally. Focal demyelination and ischemia are thought to be the etiologies of conduction block. Recovery may occur within hours, days, weeks, or up to a few months. Axonotmesis is commonly seen in crush injuries, nerve stretch injuries (such as from motor vehicle accidents and falls), or percussion injuries (such as from blast wounds). The axons and their myelin sheaths are broken, yet the surrounding stroma (i.e., the Schwann tubes, endoneurium, and perineurium) is at least partially intact. Degeneration occurs, but subsequent axonal regrowth may proceed along endoneurial tubes if they are sufficiently preserved. Recovery ultimately depends upon the degree of internal disorganization in the nerve as well as the distance to the end organ. Neurotmesis describes a nerve that has been either completely severed or is so markedly disorganized by scar tissue that axonal regrowth is impossible. Examples are sharp injury, some traction injuries, and percussion injuries or injection of noxious drugs. Prognosis for spontaneous recovery is extremely poor without surgical intervention.

Optimal timing of electrodiagnostic studies will vary according to clinical circumstances. For circumstances in which it is important to define a lesion very early, initial studies at 7 to 10 days may be useful at localization and separating conduction block from axonotmesis. On the other hand, when clinical circumstances permit waiting, studies at 3 to 4 weeks postinjury will provide much more diagnostic information, because fibrillations will be apparent on needle EMG. Finally, in cases where a nerve lesion is surgically confirmed and needle EMG is used only to document
recovery, initial studies at a few months postinjury may be most useful.

Changes may be seen in the CMAP, late responses (F and H waves), SNAP, and needle EMG. Each of these studies has a somewhat different time course, which should be understood in order to evaluate peripheral nerve injury; they will also vary according to the severity of nerve injury.

In purely neurapraxic lesions, the CMAP will change immediately after injury, assuming one can stimulate both above and below the site of the lesion (Fig. 4-15). When recording from distal muscles and stimulating distal to the site of the lesion, the CMAP should be normal because no axonal loss and no wallerian degeneration have occurred. Moving the site of stimulation proximal to the lesion will produce a smaller or absent CMAP, as conduction in some or all fibers is blocked. In addition to conduction block, partial lesions also often demonstrate concomitant slowing across the lesion. This slowing may be due to either loss of faster conducting fibers or demyelination of surviving fibers.

Electrodiagnostically, complete axonotmesis and complete neurotmesis cannot be differentiated, because the difference between these types of lesions is in the integrity of

**FIGURE 4-15.** Findings on motor nerve conductions with stimulation above and below the site of the lesion. Neurapraxia cannot be differentiated from axonotmesis or neurotmesis at day 1 but can be differentiated at day 10, after wallerian degeneration has occurred. The extent of axon loss can be determined by the CMAP amplitude obtained with stimulation distal to the lesion after day 10.
the supporting structures, which have no electrophysiologi-
cal function. Thus, these lesions can be grouped together as
axonotmesis for the purpose of this discussion. Immediately
after axonotmesis and for a “few days” thereafter, the CMAP
and motor conduction studies look the same as those seen in
a neurapraxic lesion. Nerve segments distal to the lesion remain
excitable and demonstrate normal conduction, whereas prox-
imal stimulation results in an absent or small response from
distal muscles. Early on, this picture looks the same as con-
duction block and can be confused with neurapraxia. Hence,
neurapraxia and axonotmesis cannot be distinguished until
sufficient time for wallerian degeneration in all motor fibers
has occurred, typically about 9 days postinjury (91).

After enough time has passed for wallerian degeneration
to occur, the amplitude of the CMAP elicited with distal stim-
ulation will fall. This starts at about day 3 and is complete by
about day 9 (91). Thus, in complete axonotmesis at day 9, one
has a very different picture from neurapraxia. There are absent
responses both above and below the lesion. Partial axon loss
lesions will produce small-amplitude motor responses, with
the amplitude of the CMAP roughly proportional to the num-
ber of surviving axons.

F waves may change immediately after the onset of a
neurapraxic lesion. In complete conduction block, responses
will be absent. However, in partial lesions, changes can be
more subtle, because F waves are dependent upon only 3% to
5% of the axon population to elicit a response (2). Thus,
partial lesions may have normal minimal F-wave latencies and
mean latencies, with reduced or possibly normal penetrance.
Although F waves are conceptually appealing for detecting
proximal lesions (e.g., brachial plexopathies), only in a few
instances do they truly provide useful additional or unique
information. They are sometimes useful in very early proximal
lesions when conventional studies are normal because stimula-
tion does not occur proximal to the lesion.

The SNAP and CNAP show changes similar to the CMAP
after focal nerve injury. With neurapraxia, there is focal con-
duction block at the site of the lesion, with preserved distal
distance. Immediately after axonotmesis and for a few days
thereafter, the SNAP looks the same, as with a neurapraxic
lesion. Nerve segments distal to the lesion remain excitable and
demonstrate normal conduction, whereas proximal stimulation
results in an absent or a small response. It takes slightly lon-
ger for sensory nerve studies to demonstrate loss of amplitude
than for motor studies, that is, 11 days versus 9 days, due to
the earlier failure of NMJ transmission compared with nerve
conduction.

On needle EMG, in neurapraxic lesions, the most apparent
changes will be in recruitment. These occur immediately after
injury. In complete lesions (i.e., complete conduction block),
there will be no MUAPs. In incomplete neurapraxic lesions,
there will be reduced numbers of MUAPs firing more rapidly
than normal (i.e., reduced or discrete recruitment). Because no
axon loss occurs in neurapraxic injuries, there will be no axial
sprouting and no changes in MUAP morphology (e.g., dura-
tion, amplitude, or phasicity) at any time after injury.

With axonotmesis, needle EMG will demonstrate fibril-
lation potentials and positive sharp waves a number of days
after injury. The time between injury and onset of fibrillation
potentials will be dependent in part upon the length of the
distal nerve stump. When the distal nerve stump is short, it
takes only 10 to 14 days for fibrillations to develop. With a
longer distal nerve stump (e.g., ulnar-innervated hand mus-
cles in a brachial plexopathy), 21 to 30 days is required for
full development of fibrillation potentials and positive sharp
waves (92). Thus, the electrodiagnostic medicine consultant
needs to be acutely aware of the time since injury so that
severity is not underestimated when a study is performed early
after injury, and also that development of increased fibrilla-
tion potentials over time is not misinterpreted as a worsening
of the injury.

Fibrillation potential size decreases over time since injury.
Kraft (93) has demonstrated that fibrillations initially are sev-
eral hundred microvolts in the first few months after injury.
However, when lesions are more than 1 year old, they are
unlikely to be more than 100 μV in size.

Fibrillations may also occur after direct muscle injury as
well as nerve injury. Partanen and Danner (94) have demon-
strated that patients after muscle biopsy have persistent fibril-
lation potentials starting after 6 to 7 days and extending for
up to 11 months. In patients who have undergone multiple
trauma, coexisting direct muscle injury is common and can be
potentially misleading when trying to localize a lesion.

When there are surviving axons after an incomplete
axonal injury, remaining MUAPs are initially normal in mor-
phology but demonstrate reduced or discrete recruitment.
Axonal sprouting will be manifested by changes in morphol-
y of existing motor units. Amplitude will increase, dura-
tion will become prolonged, and the percentage of polyphasic
MUAPs will increase as motor unit territory increases (95,96).
In complete lesions, the only possible mechanism of recovery is
axon regrowth. The earliest needle EMG finding in this case
is the presence of small, polyphasic, often unstable MUAPs
previously referred to as “nascent potentials.” (This term is
now discouraged because it implies an etiology; it is preferred
to simply describe the size, duration, and phasicity of the
MUAP.) Observation of these potentials is dependent upon
establishing axon regrowth as well as new NMJs, and this find-
ing represents the earliest evidence of reinnervation, usually
preceding the onset of clinically evident voluntary movement
(95). These potentials represent the earliest definitive evidence
of axonal reinnervation in complete lesions. When performing
the examination looking for new MUAPs, one must be sure
to accept only “crisp,” nearby MUAPs with a short rise time,
because distant potentials recorded from other muscles can be
deceptive.

Determining the pathophysiology of a peripheral nerve
traumatic injury can help with estimating prognosis. Those
injuries that are completely or largely neurapraxic have a good
prognosis for recovery within a few months (usually up to
3 months postinjury). Resolution of ischemia and remyelina-
tion should be complete by this time.
In axonotmesis, recovery will depend upon axonal sprouting and regeneration. Hence, there will be some early recovery followed possibly by a later recovery if or when regenerating axons reach their end organs. The amplitude of the CMAP provides some guide to prognosis. The amplitude of CMAPs or more of the other side have an excellent outcome, those with 10% to 30% have good but not always complete recovery, and those with <10% have a poor outcome (97). There is some evidence, however, that outcomes are better in peripheral nerves than in the facial nerve (74,78).

Complete axonotmesis and neurotmesis have the worst prognosis. Recovery depends solely upon axonal regeneration, which may or may not occur, depending upon the degree of injury to the nerve. In many cases of complete axon loss, it is not possible to know the degree of nerve injury except by surgical exploration with or without intraoperative recording or looking for evidence of early reinnervation after the lesion. As a consequence, it is often recommended to wait 2 to 4 months and look for evidence of reinnervation in previously completely denervated muscles near the site of the lesion (98,99). Those lesions that have some spontaneous recovery are usually treated conservatively, because operative repair is unlikely to improve upon natural recovery. Those with no evidence of axonal regrowth may have operative exploration with possible grafting.

Radiculopathy
Radiculopathy is another very common reason for assessment in the electrodiagnostic laboratory. The electrodiagnostic evaluation can be useful in assessing for the presence of radiculopathy, localizing the lesion, and exploring other possible diagnoses that may be confused with root disease (11).

The typical clinical presentation of radiculopathy is well covered in other chapters. Most commonly, it involves pain starting from the neck or the back and extending into the upper or lower limb. It is often accompanied by sensory symptoms and sometimes by weakness. Physical examination may show weakness, depressed reflexes, and positive relevant provocative signs (such as straight leg raise test or Spurling’s sign). Although sensory loss is often found, this is not as predominant or as well circumscribed as it might be in peripheral nerve lesions.

The etiology of radiculopathy can be multifactorial (100). There are certain structural lesions such as disk protrusions and bony lesions that can press upon the root. However, the electrodiagnostic medical consultant should keep in mind that there are other causes of radiculopathy as well. Neoplasms either at the root or at the spinal cord level can look like a radiculopathy. There are metabolic or inflammatory lesions such as diabetes or vasculitis, which can produce a radiculopathy without significant structural changes apparent on imaging studies. There are also infectious causes such as herpes zoster that can affect the roots. Thus, not all radiculopathies are caused by disk protrusion.

In general, the most useful electrodiagnostic assessment in radiculopathy is needle EMG. There have been studies suggesting how many muscles should be studied in a limb when assessing for possibly radiculopathy (101,102). It is generally preferable to examine at least five muscles in the lower limb plus paraspinal muscles and five muscles in the upper limb plus paraspinal muscles to reach a high degree of sensitivity on needle EMG. When the paraspinal muscles are not available, such as after surgical intervention, then one should increase the number of limb muscles to eight to achieve a similar degree of sensitivity.

The selection of muscles should be tailored to each patient depending on which root level is in question. For example, if the patient’s symptoms are more in the C6 distribution, one should include more C6 level muscles and perhaps fewer muscles at the C8 or T1 level. When symptoms are nonspecific, then one should design an electrodiagnostic evaluation that includes multiple roots and multiple peripheral nerves within the limb. Examples of commonly used screens are given below:

<table>
<thead>
<tr>
<th>Upper Limb Radiculopathy Screen</th>
<th>Lower Limb Radiculopathy Screen</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deltoid</td>
<td>Vastus medialis</td>
</tr>
<tr>
<td>Biceps</td>
<td>Tensor fascia lata</td>
</tr>
<tr>
<td>Pronator teres</td>
<td>Tibialis anterior</td>
</tr>
<tr>
<td>Triceps</td>
<td>Medial gastrocnemius</td>
</tr>
<tr>
<td>FDI</td>
<td>Biceps femoris long head</td>
</tr>
<tr>
<td>Cervical paraspinal muscles</td>
<td>Lumbar paraspinal muscles</td>
</tr>
</tbody>
</table>

The key to diagnosing radiculopathy is to find evidence of denervation in at least two muscles within a single myotome (the distribution of muscles supplied by a single root) but supplied by different peripheral nerves (11). For example, evidence of denervation in the TA and fibularis longus muscles would not be sufficient to diagnose an L5 radiculopathy because both muscles are within the common fibular nerve distribution. However, the same changes in the TA and tensor fascia lata would be much more suggestive of a root lesion.

It is important to consider carefully which types of electrodiagnostic findings are sufficient to be diagnostic. Most electromyographers would agree that evidence of the following potentials in a myotomal distribution would be diagnostic:

- Fibrillation potentials
- Positive sharp waves
- CRD
- Fasciculation potentials

It is more controversial as to which motor unit changes can be considered diagnostic. If there are very large amplitude, long duration, polyphasic MUAPs in a myotomal distribution, then one might interpret the study as suggestive of chronic radiculopathy with reinnervation. However, one should not
diagnose a radiculopathy based solely upon increased numbers of polyphasic MUAPs, since these are often found in otherwise healthy individuals in many muscles.

The paraspinal muscles can be very useful to help localize a lesion to the root level, since these muscles are supplied by the posterior primary ramus of the root. Evidence of denervation in these muscles, combined with the limb muscles, is highly suggestive of root pathology. At the same time, however, one must be cautious in interpreting paraspinal findings. Some authors have reported the presence of fibrillation potentials and positive sharp waves in otherwise healthy individuals in the lumbar and cervical paraspinal muscles, with an increasing frequency with age (27,103). Moreover, there are other causes for fibrillation potentials in the paraspinal muscles besides radiculopathy. For instance, myopathies, prior spine surgery, recent prior needle intervention, metastatic muscle lesions, and diabetes can all cause paraspinal abnormalities without significant radiculopathy. Moreover, the paraspinal muscles are not root localizing since most of the muscles are multisegmentally innervated (104). Localization is primarily achieved through analysis of limb muscle findings.

In addition to needle EMG, H waves can be useful in diagnosing S1 radiculopathies. H waves are similar in their physiology to the ankle jerk, except that stimulation occurs proximal to the muscle spindles. The primary advantage of H waves in the assessment of radiculopathy is that they will detect sensory axon loss, conduction block, and demyelination in the root, whereas needle EMG only detects loss of motor axons (19). Side-to-side differences in H-wave latency exceeding 1.2 to 1.5 ms are likely abnormal. One should be cautious, however, in interpreting H waves since any lesion along the pathway (from the soleus muscle to the S1 root) can cause abnormalities in this response. Moreover, H-wave abnormalities do not indicate chronicity, and these responses can be abnormal in very acute or very chronic or old root lesions.

Sensory responses should usually be normal in root pathology since here the lesion is proximal to the DRG. F waves are generally of little use in evaluation of radiculopathy. It is believed that since muscles are supplied by more than one root, the responses traverse more than one root proximally, and a single intact root can produce normal F-wave latencies.

The electrodiagnostic medical consultant is often asked about the sensitivity of needle EMG in comparison to imaging studies. Generally, the EMG is considered to have sensitivity for radiculopathy in the range of 70% to 80% compared to clinical presentation (101,105,106). This is in contrast to imaging studies that are often reported with sensitivities exceeding 90%.

Sensitivity, however, is not the only figure of merit. One should be equally or even more concerned about specificity, that is, the ability to detect absence of disease and to avoid false-positive findings (59). Specificity of needle EMG, using the approach discussed above, is excellent, likely exceeding 95%. If one were to make a diagnosis based solely upon paraspinal findings or a few polyphasic motor units, the specificity would be less. By contrast, the specificity of imaging is much lower. Many studies have reported imaging specificities in the 60% to 70% range, with a false-positive rate of 30% to 40% (107). Thus, a primary advantage of needle EMG is the high specificity it offers in comparison to imaging studies.

**Plexopathy**

The assessment of brachial and lumbosacral plexopathy is one of the more challenging aspects of the electrodiagnostic examination (108). An important and elementary requirement is that the electrodiagnostic physician be well acquainted with the anatomy of the brachial plexus or in some cases, the lumbosacral plexus. In this author’s experience, it is usually necessary for the electrodiagnostic consultant to be able to draw the brachial plexus with all branches in less than a minute in order to have a sufficient working knowledge of the anatomy. There are Web sites available to help with drawing the anatomy easily (e.g., http://www.ama-assn.org/ama1/pub/upload/mm/15/brachial_plex_how_to.pdf). Although one could look up the brachial plexus online or in text books, without this solid working knowledge, one will not think of patterns that represent plexus lesions and will potentially miss some important diagnoses.

The etiologies of brachial plexopathy are varied. Most common is trauma, both in children and in adults. Motorcycle crashes, motor vehicle crashes, falls, and industrial injuries account for the great majority of brachial plexus lesions. It is estimated that 2% to 3% of patients coming to level 1 trauma centers have a brachial plexus injury (88). There are other causes of brachial plexus lesions such as neuralgic amyotrophy, radiation plexitis, tumor invasion, true neurogenic thoracic outlet syndrome, and others.

Lumbosacral plexopathies are much less common. Traumatic lumbosacral plexopathies occur in less than 1% of patients with pelvic fractures (109). Moreover, the lumbosacral plexus is less affected by other diseases than the brachial plexus. The placement of the lumbosacral plexus within the pelvis offers a great deal of protection compared to the relatively vulnerable brachial plexus in the upper limb and shoulder. The lumbosacral plexus nevertheless can be affected by trauma, radiation plexitis, diabetes, and neurogenic amyotrophy, all of which are rare.

Patients with plexopathy can present with a varied presentation. Typically, there is pain in the limb (but not the spine) and there are sensory symptoms in the limb. Occasionally, patients present with a plexopathy after breast cancer and subsequent radiation; those with radiation plexitis typically have upper trunk distribution paresthesias, whereas those with tumor invasion present with lower trunk painful lesions (34). Physical examination is usually remarkable for weakness, significant sensory loss outside of a single peripheral nerve distribution, and reduced reflexes.

Sensory nerve conduction studies can be very useful to differentiate plexus lesions from more proximal root pathology. As has been discussed above, root lesions will have normal SNAPs, whereas plexus lesions (distal to the DRG) will have
abnormal or absent sensory responses. This fact can be used to screen the brachial plexus. For example, one can evaluate the upper trunk through the use of median and radial nerve sensory testing to the thumb (C6). The middle trunk can be evaluated by stimulating the median nerve and recording from the long finger (C7), and the lower trunk can be evaluated by testing the ulnar sensory potential to the small finger (C8). If these are all intact, it would argue strongly against a brachial plexus lesion affecting the upper, middle, or lower trunk. Similarly, one can use SNAPs in the lower limb to evaluate for peripheral nerve or plexus lesions versus root pathology.

The reader should be aware of a rather unique setting in which the presence of normal sensory potentials is a very poor prognostic sign. When a patient is evaluated for brachial plexopathy, at times they will have complete root lesions such as root avulsion rather than a more distal brachial plexus lesion. When the roots have been avulsed from the spinal cord, this is a very poor prognosis, and there is essentially no likelihood of recovery. In this setting, the SNAPs will be normal because the dorsal root ganglia have been avulsed along with the plexus, and there is no subsequent degeneration of sensory axons. However, needle EMG will be consistent with complete denervation and no motor responses will be obtainable. In this case, presence of normal sensory potentials generally suggests root avulsion and a very poor prognosis for recovery, though these findings do not always agree with imaging studies (110).

Motor nerve conduction studies are also useful in the assessment of brachial plexopathy, particularly for prognosis. One can determine the degree of axon loss by recording from an affected muscle and stimulating the nerve distal to the lesion. For example, in the case of an upper trunk plexus lesion, one might stimulate the musculocutaneous nerve in the axilla and record from biceps. A large amplitude response here would indicate a rather good prognosis since the lesion does not have severe axon loss.

Needle EMG can also be helpful in diagnosing and localizing brachial plexopathy. These are usually very extensive evaluations. This author generally finds it helpful to draw out the plexus in question and make sure that a muscle supplied by each of the significant branches off the plexus is tested on needle EMG. Paraspinal muscles are also very useful because abnormalities there will suggest a more proximal root lesion.

F-wave responses can be helpful early on in a possible brachial plexus lesion, before motor or sensory conduction studies have changed and before needle EMG shows abnormalities. Absence of F waves may indicate a proximal lesion particularly if distal conduction is intact. However, it should be remembered that an abnormality anywhere along the course of the F wave will produce an abnormal response. In the setting of more chronic pathology, it is rare that the F wave will show a unique abnormality that is not apparent on other nerve conduction studies or needle EMG.

**Polynuropathies**

The goal of electrodiagnostic evaluation in assessing peripheral polyneuropathies is to both detect abnormalities and classify the neuropathy into one of six groups. One should also determine whether there are other lesions (such as spinal stenosis) that can produce similar clinical presentations.

There are generally six categories in which one can place peripheral polyneuropathies. These are well covered in other references (111). The six categories will be discussed briefly below.

**Uniform Demyelinating Neuropathies**

Uniform demyelinating neuropathies are inherited neuropathies in which myelin is abnormal throughout the peripheral nervous system. Examples include Charcot-Marie-Tooth disease (type 1), Refsum’s disease, Dejerine-Sottas disease, and others. Because of the diffuse abnormality of myelin, one sees diffuse nonsegmental changes. In these uniform demyelinating neuropathies, one sees diffuse slowing of nerve conduction velocities (without patchiness), prolonged distal latencies, and prolonged or absent F-wave responses. There is no conduction block or temporal dispersion. Generally, motor fibers are more affected than sensory fibers. Needle EMG can reveal mild evidence of distal denervation, but this is not as marked as the slowing that is apparent on nerve conduction studies.

**Segmental Demyelinating Neuropathies**

Segmental demyelinating neuropathies are acquired neuropathies (not inherited) that present with patchy demyelination along the peripheral nervous system. Examples include Guillain-Barre syndrome, monoclonal gammopathies of undetermined significance, chronic inflammatory demyelinating polyradiculoneuropathy (CIDP), multiple myeloma, and others. The hallmark of this category is patchiness of abnormalities. Generally, there are five findings seen on nerve conduction studies in this group including (a) prolonged distal latencies, (b) prolonged or absent F-wave responses, (c) patchy slowing (faster in some areas, slower in other areas of the same nerve), (d) conduction block, and (e) temporal dispersion. This can be differentiated from the inherited neuropathies mentioned above by the presence of patchy slowing, conduction block, and temporal dispersion. Needle EMG will often show patchy denervation in more severe cases. Motor findings are more prominent than sensory findings in these patients. Many of these patients will have sparing of sural SNAPs but relatively greater abnormalities noted in the upper limb sensory responses. Prognosis in these cases has been shown to correlate best with the distal CMAP amplitudes (112). When amplitudes are larger than 10% of the lower limit of normal at the nadir of the illness, prognosis is generally good.

**Sensory Neuropathy or Neuronopathy**

Sensory neurons can be selectively affected at their cell bodies (neuronopathy) or axons (neuropathy). There are a variety of lesions that predominantly affect sensory nerves. At the cell body level, there is an entity that is known as dorsal root ganglionitis, or sensory neuronopathy. This represents death of the cell body within the DRG and subsequent loss of sensory axons. Most commonly, this is seen as a paraneoplastic presentation, but occasionally, autoimmune diseases can present with the same findings (113).
Selective loss of sensory axons distally can be seen in a variety of toxins such as B6 toxicity, cis-platin exposure, and others.

The hallmark of this category is the absence of SNAPs or reduction in their amplitudes. When these presentations are severe, one will be unable to obtain a sensory response anywhere in the limbs. In general, motor nerve conduction and needle EMG are normal as expected with selective involvement of sensory nerves.

**Motor Greater than Sensory Axon Loss**

Motor greater than sensory neuropathies are rare. In evaluating these patients, one should be wary of other diseases that can present with similar findings such as motor neuron disease, NMJ disease, or myopathy. Selective motor neuropathies can be seen in acute porphyria, heavy metal exposure, and exposure to vinca alkaloids, amongst other conditions. Occasionally, inherited neuropathies (e.g., Charcot-Marie-Tooth disease type 2) also present this way.

The primary abnormality in this category is reduced amplitude or absent motor responses with relatively normal SNAPs. Needle EMG usually shows length-dependent evidence of denervation in the limb muscles with more distal muscles showing greater abnormalities than proximal limb muscles. There is usually no significant slowing of nerve conduction studies in this group, although mild slowing could be present due to loss of faster conducting fibers.

**Motor and Sensory Axon Loss**

This is the most common presentation of polyneuropathy and has the greatest number of etiologies amongst all the categories. This group includes the majority of toxic exposures, paraneoplastic conditions, infectious diseases (such as HIV), and other etiologies. One sees reduced amplitude sensory and motor amplitudes on nerve conduction studies with relatively little slowing. Needle EMG also shows length-dependent evidence of denervation worse in the distal muscles of the limb. In many cases, the exact etiology of these polyneuropathies remains undetermined even after extensive investigation. Only about 60% of these patients will ultimately have an etiologic diagnosis.

**Motor and Sensory Axon Loss and Demyelination**

There are two common diseases that present with motor and sensory axon loss and demyelination: diabetes and uremia. These two diseases present with a mixture of reduction of SNAP amplitudes, reduced amplitude CMAPs, mild to moderate slowing of conduction, and evidence of denervation worse distally in the limbs.

When evaluating possible polyneuropathy, it is reasonable to start by performing motor and sensory conduction studies and F waves in one upper limb and one lower limb. This author’s approach is generally to perform the following nerve conduction: sural sensory, fibular motor, fibular F wave, ulnar sensory, ulnar motor, and ulnar F wave (the advantage of ulnar nerve studies is that it avoids possible misinterpretation resulting from median neuropathy at the wrist). If the symptoms are primarily in the lower limb and the lower limb nerve conduction studies are normal, it may not be necessary to study the upper limb.

Needle EMG should be performed routinely because it is more sensitive at detecting motor axon loss than nerve conduction studies. The TA and soleus muscles are useful muscles, and occasionally, one may study an interosseous muscle in the foot. When there are abnormalities seen in the lower limb, other muscles should be studied in the lower limb as well as some upper limb muscles to determine whether this represents a distal greater than proximal gradient.

The electromyographer needs to be aware of the effects of temperature upon nerve conduction studies. If the limb is excessively cold, less than 32°C, it is generally preferable to warm up the limb with a hairdryer, warm water, or other method. The use of correction factors, particularly for correction across many degrees of temperature, is discouraged.

In writing the report, it is usually not possible to definitively indicate the cause of the polyneuropathy in the patient being studied. However, one can indicate the category as discussed above and include a differential diagnosis that fits with the clinical presentation.

**Myopathy**

EMG can be very useful in the assessment of patients with possible muscle disease. Generally, patients presenting with a possible myopathy will primarily report symptoms of proximal weakness. They often report difficulty climbing stairs, difficulty brushing their hair or teeth, and at times difficulty with breathing or coughing. Some patients will have muscle pain or tenderness. If there are any sensory symptoms, then an alternative diagnosis should be considered. On examination, one will generally note proximal weakness greater than distal weakness. A Trendelenburg sign is often present and may be either compensated or uncompensated. Reflexes are usually preserved until late in the progression of the disease.

The diagnosis of myopathy depends upon multiple evaluations. First, the clinical presentation needs to be consistent with myopathy. In addition, laboratory values should usually show elevated muscle enzymes such as creatine kinase. Muscle biopsy is often important to arrive at a more definitive diagnosis in terms of what type of myopathy is present. Finally, EMG can demonstrate abnormalities that can help with determining the presence, activity, and sometimes help with diagnosis of the myopathy (114,115).

One should approach the patient with myopathy by studying only one side of the patient (right or left). This is important because many patients will have a subsequent muscle biopsy, and it is preferable to avoid histological examination of muscles that have had a recent needle EMG examination. Muscle trauma from the EMG needle can sometimes mimic changes from inflammation.

It is often helpful to study both proximal and distal muscles in the limb to look for a proximal greater than distal gradient since proximal muscles should be affected considerably more
than distal muscles. One major exception to this rule is myotonic dystrophy that presents with distal findings greater than proximal (116). Even though they are not easily biopsied, the EMG evaluation should usually include paraspinal muscles as they are very sensitive. It is also wise to include muscles that can be biopsied on the contralateral side such as quadriceps, biceps, or deltoid.

On needle EMG, many acquired myopathies and inherited dystrophies will be associated with fibrillations and positive sharp waves, especially when there is inflammation. It is hypothesized that inflammation induces segmental necrosis. As a result of this focal necrosis, distal portions of the muscle fibers are functionally denervated and fibrillate. When there is no inflammation, such as with chronic steroid myopathy (which produces selective type II fiber atrophy), needle EMG will be relatively normal at rest. In inflammatory myopathies, the presence of fibrillation potentials and positive sharp waves is suggestive of more active disease, whereas muscles studied in those with treated disease will often have no significant spontaneous activity (117,118). CRDs are commonly seen in myopathies, especially in chronic inflammatory myopathies, though they are not specific for muscle disease. As mentioned earlier, MUAPs in myopathies are generally small in amplitude, brief in duration, and sometimes polyphasic. They are recruited in an “early” pattern with many MUAPs firing, despite small amounts of force.

On nerve conduction studies, SNAPs should generally be normal without any significant slowing or changes in amplitude. On the other hand, CMAPs may be reduced in size, proportional to the degree of muscle fiber loss. Conduction velocity is usually preserved. At times, one will see unusual potentials in muscles that will help refine the diagnosis. For instance, the presence of myotonic discharges will indicate that the patient may have a myotonic syndrome such as myotonic dystrophy, myotonia congenita, acid maltase deficiency, or other rare diseases.

After the needle EMG has been completed, it is often helpful for the electromyographer to suggest some potential contralateral muscles for subsequent muscle biopsy. Generally, it is good to select a muscle with moderate disease. If one selects a muscle with very severe electromyographic changes, often the muscle biopsy will be read as end-stage muscle disease and will offer little specificity as to the etiology.

It may be useful to repeat the EMG in inflammatory myopathies treated with steroids. In those patients with polymyositis who have initial recovery but later weakness, there is often a question as to whether this represents recurrence of the inflammatory disease or new steroid myopathy. In disease recurrence, one sees increased spontaneous activity such as fibrillation potentials and positive sharp waves (119). However, if a patient has steroid myopathy, which does not cause segmental necrosis or muscle fiber loss, one will see a relatively normal needle EMG in terms of both spontaneous activity and initially recruited MUAPs (120).

Critical illness myopathy is probably the most common myopathy seen in patients admitted to the hospital. This commonly coexists with critical illness polyneuropathy, but the former has a better prognosis (121). Findings are not always pronounced but usually demonstrate small amplitude CMAPs, normal SNAPs (unless there is coexisting polyneuropathy), and proximal greater than distal needle EMG abnormalities. It is often difficult to examine MUAPs since these patients are often not alert enough to produce a muscle contraction. There are recently described techniques for direct muscle stimulation that have suggested that much of the weakness noted in the ICU is due to critical illness myopathy (122).

**Motor Neuron Disease**

Motor neuron disease can have multiple presentations, which are well covered in other chapters in this textbook, the most common of which is amyotrophic lateral sclerosis (ALS). ALS is characterized by both upper and lower motor neuron loss as well as involvement of both bulbar and limb muscles. ALS usually starts off with focal weakness without any sensory loss and with little or no pain. The weakness can be in the distal upper or lower limb or, especially in the older individuals, in the bulbar musculature. Fasciculations are often noted by the patient and reported to the physician.

There are several symptoms and/or signs that make the diagnosis of ALS unlikely including presence of sensory symptoms such as numbness or tingling, presence of urinary incontinence, or deficits in extraocular muscles.

There are other variants of motor neuron disease in addition to ALS (123). Primary lateral sclerosis (PLS) presents with selective involvement of the corticospinal tracks in the brain and spinal cord. These patients have upper motor neuron syndromes without significant atrophy or evidence of lower motor neuron loss. They have a somewhat better prognosis than ALS but some eventually progress into ALS. Progressive bulbar palsy (PBP) starts off with weakness primarily in the bulbar muscles. Most commonly, patients present with slurring of speech or difficulty swallowing. This often quickly evolves into full blown ALS and has a worse prognosis than if the onset were in the limbs instead. Spinal muscular atrophy (SMA) is a predominantly lower motor neuron loss that has a slower disease progression and somewhat better prognosis than ALS. It does not have prominent upper motor neuron features. Monomelic amyotrophy looks like ALS clinically but is confined to one limb and has a much better prognosis. It usually has both upper and lower motor neuron features in one limb without any sensory changes.

With respect to ALS, physical examination usually shows weakness in either one limb or, as the disease becomes advanced, more diffusely. There is often dysphagia and slurring of speech, which are worrisome signs. There should be no significant sensory loss unless there is a coexisting lesion such as sensory polyneuropathy. Reflexes are typically brisk, consistent with the upper motor neuron involvement. However, at the same time, there is usually atrophy that suggests lower motor neuron loss.

When evaluating a patient with possible motor neuron disease, it is most fortunate for the patient if the electromyographer can identify some other (treatable) cause of the patient’s
symptoms. At times, myopathies, NMJ diseases, cervical spinal stenosis, and multifocal motor neuropathy with conduction block (124) can all mimic ALS. In contrast to ALS, these are treatable lesions without rapid progression to death.

Nerve conduction studies are first performed to evaluate for polyneuropathy or to look for multifocal motor neuropathy with conduction block. It is generally useful to study one motor and one sensory nerve in an upper and lower limb, with multiple sites of stimulation, in addition to F waves. In ALS, these are usually normal except for potentially reduced amplitudes of the CMAPs. If there is conduction block, then other diagnoses should be considered.

Needle EMG is usually performed according to the diagnostic requirements of the El Escorial criteria (125). These criteria divide the body into four regions: bulbar, cervical, thoracic, and lumbosacral. The electromyographer should study muscles from each of these regions to look for evidence of denervation, fasciculation potentials, or reinnervation. Presence of abnormalities in three or four of these regions is consistent with a diagnosis of ALS. When only one or two regions are involved, then the diagnosis is only possible or probable. There is recent debate about how these criteria should be used and/or modified (126) and whether or not fasciculation potentials must be recorded to be diagnostic.

Examples of muscles that can be studied in patients with possible ALS include

<table>
<thead>
<tr>
<th>Bulbar</th>
<th>Cervical</th>
<th>Thoracic</th>
<th>Lumbosacral</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tongue</td>
<td>Biceps</td>
<td>Paraspinal muscles</td>
<td>Vastus medialis</td>
</tr>
<tr>
<td>Facial muscles</td>
<td>Flexor carpi radialis</td>
<td>Tensor fascia lata</td>
<td></td>
</tr>
<tr>
<td></td>
<td>First dorsal interosseous</td>
<td>Medial gastrocnemius</td>
<td></td>
</tr>
</tbody>
</table>

Muscles such as these are selected to give a sampling of multiple roots and multiple peripheral nerves to diffusely screen for motor neuron loss. The thoracic paraspinal muscles are especially important to study because they are commonly affected in motor neuron disease (127). Moreover, when abnormal, they help to exclude the presence of combined cervical and lumbar spinal stenosis, which could result in abnormalities in both the upper and lower limbs. Likewise, the bulbar muscles are very important because, when abnormal, they help to exclude spine disease as the cause of the patient’s presentation.

When abnormalities are detected, the needle examination may need to be expanded beyond that summarized above. The primary finding on needle EMG is that of acute denervation, with positive sharp waves and fibrillations. Although fasciculation potentials are not required for a definitive diagnosis, their presence is much more suggestive of motor neuron disease than when they are absent. The presence of polyphasic, large-amplitude, long-duration MUAPs is suggestive of chronic reinnervation and should be seen in the limbs as well as evidence of denervation. Often, reduced recruitment with rapid firing of MUAPs is one of the first findings in ALS. Since progressive motor neuron loss can often be compensated by distal sprouting, there may not be prominent fibrillation potentials early in the disease.

Generally, when evaluating the patient with motor neuron disease, it is preferable not to discuss the findings or even the differential diagnosis with the patient unless he or she has already been discussing this with a referring physician. Getting into a discussion during the study often prompts questions that the electromyographer is not prepared to answer.

**REPORT WRITING**

After the electrodiagnostic examination has been completed, one then writes the electrodiagnostic report. This usually includes several elements such as

- Identifying information (name, medical record number, date of birth, name of referring physician, name of examining physician, and date of study).
- Brief history and physical (enough to support your differential diagnosis but not so long as to repeat extensive notes from the medical record).
- Electrodiagnostic data should be presented in tabular form with pertinent findings listed in the tables. Most recent electrodiagnostic instruments can automatically prepare these reports in a word processing document, but the electromyographer should use caution when employing this feature. Because the instruments will automatically place cursors on all wave forms, at times the automated reports will give cursor information when no response is in fact present. The electromyographer should review these tables to make sure that they are accurate and that responses are listed as absent when they are unobtainable.
- Summary of findings (how you interpret the findings—for example, “there is evidence of acute denervation in the C7 myotome” or “there is slowing of median motor and sensory conduction across the wrist”). The physician should not repeat what is in the data tables in this section.
- The impression or conclusion should be succinct and should clearly address the referring physician’s question. It should include whether the study was normal or abnormal, which diagnoses were ruled in, important diagnoses that were ruled out, pathophysiology, and prognosis when appropriate. This section should be able to stand alone, since some referring physicians will only read this part of the report. It should also give the side (right or left) and the pathophysiology (e.g., axon loss or demyelination). Sometimes, the electromyographer will not have a final diagnosis but will have a differential diagnosis, and this should be clearly stated in the impression (it is better to give a differential diagnosis than to simply guess what it might be).
The study will sometimes be normal. In these cases, the electrodiagnostic medical consultant will want to report "There is no electrodiagnostic evidence of..." Other times, the electromyographer will have a different diagnosis than the referring diagnosis. In these cases, one should comment on the referring diagnosis as well as the diagnosis found during the study.

Guidelines for what should go in the report are available on the Web site of the American Association of Electrodiagnostic and Neuromuscular Medicine (http://www.aanem.org/practiceissues/practiceguidelines/practiceguidelines.cfm).

In most cases, one should assume that the patient may read the medical record, and take care in the use of wording that might be offensive or might be interpreted as derogatory (e.g., obese or arguative).

SUMMARY

In summary, the electrodiagnostic evaluation is critical to the understanding of nerve and muscle disease, to the management of peripheral nervous system disease, and to assessing prognosis for a variety of lesions that the physiatrist will encounter.

ACKNOWLEDGMENTS

The author wishes to thank Paula Micklesen for the preparation of figures, Kevin Bleich for manuscript preparation, Dr. Leilei Wang for review of the manuscript, and Michelle Walter for her editing assistance.

REFERENCES


Human Walking

Walking is one of the most basic and ubiquitous forms of human motion. While studying gait, one can simply observe the world in motion. Airports and shopping malls are excellent places to watch walking. Observe the changes that occur based upon each person’s gait speed, height, or age. Compare those who are carrying loads with those who are unfettered. And observe yourself walking in different surface conditions and hills and at different speeds. Which pattern causes fatigue and shortness of breath, and which muscles become fatigued? Building up this personal knowledge base will help the physician to evaluate complaints of walking difficulty and gather information that might clarify the problem.

The basic unit of walking is the gait cycle, which is typically recorded from the time one foot strikes the ground until that episode recurs and starts the next, repeating cycle. During one gait cycle, the body traverses a distance of one stride (Table 5-1). Each stride is made up of one step by each foot, and these two steps are normally symmetrical in length. The frequency of stepping is known as the cadence (steps/min). The speed of walking is calculated as the cadence times the step length. In practice, walking speed is often computed by recording the time needed to traverse a measured distance (e.g., 50-ft walking speed) (1).

The gait cycle is divided into segments (Fig. 5-1) that serve specific functions, each for a limited time during the gait cycle. Each lower limb supports the body during its stance phase and then leaves the floor for its swing phase, during which it advances (or steps). Although the swing phase is the “motion” portion of the gait cycle, one must remember that the horizontal push for the forward motion is the responsibility of the contralateral limb, which is in stance. This forward push requires adequate friction between the foot and the walking surface so that hip extension torques are translated into forward propulsion of the pelvis and swing limb.

NORMAL HUMAN GAIT

The human gait pattern is normally fluid and shows continuous movement in the direction of travel. It has a natural repeating character, with the basic unit of the gait cycle, including one step with each foot. The basic function of walking involves each foot in turn either advancing forward as a step or supporting the body weight and balancing during the advancement of the contralateral lower limb. A period of double-limb stance (DLS) occurs between each step. During DLS, the weight is transferred from one foot to the other in a complex coordinated pattern known as weight acceptance (or loading) and weight release (preswing) for each of the respective limbs (see Fig. 5-1). Although related, the associated actions of arising from sitting and starting and stopping walking will not be addressed in this chapter.

During each gait cycle, a carefully timed pattern of acceleration and deceleration is produced by the muscles. This muscle activity must overcome gravity, as represented by the vertical component of the ground reaction force (GRF), and provide forward propulsion (Figs. 5-2 and 5-3). The entire body is seen to be involved in the gait cycle, with movement occurring simultaneously in each of the three planes of movement, as well as each of the three axes of rotation. These six degrees of freedom of movement need to be considered throughout the gait cycle. For convenience in understanding, the gait cycle is divided into natural subunits called phases that have distinct functions in the gait cycle and have both propulsion and control aspects.

The stance phase of gait begins with the period of weight acceptance. Weight acceptance, also called loading, is a decelerating portion of the gait cycle where the foot must stop after traveling at about 4 m/s during the end of swing phase. This sudden stop requires controlled braking involving simultaneous action of the ankle, knee, and hip. (At the same time, the left leg is involved in the equally complex task of liftoff to initiate its swing phase.) The ankle is in a few degrees of dorsiflexion at heel strike and then rapidly plantar flexes under the control of an eccentric (lengthening) contraction of the anterior tibialis muscle, as well as the lesser dorsiflexors, until the foot is flat on the ground. Simultaneous with this, the knee begins to flex under the eccentric control of the quadriceps (the loading response), and the trunk reaches its lowest point during the cycle. The hip, which was flexed approximately 40 degrees at heel strike, immediately begins to extend on the pelvis as the trunk smoothly continues forward. The forward momentum of the trunk is similarly controlled by contraction of the hip extensors, both the gluteus maximus and the long hamstrings, resulting in controlled hip extension as the trunk moves forward. The hip and pelvis rotate opposite to one another at the same time, the hip rotation of the stance limb being internal. The critical role that the stance limb’s hip extension and internal rotation has on forward propulsion was demonstrated by Sadeghi (2). He also reported that the ankle power is much less important to forward progression or propulsion.
As the left foot leaves the ground, the right hip continues to rotate internally as that limb enters the single-limb support portion of stance phase. The right ankle begins to passively dorsiflex as the tibia tilts forward and the pelvis lowers on the left in the familiar Trendelenburg tilt. The pelvis is also moving laterally to the right so that the center of mass (COM) of the body is aligned over the right foot to allow balance. During this period, the pelvis and COM of the body are rising. To achieve this, some of the kinetic energy of forward motion is converted into potential energy as the trunk rises against the force of gravity (3).

In the middle of single-limb stance (SLS) phase, the ankle reaches maximum dorsiflexion and the heel begins to rise from the floor. This heel rise signals a major shift in the right leg’s function in the gait cycle as it ends its decelerating role and begins to serve as the acceleration for its step phase. The foot now rocks over to the forefoot as the knee extends, and the pelvis (not the hip) maximally rotates externally and the right hip maximally extends in preparation for opposite heel strike. The left lower limb is fully stretched out for its heel strike so that the pelvis is supported and the drop of the COM of the body is minimized.

The weight-release phase on the right must now shift weight support to the left lower limb while accelerating the right leg for its forward propulsion or swing phase. During weight release, the ankle is actively plantar flexed by concentric action of the gastrocnemius, soleus, posterior tibialis, and lesser plantar flexors. These forces contribute primarily to the

---

**TABLE 5.1 Typical Temporal Gait Parameters for Comfortable Walking on Level Surfaces in Normal Adult Subjects**

<table>
<thead>
<tr>
<th>Temporal Gait Parameter</th>
<th>Average Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Velocity (m/s)</td>
<td>1.33</td>
</tr>
<tr>
<td>Cadence (steps/min)</td>
<td>113</td>
</tr>
<tr>
<td>Stride length (m)</td>
<td>1.41</td>
</tr>
<tr>
<td>Stance (% gait cycle)</td>
<td>62</td>
</tr>
<tr>
<td>Swing (% gait cycle)</td>
<td>38</td>
</tr>
<tr>
<td>Double support (% gait cycle)</td>
<td>24</td>
</tr>
</tbody>
</table>


---

**FIGURE 5-1.** Periods of the gait cycle. The gait cycle is separated into two distinct periods of stance and swing. Functional tasks include weight acceptance and single-limb support during stance and limb advancement during swing. The stance period of the gait cycle includes initial contact, loading response, midstance, terminal stance, and preswing. The swing period includes initial swing, midswing, and terminal swing. (Reprinted from SLACK Incorporated: Perry J. Gait Analysis: Normal and Pathological Function. Thorofare, NJ: Slack; 1992, with permission.)

**FIGURE 5-2.** Quiet standing. The GRF, represented by the solid line with an arrow, is located anterior to the knee and ankle and posterior to the hip. The soleus muscle is active to stabilize the lower limb. (Courtesy of D. Casey Kerrigan, MD, with permission.)
vertical components of the GRF. At the same time, right hip flexion is being produced by the iliacus, psoas, and tensor fascia lata muscles, and the left hip internal rotators are causing forward rotation of the right pelvis. These left hip muscles contribute much to the horizontal forces or forward propulsion of the right lower limb. As the toe finally leaves the ground as a result of this combined push and pull on the right lower limb, the swing phase is initiated, causing a step to occur.

During the beginning of swing phase, the right leg continues to accelerate as hip flexion, knee flexion, and ankle dorsiflexion combine to cause the toe to pass over the ground cleanly. Typically, the toe reaches a minimum height of less than 2.5 cm. at the middle of the swing phase. This minimal elevation is energy conserving by reducing the work done to elevate body parts against gravity. This closeness to the ground can be a safety problem, causing stumbles on uneven ground, but the gains of energy efficiency caused by reducing the step height are important enough for this safety risk. The energy-conserving nature of the gait cycle, in part attributable to these mechanics and its highly repetitive and symmetrical nature, will be discussed in more detail, but most of us have experienced the rapid fatigue caused by high stepping (“steppage gait”) when traversing deep snow or the lack of repetitive steps when walking on very uneven ground.

The second half of the swing phase returns the right leg into a role of decelerating as the forward motion is slowed in preparation for heel strike. The ankle is held in dorsiflexion while the hip flexes and the knee extends. This combination, along with the forward rotation of the right pelvis and external rotation of the right hip, results in maximal length of the step. Any injury or dysfunction of a joint or muscle-tendon unit that reduces the step length will have a major impact on the efficiency of walking. At the end of swing, the hip extensors serve to brake the forward flexion of the hip; and, at faster gait speeds, the hamstrings slow and control the knee extension. The speed of the leg and foot must be controlled to prevent slipping at heel strike.

**ENERGY CONSERVATION AND GAIT**

Energy is consumed in three different categories during walking (4). First, there is the work of moving the body’s mass through the required distance in a period of time. Second, there is the work done to accomplish the up-and-down motion of the trunk for each step as the pelvis rises to a position above a single supporting leg during midstance and is lowered during double support when it lies between the two lower limbs supporting it at opposite angles. Third, energy is being consumed by the body for general, or basal, metabolism. Because work requires energy and faster walking involves more work, there is an energy cost. However, since the constant basal metabolic rate continues, as well as some muscular energy consumption for standing, the total energy consumed per unit distance traveled is seen to decrease initially as speed increases from a very slow walking speed. As an optimal speed is reached that is traditionally called comfortable walking, the combined metabolic rate is most efficient for traveling over the ground. The higher energy requirements of faster walking and running require greater work output by the muscles.
The vertical displacement of the body is minimized by a number of factors that are known as the determinants of gait (5). These six determinants operate independently but simultaneously to produce a smooth sinusoidal vertical and horizontal path, which has one vertical peak and trough for each step. There is also one lateral sinusoid, or curve, for each stride as the body moves toward the supporting foot during each cycle. To illustrate the effects of the determinants of gait, they are removed and replaced in models of the gait cycle as shown in Figures 5-4 to 5-8. This "compass gait" analysis was restudied by Della Croce et al. (6) with the inclusion of data recorded in a contemporary gait lab, largely reaffirming the model.

The first determinant of gait is the rotation of the pelvis (Figs. 5-4 and 5-5). During each step, the pelvis rotates forward on the side of the swinging limb. The axis of this rotation...
is the hip joint of the stance leg, which undergoes internal rotation. As the pelvis forms a bridge between the two hips, it reduces the angle of intersection of the thighs to reduce the vertical descent of the trunk.

The pelvic Trendelenburg motion, or pelvic list, is the second determinant (Fig. 5-6). The pelvis drops a few degrees so that the hip of the leg in swing phase is lower than the hip of the stance limb. This reduces the vertical rise of the COM of the trunk and reduces the work of lifting this mass. The cost of this pelvic list is to reduce the space for toe clearance, but this is an affordable cost. Pelvic list, as well as rotation, has been shown to decrease at slower speeds (7). This fact, and the data that reveal that these motions are less important modifiers of the vertical movement of the COM (6), suggests that these movements are important to the control of the momentum during forward propulsion.

Knee flexion in stance phase is an important determinant of gait for two reasons. First, it provides a shock-absorbing mechanism at the beginning of the stance phase. The reduction of the shock of foot impact on the floor helps to maintain momentum and thereby reduces energy loss of stopping and restarting the gait cycle. Also, the knee flexion in stance reduces the height of the hip joint in midstance. This additional height reduction prevents energy loss from lifting the body but at the cost of quadriceps muscle work.

Lateral displacement of the pelvis also occurs during each step (Fig. 5-7). The pelvis and trunk must move to the stance side to balance the COM of the trunk above the stance foot. This also aligns the tibia into the vertical position during stance. This determinant of gait is a net loss of energy since it causes upward movement of the body, but it is necessary for balance in bipedal gait.

The trunk and shoulder rotate during normal gait in a direction opposite to the pelvic rotation. This 180-degree phase shift of total trunk movement balances the angular acceleration so that balance and forward momentum are maintained. Smooth, coordinated movement here is an energy advantage.

The obliquity of the subtalar joint provides a unique relationship between the motion of the foot and that of the shank. Dorsiflexion of the foot causes lateral movement of the forefoot and vice versa. During stance, the passive dorsiflexion of the ankle, therefore, causes internal rotation of the tibia to partially match that similar movement in the hip. This combined rotation is then reversed during the end of stance or weight release.

Sagittal plane foot and ankle movement is referred to as the three “rockers.” The three components are at heel strike, during foot flat, and during toe-off. Each of these has work- and energy-saving factors (Fig. 5-8). For the first rocker, the dorsiflexion of the foot causes the heel to stick out and produce a net lengthening of leg length to maximize the length of the step. This extra length is quickly lost during weight acceptance, when it is no longer needed. Also during this time, the resisting ankle dorsiflexion muscles provide a shock-absorbing descent of the forefoot. During the second rocker, dorsiflexion of the foot occurs during midstance. This serves to reduce the length of the leg until the pelvis passes in front of the ankle. As the heel rises after midstance, the third rocker occurs. The elevation of the heel increases the leg length during push-off and so limits the amount of drop that is experienced by the pelvis. The rockers of the heel, midfoot and forefoot are therefore useful in minimizing the vertical work of the body movement. The second utility of the rockers...
A complete understanding of gait requires knowledge of the kinetics of movement in addition to the kinematics of motion discussed in the preceding sections. Kinetics is the science of forces acting on bodies to cause motion. Gait kinetics explains the causes of the motions of gait.

The basic principles of kinetics are Newton's three laws of motion. The first is that a body (of mass $m$) will change velocity (accelerate ($a$) or decelerate) only if a force is applied to it. Also, this change in velocity ($\Delta v$) is proportional to the force ($F = ma$ and $a = \Delta v$). Newton’s third law, the law of action and reaction, is very important for the study of gait and other aspects of biomechanics. This law relates the forces interacting between the foot and the floor as always being equal and opposite. One can therefore measure the forces of ground reaction on the foot with a force plate and begin to understand the net forces acting on the lower limb and on the body as a whole (Fig. 5-9).

The GRF is represented by three directions perpendicular to each other: horizontal (fore/aft), side to side, and vertical. One can also measure the twisting or rotational load on the force plate as the rotating limb is constrained from moving by friction. The force plate also allows the direct calculation of the center of pressure on the foot. During stance phase, the center of pressure starts at the medial heel. As foot flat occurs and the limb progresses to single stance, the center of pressure typically moves laterally as it progresses forward. Then, after heel rise and into weight release, the center is in the forefoot, progressing to the medial side again.

From the center of pressure, the net vector of the three GRFs can be located at each instant during gait (see Fig. 5-3). Through the process known as inverse dynamics, these forces can be calculated sequentially from the ground and foot up to the trunk. Changes in vertical position of the trunk COM result in changes in the potential energy, and forces
recorded and amplified so that the muscle activity can be correlated with the kinematics and kinetics of gait. In carefully controlled experiments of isometric contraction, a linear relationship exists between the EMG signal and the muscle force. However, this relationship is greatly disrupted by the movement of the muscles during gait, so EMG does not give accurate information about muscle forces (8). For example, EMG analysis of the rectus femoris’ role as a hip flexor in early swing phase demonstrates that higher amplitude EMG signals are produced at faster walking speeds (9). The muscle must exert greater force to flex the hip more rapidly, but the relationship is nonlinear and complex, and one is still unable to make quantitative conclusions about muscle force from EMG signals alone.

Multichannel amplifiers are used in gait laboratories to record from several muscles simultaneously to the recording of kinematic motion (Fig. 5-11). The electrodes used for this recording may be either surface electrodes or flexible intramuscular (IM) wire electrodes (8). Surface electrodes are noninvasive and record a larger volume of the target muscle. They are unable to record signals from deep muscles and may record unwanted signals from muscles that are adjacent to the target muscle. Wire electrodes can be located to record more precisely from any muscle, including those that are deep. However, in addition to being more painful, the wire electrodes may dislodge or break during activity. EMG also allows the measurement of muscle fatigue during sustained activity. As lactic acid accumulates in a muscle with exercise, the membrane propagation of the action potential (AP) is slowed. This slowing of velocity is seen as a reduction in the frequency of the recorded EMG signals (8).

**DEVELOPMENT OF GAIT**

Normal adults perform the task of walking without significant active thought or effort. It is a process that is learned and eventually mastered in childhood. Understanding the development of gait is critical to any physician performing neuromuscular examinations of children. To appreciate subtle normal variations or pathologic findings, a physician must have a thorough knowledge of normal development, understand its relationship to gait, and use a systematic approach (10).

**Early Infant Development**

The first few years of life are characterized by the most rapid growth in the human life span (11). During this period of maturation of the central and peripheral nervous systems, the development of gait emerges according to a typical pattern of motor development: rostral to caudal, proximal to distal, and mass movement to specific action (11,12). Although there is variation in the development of each child, some general statements concerning milestones of development are appropriate. Head control is a prerequisite to ambulation that begins to improve by the age of 2 months (12,13). Further improvement in head control is seen at the age of 3 months, along with an integration of primitive reflexes...
FIGURE 5-10. Kinetics and kinematics at the hip, knee, and ankle. Sagittal joint motion, moments, and power are shown.
(Courtesy of D. Casey Kerrigan, MD, with permission.)

FIGURE 5-11. General muscle group activity as a percentage of the gait cycle in normals.
(Courtesy of D. Casey Kerrigan, MD, with permission.)

<table>
<thead>
<tr>
<th>Muscle Groups</th>
<th>0</th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>50</th>
<th>60</th>
<th>70</th>
<th>80</th>
<th>90</th>
<th>100</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tibialis Anterior</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>100</td>
</tr>
<tr>
<td>Vasti, Long Hamstrings, and Hip Extensors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>100</td>
</tr>
<tr>
<td>Rectus Femoris</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>100</td>
</tr>
<tr>
<td>Ankle Plantarflexors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>100</td>
</tr>
<tr>
<td>Hip Flexors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>100</td>
</tr>
</tbody>
</table>

□ Indicates that a muscle or muscle groups are active
with occasional bobbing during supported sitting (11). From the ages of 4 to 6 months, children learn prone prop-
ing followed by rolling. Rolling is first achieved from prone to supine, followed by supine to prone, and by this time, independent sitting is usually achieved with arm support. By 9 months, a significant increase in mobility ensues with the advent of belly crawling and standing up by pushing onto hands and knees (11).

Achievements of these skills as well as sitting, creeping, and pulling to a stand are prerequisites that a child must master before supported and independent walking. Balance and support, two major skills required for independent walking (14), are practiced as a child progresses toward supported walking.

**Supported Walking**

Around 8 to 10 months of age, children begin to walk by holding onto objects. For example, a child may grab a crib surface or table while attempting to maneuver across a room. On the average, supported walking or cruising lasts 2 months, with significant variability (15). This stage is characterized by the use of arms for support and is a transitional phase that precedes independent walking. In comparison to independent walking, supported walking is characterized by a longer gait cycle time, longer stance phase, increased variability in successive strides, and decreased walking speed. The average walking speed in supported walking is noted to be 39 cm/s, with stride length of 44 cm and 65% of gait cycle being spent in the stance phase (16).

When compared with normal adults, children reveal greater angles of hip flexion during walking. The hip exten-
sion pattern shows increased variability, poor smoothness, and decreased total amplitude. There are also striking differences in knee and ankle patterns between normal adults and children during supported walking (16). Knee pattern and amplitude of flexion differences are most impressive during the stance phase. At initial floor contact, Statham and Murray (16) found average knee flexion of children in supported walking (stance) to be 21 degrees, which is in sharp contrast to that of adults where knee flexion was 1 degree. Knee flexion angles during swing were similar to that of adults.

In supported walking, there is no clear heel strike, as a child may make initial floor contact with the heel, flat foot, or forefoot (i.e., toe walking) (16–18). This is in contrast to normal adults, where heel strike is very consistent and is followed by ankle plantar flexion rather than ankle dorsiflexion. The lack of full knee extension and ankle dorsiflexion at the end-of-swing phase have been cited as possible causes for variability at foot strike (15). Initial floor contact is followed by ankle dorsiflexion (modified loading response) in supported walking regardless of which part of the foot makes initial contact. The maximums of ankle dorsiflexion and plantar flexion angles, the degree of ankle dorsiflexion at initial floor contact, and the maximum ankle dorsiflexion during swing phase are also greater in supported walking than they are in adults (16).

In summary, prominent features of supported walking in children that are in contrast to those of adults include increased hip flexion, decreased knee extension, increased ankle dorsiflexion, decreased walking speed, increased variability, and decreased smoothness.

**Independent Walking**

Following supported walking, children progress to independent walking with a premature gait pattern and, during this phase, are referred to as “toddlers” (15). This is a transitional period between supported walking and a mature gait pattern. Collection of data using quantitative gait analysis in this age group often proves to be very difficult, as these children do not follow simple instructions and display significant variability from one gait cycle to the next. Inconsistent gait pattern, wide-based support, increased hip flexion, and lack of reciprocal arm swing are some of the hallmarks of gait in toddlers. In general, the development of walking skills is completed by the age of 5 years (19,20).

In the progression from supported walking to independent walking, there is a trend toward decreasing excessive hip flexion and greater extension at initial floor contact; however, these movements are still not identical to those of adults (16). On initial floor contact, the knee in independent walkers is held in less flexion, and there is the emergence of an adultlike knee flexion during landing. This knee flexion is established in most children by the age of 2 years (21). Consistent initial floor contact with the heel is observed by 22.5 weeks after initiating walking and at the average age of 18.5 months. A variable amount of toe walking continues during this age range. Subsequent to this, there is the development of a mature foot and knee mechanism during stance phase, which develops about 27 weeks after the initiation of independent walking, at a mean age of 19.5 months (17). There is a significant decrease in the excess ankle dorsiflexion movement at initial floor contact, compared with that seen in supported walking. In mature adult gait, the swing phase is marked by hip and knee flexion initially, followed by continued hip flexion and knee extension during mid and terminal swing, whereas the ankle begins dorsiflexion during midswing. These same swing phase motions are present in most toddlers, with slight modifications until the normal adult motions are achieved, typically by the age of 3 years (15).

Sutherland et al. (19,21) proposed five important determinants of mature gait: duration of SLS, walking velocity, cadence, step length, and the ratio of pelvic span to ankle spread. The gait of a toddler demonstrates significant changes in each of these determinants as he or she approaches gait maturity. The duration of SLS increases steadily from 32% of the gait cycle at 1 year to 38% at age 7 (19,21). The most rapid changes were observed in children before 2½ years of age.

Walking speed increases as a child progresses from supported to independent walking; however, children are inconsistent in their ability to manage their walking speed (20). Average speed in independent walkers is found to be 64 cm/s, in contrast to 39 cm/s for supported walkers (16). Furthermore, increased cadence seems to be the main reason for the noticeable increase in walking velocity in the toddler. Sutherland (21) reported that walking velocity increases with age in a linear manner from 1 to 3 years, at a rate of 11 cm/s/y, and from 4 to 7 years, the rate of change diminishes to 4.5 cm/s/y with a continued
linear relationship. As independent walking matures, there is a
decrease in cadence with an increase in stride length with a net
result of faster walking speed (15,21). The primary reduction
in cadence occurs between the ages of 1 and 2 years, with a
gradual decrease thereafter. The cadence at 7 years is noted to be
26% greater than in a normal adult (21). As one would expect,
the stride length increases with height, and this leads to faster
walking for taller children (and adults). Todd has developed a
mathematical equation that defines this nonlinear relationship
for children during development without regard to age (22).
Sutherland et al. (19,21) defined the ratio of pelvic span to
ankle spread, their fifth determinant of mature gait, as the ratio
of body width at the anterior superior iliac spine and distance
between ankle centers during double support. They found that
the ratio increased in a linear fashion from ages 1 to 3 years
and then remained constant. Burnett and Johnson (17) found
that foot placement moves within the lateral dimensions of
the trunk by 17.5 weeks after initiating independent walking
and continues to narrow for several more weeks. Reciprocal
arm swing, another feature seen in mature gait, is universally
seen in children by the age of 4 years and rarely before the first
birthday (21). Burnett and Johnson (17) reported the develop-
ment of reciprocating arm swing by 22 weeks after initiating
independent walking.
The trend of muscle timing during the toddler years is to
shorten the duration of muscle action during the gait cycle and
establish normal phasing by reducing unnecessary muscle activ-
ity (15). Because of convenience as well as other obvious issues,
EMG data in toddlers have been collected primarily from sur-
face electrodes rather than from IM wires. Quadriceps and
hamstring activities are prolonged in a toddler but approach
normal timing by 2 years (21). No change was seen in the
gluteus medius EMG activity with increasing age; however,
there was shortening of the time that the gluteus maximus
was active during stance phase. Medial and lateral hamstring
activities were primarily prolonged, during stance phase, in the
immature gait of 12- to 18-month children, with mature pat-
terns emerging around the age of 2 years. Similar findings were
also observed for the tibialis anterior and gastrocnemius-soleus
muscle complex when referring to the late swing phase and
premature stance phase activities seen in these muscles (21).
Normal gait development is a complex process. Pathologic
gait is even more complex, and walking is considered to be a sen-
tive measure of neuromuscular development (23). By under-
standing gait development and utilizing both careful observation
and experience, the physiatrist can detect pathologic gait in this
changing background. An understanding of normal gait devel-
opment, including age-related changes, is crucial for any physi-
cian performing a pediatric neuromuscular examination.

ABNORMAL GAIT
Several examples of abnormal gait have been selected to help
reinforce important basic biomechanical concepts of human
motion. When presented with a patient who has an abnormal
gait pattern, it must be determined whether the abnormal
movement patterns have (a) potentially detrimental functional
consequences, such as increased fall risk, excessive energy expe-
diture, or adverse joint stresses that may predispose to degener-
ative changes; (b) only cosmetic significance; or (c) advantages
in adapting and are functionally beneficial. Often deciding
that a gait pattern is abnormal is apparent from the patient’s
history and observational gait analysis. Quantitative kinematic
and kinetic data from gait analysis are used to differentiate
the primary gait deviations from those that are secondary (or
compensatory) (24,25).

The general appraisal of a patient with an abnormal gait
involves assessing the symmetry and smoothness of movements,
including trunk movement, arm swing, stride length, width of
gait base, balance, and degree of effort. Useful additional infor-
mation will result from having the patient increase and decrease
his or her gait speed, observing differences with and without
use of any hand-held gait aid or orthosis (if going without such
devices is feasible), and if available, observing the patient walk-
ing up an incline and down a declined surface. Then, a more
systematic and methodical evaluation should be performed,
focusing on specific segments and joints, from head to toe or
vice versa. A patient’s gait should be observed from the front,
back, and both sides, and the examiner should note the portion
of the gait cycle during which a particular deviation occurs.

When interventions are planned to improve a patient’s
gait, specific goals need to be determined from the outset,
including the patient’s ultimate ambulatory goal (full com-
unity, limited community, or household level ambulation).
When muscular weakness is a contributor to an abnormal gait,
strengthening exercises for the agonist and stabilizing muscles
may be appropriate. However, with a progressive disease pro-
cess, or following the destruction of most motor units of a
particular muscle group, compensating for the weakness with
appropriate orthoses and hand-held gait aids should be con-
sidered along with a manual wheelchair or powered mobility
device. If spasticity is present, spasticity-reducing interven-
tions include physical therapy; oral, injectable, and intrathe-
cal pharmacologic agents; and tendon surgery. It is important
to remember that spasticity may actually have limb-stabilizing
effects that can be beneficial from a functional perspective
(26). For patients who wear a lower limb prosthesis or brace,
gait abnormalities with a prosthetic or orthotic cause should be
addressed through adjusting or replacing prosthetic or orthotic
components, whereas patient-related causes may require train-
ing the person in optimal prosthetic or orthotic ambulation,
skin care, muscle stretching and strengthening, balancing, and
endurance and aerobic exercises.

Transfemoral (Above-Knee) Amputation
The first example will consider a patient who has undergone
a transfemoral (TF) amputation. More detailed discussions of
TF amputee gait patterns are available elsewhere (27–31), and
Chapter 74 provides information about potential gait deviations
in the TF amputee. For the purposes of this example, we will
assume a mid thigh level of amputation, no significant hip
flexion contracture, use of an ischial containment socket, and optimal prosthetic suspension. The shorter the residual limb is, the shorter the lever available to control the prosthesis, resulting in the hip abductors on the prosthetic side becoming less effective in stabilizing the pelvis (32). The ischial containment or narrow medial-lateral socket is usually preferred to the older quadrilateral socket design, because of advantages in both comfort and biomechanics (33). Finally, suboptimal suspension results in a functionally long prosthetic limb, with resultant difficulty clearing the foot during midswing phase unless compensatory motions (e.g., hip hiking or an abducted gait on the prosthetic side, or vaulting on the intact side) are performed.

Prosthetic knee stability is of paramount importance to the TF amputee, whose remaining quadriceps can no longer serve the function of controlling knee flexion during loading. During the initial, weight acceptance portion of stance phase, the prosthetic knee unit will be more stable when the GRF line passes anterior to the knee. For a TF amputee with a longer, stable residual limb and normal strength, the prosthetist may elect to place the axis of rotation of the knee unit anterior to the trochanter-knee-ankle (TKA) line—the line drawn during static alignment that passes through the greater trochanter, knee, and ankle joint centers in the sagittal plane—to provide “voluntary control” over prosthetic knee stability (32). Although this situation requires greater voluntary activation of the hip extensors beginning at the moment of heel strike, which via closed-chain kinetics affects an extension moment at the knee, the advantage is easier knee flexion during terminal stance and preswing, since the GRF line is more readily positioned posterior to the knee joint axis (29,30). On the other hand, for a TF amputee with a weakened residual limb, the prosthetic knee joint axis would be placed posterior to the TKA line to afford greater knee stability but with the disadvantage of greater difficulty achieving knee flexion during late stance. Current prosthetic technology provides greater knee stability during early stance and more readily achieved knee flexion during late stance with mechanical polycentric knees, simple fluid or pneumatic-controlled knees, and the increasingly popular microprocessor-controlled knee units (34).

With regard to the prosthetic foot and ankle, the more rapidly the prosthetic foot advances from heel strike to foot flat, the more rapidly the GRF line progresses to a position anterior to the knee joint axis, with a resultant extension moment at the knee. Any restriction in ankle plantar flexion, such as may result from an excessively stiff plantar flexion bumper in a single-axis prosthetic foot, would serve to destabilize the prosthetic knee during early stance. Given that there is an ankle dorsiflexion moment from midstance to terminal stance, unchecked dorsiflexion, such as can result from a worn dorsiflexion bumper in a single-axis prosthetic foot, would predispose to late-stance phase buckling of the prosthetic knee because of unrestricted forward motion of the prosthetic shank (32).

Foot Drop Gait
A “foot drop” condition results from weakness of the anterior compartment (dorsiflexor) muscles of the leg, ankle plantar flexor spasticity, or an ankle plantar flexion contracture. With the exception of the last condition, an ankle-foot orthosis (AFO) is typically used to compensate for foot drop (35). Without an AFO, foot drop results in gait deviations, resulting in increased energy expenditure from such compensatory motions as hip hiking or excessive hip and knee flexion (i.e., steppage gait) to achieve foot clearance during swing phase. Foot slap may occur immediately following heel strike in mild weakness, or, if severe, the forefoot makes initial contact in stance phase (36). Spastic plantar flexors usually necessitate a rigid, solid (nonarticulated, to reduce activation of the plantar flexors’ stretch reflex) AFO designed to maintain a 90-degree (neutral) position of ankle dorsiflexion (if achievable). In the absence of significant spasticity, however, spring-assisted dorsiflexion is often used. This type of component can be incorporated into both plastic and double metal upright AFO designs and allows partial ankle motion—from neutral to up to full ankle dorsiflexion—with less adverse effect during stance phase compared with a solid plastic AFO or fixed-ankle metal AFO. Either an AFO with assisted dorsiflexion or an AFO that allows no ankle motion will usually prevent ankle plantar flexion beyond neutral (via a 90-degree “plantar flexion stop”), resulting in little difference between the two from initial contact through loading response. The inability of the ankle in an AFO to plantar flex beyond neutral keeps the GRF line posterior to the knee, resulting in a knee flexion moment from initial contact through midstance. However, from midstance to terminal stance, a solid AFO does not allow the tibia to move progressively forward via dorsiflexion, which interferes with the second of the “three rockers” described earlier, and instead necessitates early heel rise from knee flexion. Neither AFO design allows ankle plantar flexion to occur as would normally take place from terminal stance through preswing. Maintaining foot clearance in a narrow range is critical during initial and midswing phases, and both of these AFO designs prevent toe drag by restricting plantar flexion.

Weak Calf Gait
In contrast to ankle dorsiflexor weakness, which affects the entire swing phase and the first half of stance phase, weakness of the ankle plantar flexors results in deviations involving the second half of stance phase. In normal gait, the ankle plantar flexors contract eccentrically to control the rate at which the tibia advances forward over the supporting foot. Weakness of this muscle group, as may result from such conditions as Achilles tendon injury, tibial nerve injury, S1 radiculopathy, or lower lumbar myelomeningocele, results in excessive and untimely forward progression of the tibia during the midstance to late-stance phase (2). This excessive passive ankle dorsiflexion during terminal stance prevents normal heel rise, which functionally shortens the stance limb and reduces the effect of the toe (or third) rocker. This leads to premature contralateral initial foot contact that decreases stride length (36). It is shortened step, not deficient “push-off,” that impairs forward progression in the presence of calf weakness (2). The failure of vertical support from the calf also causes ipsilateral
Gait deviations following stroke may manifest in a variety of ways, depending on the severity, location, and extent of damaged brain tissue; the amount of time that has passed since the stroke; and the rehabilitative interventions that have been implemented. For purposes of illustration, this discussion will relate to typical gait abnormalities that may follow the common middle cerebral artery distribution stroke, with resultant unilateral weakness and spasticity. Typical abnormal movement patterns include reduced knee flexion during swing and stance phases (spastic paretic stiff-legged gait), knee hyperextension (dynamic recurvatum) during stance, and excessive ankle plantar flexion (equinus) during swing and/or stance (38). Each of these has the potential detrimental effect of raising the energy requirement for walking.

Reduced knee flexion during swing phase may necessitate such compensatory maneuvers as hip circumduction, hip hiking, and contralateral vaulting with excessive elevation of the pelvis to avoid toe drag. Insufficient knee flexion during stance results in a relatively elevated position of the body’s COM. Dynamic knee recurvatum may predispose the knee to injury from overstretch of ligament and posterior capsular structures as the (external) GRF and the (internal) muscle force combine to create a knee extension moment (39).

Ankle equinus during swing phase may be due to ankle dorsiflexor weakness, plantar-flexor spasticity, or ankle plantar flexion contracture. During stance phase, the latter two causes may contribute to the development of dynamic knee recurvatum by preventing the progressive ankle dorsiflexion that normally begins during loading response and continues through terminal stance (36). All three causes of ankle equinus, depending on severity, may necessitate one or more of the compensatory motions during the swing phase (mentioned above) that can reduce toe drag.

Another feature of gait following stroke is reduced speed, with marked asymmetry of stepping (40). Many authors have reported a reduced percentage of the gait cycle in single-limb stance (SLS) on the affected side (40,41). However, the most striking deviation from normal gait is the prolonged period of weight release for the hemiparetic limb during DLS (38). Poor weight release is associated with both a low angular velocity of hip flexion and a delay in initiating flexion, which is seen as a long period of double-limb stance. We also observed that SLS on the affected side is the same as SLS in normal gait when measured in absolute time (Fig. 5-12) (38). SLS lasts just long enough to allow advancement of the sound limb, which is a relatively fixed period of 0.34 to 0.38 seconds. As gait slows, this fixed time may become a very small portion of the gait cycle time. The poor angular velocity of the weak hip contributes to poor knee flexion during swing, which is passive at normal gait speed. This compounds any knee extensor muscle contraction caused by spasticity (38,42).

Diplegic Cerebral Palsy

Of the various types of cerebral palsy (CP), the spastic form is the most prevalent. Sutherland and Davids classified the gait abnormalities at the knee in patients with CP into four types: jump, crouch, recurvatum, and stiff (43). Some of the features of a spastic gait with knee recurvatum, or a "stiff-legged” knee gait, have been discussed under post-stroke gait. In jump gait, excessive knee flexion is present during early stance phase but with normal or near-normal knee extension during the midstance and late-stance phases. Crouch gait typifies the spastic diplegic form of CP and is...
characterized by excessive knee flexion throughout the stance phase and variable knee alignment during the swing phase (44). In addition, the hips are often adducted and internally rotated (scissor gait), and there may be equinus and forefoot abduction (44). Although one might assume that the spasticity evident on static examination is the main culprit responsible for the abnormal movement patterns, gait analysis has revealed abnormal dynamic muscle activation patterns, including excessive co-contraction of agonists and antagonists, with resultant increased joint stiffness and weakening of agonist force production as playing a more central role (45). Hamstring spasticity has been proposed as being primarily responsible for the crouch position, but gait analysis has additionally identified hip or knee flexion contractures, ankle plantar flexion contractures, and ankle plantar-flexor weakness as being contributory (36,44,45).

Changes in Spastic Gaits Following Spasticity-Reducing Interventions

The mere presence of spasticity does not necessitate the initiation of spasticity-reducing interventions to improve a spastic gait pattern. As noted for CP, disordered muscle activation patterns may be partially or primarily responsible for abnormal movement patterns. And as noted for gait following stroke, delay in the initiation and speed of hip flexion during the initial swing phase, rather than quadriceps spasticity, may be responsible for reduced knee flexion during swing phase. Even if quadriceps muscle spasticity is responsible for stiff-legged gait, only one or two heads of the quadriceps may be involved. Although intense and prolonged activity in the rectus femoris at the junction of preswing and initial swing phases is often identified, similar abnormal firing of the vastus intermedius, with or without co-contraction of hamstring muscles, may also be occurring. Differentiating these muscle contraction patterns necessitates dynamic, multichannel EMG recording during quantitative gait analysis (36,46).

Surgical procedures such as rectus femoris release have been used to eliminate the detrimental effect of abnormal rectus femoris activity, with resultant enhanced initial swing phase knee flexion in CP patients and hemiplegic stroke patients (47–53). Motor nerve block of the rectus femoris has been shown to improve maximal swing phase knee flexion, the slope of the knee motion curve at toe-off, and gait speed in patients with spastic stiff-legged gait who also had sufficient hip flexor strength and no abnormal EMG activity of the vastus muscles during the initial swing phase (54). Alcohol neurolysis of the sciatic nerve to reduce hamstring spasticity in hemiplegic stroke patients has produced beneficial effects lasting 6 months (55). Selective dorsal rhizotomy in CP patients has been shown to normalize co-contraction during knee extension but was without benefit to co-contraction during ankle plantar flexion the majority of the time (56).

Electrical stimulation of nerves and muscles has a long history in attempts to improve the functional walking in persons with upper motor neuron injury and spastic paralysis (57–59). Liberson envisioned an electrical replacement for the ankle-foot

FIGURE 5-12. Mean duration (absolute time) of the phases of the gait cycle for each performance level in adult hemiplegic walking. The percentage of the gait cycle that each phase represents is printed below each bar. The four bars represent three different functional levels of hemiplegia after stroke in comparison to normal gait. The bars are arranged to allow comparison of the duration of the SLS phase for the hemiplegic side. HFS—foot strike on the hemiplegic side; OTO—toe-off on the contralateral (unaffected) side; OFS—foot strike on the contralateral side; HTO—toe-off on the hemiplegic side. (Reprinted from de Quervain IA, Simon SR, Luergans S, et al. Gait pattern in the early recovery period after stroke. J Bone Joint Surg. 1996;78A:1506–1514, with permission.)
orthoses (59). Later work expanded this concept into complex systems that use multiple site of stimulation with implanted wires to produce more complex movements (60,61). Some of these devices functioned with added stabilizing support of mechanical orthotics, and many were successful in the laboratory setting (60,61). Further development of this technology has redirected attention on Liberson’s concept of a device that is simple and safe enough for home use as an exercise and training device (62,63). In persons with hemiplegia after stroke and incomplete spinal cord injury (SCI), these electrical stimulators activate the fibular nerve during swing phase to produce walking patterns that are faster, farther and more efficient, and with improved toe clearance (60,61). The improved toe clearance comes with the finding that fibular nerve stimulation improves knee flexion as well as ankle dorsiflexion (57,60). In addition, the evidence makes it clear that this method of treatment has a training effect and results in sustained improvement in walking when the stimulator is inactive (60,63,64). Researchers suspect that this training effect represents central nervous system reorganization rather than just peripheral improvements in muscle tone, range of motion, and muscle strength (64).

Another recent advance in the management of gait disorders associated with spasticity and dystonia has been the application of botulinum toxin (BTX) types A and B. BTX is available in the United States and many other countries for therapeutic use. When compared to other muscle tone–reducing options, these toxins have the advantage of producing selective, graded weakness in individual muscles with a predictable recovery over several weeks. Gait analysis may identify a muscle or muscles that reduce the efficiency of gait (65), and BTX can reduce the contractile force of the selected muscle. The analysis might be a simple, clinical observation of gait as part of a neurological exam, or it might be a sophisticated, multichannel EMG recording of muscle activity during the gait cycle. At the time of injection, needle electromyography may confirm motor unit activity at rest, or electrical muscle stimulation may be used to direct the injection into the muscle identified by the exam during walking.

An example of this type of treatment includes treatment of ankle inversion by BTX injection of the tibialis posterior muscle (in some cases with additional dosing of the tibialis anterior) (26). During the ensuing weeks, the weakened muscles will require orthotic support so that stable walking can be practiced (66); however, the treatment can reduce pressure at the lateral malleolus and other areas so that pain and pressure ulcers are improved. The result can include improved walking speed and step length (67,68). Adductor muscle scissoring of the thighs can also be improved with BTX treatment. A risk to be considered is the potential loss of extensor tone in the lower limbs upon which the patient has been relying for standing and walking. While the rehab team may be able to teach a compensation for this weakness, in many cases, the patient will need to wait for the BTX to wear off so that the useful tone returns to the legs. Dosing information for BTXs A and B are available for both adults and children (69,70). Typically, injections are repeated at 3 to 4 month intervals if the offending muscle tone reoccurs (71). Side effects, in addition to muscle weakness or fatigue, include nausea or “flu-like symptoms” and symptoms associated with anticholinergics such as dry mouth. BTX treatment has eliminated the use of tendon transfer surgery for many patients (70). It needs to be noted, however, that BTXs have not yet been approved for this use by the U.S. Food and Drug Administration and that serious complications, including some deaths, have been reported. However, numerous authorities have produced practice guidelines for this therapy (69–72) and most U.S. health care plans pay for the treatment (73).

Treatment with BTX is an adjunct to other physical and pharmacologic treatment of specific conditions resulting in disordered muscle tone, covered in more detail in other chapters of this text. Oral and intrathecal baclofen therapies have been shown to improve walking velocity in appropriately selected patients with upper motor neuron disorders, for example (67,71). However, it should always be remembered that the primary treatment of abnormal tone is a consistent stretching exercise program (with active exercise as is practical for the individual) that improves the health of the muscles (69,70).

### Gait Abnormalities Associated with Aging

Many of the gait changes seen in elderly individuals may be attributed to their reduced walking speed (74,75). However, differences that persist when healthy elderly individuals walk both at comfortable walking speeds and at speeds comparable to those of young adults include reduced peak hip extension, increased anterior pelvic tilt, and reduced ankle plantar flexion and power generation, possibly caused by ankle plantar-flexor muscle weakness (75). Reduced gait velocity, resulting from reductions in both stride length and cadence, has been shown to correlate with lower-limb muscle strength in older persons, but the muscle group that is primarily responsible has been variably suggested to be the ankle dorsiflexors (76), ankle plantar flexors (77), knee extensors and hip flexors (78), and more recently, hip extensors (79).

Falls in the elderly may be due to a variety of factors impacting one or more of a number of organ systems related to balance, such as disturbances of the visual, vestibular, cognitive, sensory, and other systems. However, falls occur in elderly individuals without evident predisposing factors. Although kinematic gait studies have not revealed significant differences between fallers and nonfallers, kinetic gait studies have shown increased peak external hip flexion moment in stance, reduced peak hip extension moment, reduced knee flexion moment in preswing, and reduced knee power absorption in preswing phase (80).

## BODY WEIGHT–SUPPORTED TREADMILL TRAINING

We will conclude this chapter with a brief discussion of body weight–supported treadmill training (BWSTT), a more recently developed approach to gait rehabilitation being utilized with
increasing frequency for patients with diminished or absent supraspinal control. Although this technique was initially developed and then used with varying degrees of success following SCI (81–88), BWSTT has since been a component of locomotor rehabilitation for patients with stroke, traumatic brain injury (TBI), Parkinson’s disease, and lumbar stenosis (89–99), among other conditions. In addition, improvement in cardiovascular fitness from BWSTT has been documented in patients after a stroke or TBI (100–102).

BWSTT employs upright walking on a motorized treadmill while the patient wears a suspension harness to reduce ground reaction force (or GRF) by a specified percentage. Typically, two therapists (and sometimes a third to facilitate upright posture) manually position and guide each lower limb to achieve repetitive, rhythmical stepping motions. To reduce therapist effort and improve the repeatability of locomotor training, commercially available robotic devices have been developed to increase the volume of stepping practice (103–105). Another proposed technologic refinement in the setting of hemiparetic gait has been a computer-controlled, dynamic system to precisely regulate the magnitude and timing—with respect to gait cycle events—of body weight support provided to the user (106).

ACKNOWLEDGMENTS

Vivek Kadyan, MD was our co-author in this chapter as it appeared in the predecessor edition of this textbook. Some of his previous phrasing and other contributions are reproduced here with his permission.

D. Casey Kerrigan, MD created some figures for a predecessor edition that she authored that have been reproduced here with her permission.

REFERENCES


A brief presentation of imaging techniques of interest to the physiatrist must necessarily be selective. Because the diagnosis and initial treatment of fractures are primarily the responsibility of the orthopedic surgeon, with the rehabilitation professional typically involved only later in the course, a full discussion of fractures is not presented in this chapter. Only those fractures that bring patients under the long-term care of the physiatrist are included (e.g., vertebral fractures with the potential to damage the spinal cord). Similarly, tumors and infectious processes are de-emphasized. Rather, emphasis is placed on imaging degenerative musculoskeletal processes, spine and head trauma, stroke, and degenerative central nervous system (CNS) diseases commonly seen by the physiatrist. We will also cover imaging in sport medicine as this is a rapidly changing area in radiology and review the current applications of diagnostic ultrasound in the evaluation of musculoskeletal disorders.

In the past two decades, computed tomography (CT) and magnetic resonance imaging (MRI) have become the most sophisticated imaging modalities for evaluating the musculoskeletal system and the CNS. Therefore, this chapter focuses mainly on the recent applications of CT and MRI in the imaging of musculoskeletal and neural pathology of interest to the physiatrist. In the final section, we will introduce some relatively new imaging technologies of interest to the physiatrist, including advanced MRI methods and ultrasound imaging (USI).

The role of plain film examinations in the assessment of abnormalities of specific joint disorders is well established in the medical literature. A brief review of the most commonly performed radiographic examinations of the extremities will be done when addressing the specific subject.

MUSCULOSKELETAL IMAGING

The advent of the multidetector CT (MDCT) scanner has increased the applicability of this imaging technique for the assessment of the musculoskeletal system. This technology allows for the acquisition of large data set in the axial plane that can be reconstructed in multiple planes of imaging with the use of multiplanar reconstruction (MPR) algorithm. Any anatomical part in the human body can now be scanned in the axial plane and the anatomical information can later be reconstructed in the sagittal, coronal, or any orthogonal plane desired in order to better assess complex anatomical structures such as the joints of the axial skeleton and the spine (1).

CT provides poor contrast resolution to evaluate the musculoskeletal system. Since the relative soft-tissue density of cartilage, tendons, and muscle is similar, we cannot resolve adequate soft-tissue differences between these structures. For example the articular cartilage can only be assessed with CT after a positive contrast is introduced in the joint space such as is the case with CT arthrography. CT however provides a superb spatial resolution that allows for the accurate evaluation of fine soft-tissue and bone trabecular details.

CT images may be displayed with various windows suitable to resolve different structures. Bone window images provide the highest resolution of compact and cancellous bone. Soft-tissue window offers moderate resolution of muscle, tendon, ligament, fat, cartilage, and neural structures.

The good resolution and enhanced contrast of MRI for soft-tissue structures, together with its direct multiplanar imaging capability, make it a superb modality for evaluating all the principal constituents of the musculoskeletal system. Although a technical discussion of the physics of MRI is beyond the scope of this chapter, the physiatrist should know the normal and abnormal MRI appearance of various tissues to be able to look at an MR image with confidence and explain the findings to a patient. The MRI signal intensity of any tissue primarily reflects its proton density, its T1 relaxation time, and its T2 relaxation time. Various techniques, including manipulating the repetition time (TR) between the application of radiofrequency pulses or the echo time (TE) between the radiofrequency pulse and the recording of a signal (i.e., echo) produced by the tissue, can emphasize the proton density, T1 relaxation time, or T2 relaxation time features of any tissue (2). The TR and TE are expressed in milliseconds. The most commonly used technique is spin echo, in which short TR and TE will emphasize the T1 relaxation time of a tissue, the so-called T1-weighted image. In general, an image is said to be T1 weighted if TR is less than 1,000 ms and TE is less than 30 ms (e.g., TR = 500 ms, TE = 20 ms). A T2-weighted image
moderate signal intensity and appears light gray. Fat produces very high signal intensity and appears bright on T1 and T2. Because fat is frequently situated adjacent to ligaments and tendons, it can provide a high-contrast interface for evaluating the integrity of these structures. Adult bone marrow also shows high signal intensity because of its high fat content. Most normal body fluids that are not flowing show low signal intensity on T1-weighted images and high signal intensity on T2-weighted images.

Pathologic processes such as tumor, infection, and abnormal fluids (e.g., edema, joint effusion) show intermediate signal intensity on T1-weighted images and become very hyperintense on T2-weighted images. Pathologic calcifications generally is accomplished with a TR longer than 1,500 ms and a TE greater than 60 ms (e.g., TR = 2,000 ms, TE = 85 ms). Proton density images are obtained with a long TR and a short TE (e.g., TR = 2,000 ms, TE = 20 ms).

Most normal tissues demonstrate similar signal intensities on both T1- and T2-weighted images. Compact bone, fibrocartilage, ligament, tendon, and the rapidly flowing blood within the blood vessel typically produce very low signal intensity, referred to as a signal void, and appear black both on T1 and T2 (Fig. 6-1). Muscle demonstrates moderately low signal intensity and appears dark gray. Peripheral nerves demonstrate slightly higher signal intensity than muscle because of the fat content of their myelinated fibers. Hyaline cartilage produces moderate signal intensity and appears light gray. Fat produces very high signal intensity and appears bright on T1 and T2. Because fat is frequently situated adjacent to ligaments and tendons, it can provide a high-contrast interface for evaluating the integrity of these structures. Adult bone marrow also shows high signal intensity because of its high fat content. Most normal body fluids that are not flowing show low signal intensity on T1-weighted images and high signal intensity on T2-weighted images.

Pathologic processes such as tumor, infection, and abnormal fluids (e.g., edema, joint effusion) show intermediate signal intensity on T1-weighted images and become very hyperintense on T2-weighted images. Pathologic calcifications

FIGURE 6-1. Normal shoulder MR images. A: An axial scout film with cursors displays the oblique coronal planes parallel to the plane of the scapula, which allow optimal visualization of the supraspinatus. B: An oblique coronal image demonstrates the supraspinatus muscle belly (SsB), supraspinatus tendon (SsT), subacromial-subdeltoid fat plane (FP), acromioclavicular joint (ACJ), deltoid muscle (D), articular cartilage of humeral head and glenoid (AC), glenoid (G), and humeral head (H). C: An axial image displays humeral head (H), glenoid (G), glenoid labrum (L), the inferior glenohumeral ligament (IGHL), deltoid muscle (D), subscapularis tendon (Sct), and biceps tendon (BT). D: A sagittal section demonstrates good resolution of the coracoacromial ligament (CAL) extending from the coracoid process (CP) to the acromion; coracoid process (CP), supraspinatus (SS), infraspinatus (IS).
demonstrate very low signal intensity on both T1- and T2-weighted images.

The direct multiplanar imaging capability of MRI is particularly useful in evaluating obliquely oriented musculoskeletal structures such as the supraspinatus tendon, the cruciate ligaments, and the lateral collateral ligaments of the ankle.

MRI has proved useful in evaluating traumatic, degenerative, inflammatory, and neoplastic pathology of the limbs and spine. It is useful in detecting acute or chronic traumatic injuries and degenerative conditions involving bones, muscles, tendons, ligaments, fibrocartilage, and nerves. Bone pathology, particularly well detected by MRI, includes contusions, osteochondral injuries, stress fractures, marrow replacement by neoplastic cells, and ischemic necrosis. Muscle lesions that MRI is especially sensitive at identifying include strain or contusion, complete rupture, compartment syndrome, myopathies, and atrophy (3). Tendon conditions well depicted by MRI include partial and complete tear, tendinitis, and tenosynovitis. MRI is also very sensitive for detecting partial or complete ligament tears. Fibrocartilaginous injuries or diseases well delineated by MRI include pathology of the menisci, the glenoid labrum, the triangular fibrocartilage of the wrist, and the intervertebral disc. Nerve entrapments well visualized by MRI include spinal nerve encroachment by disc disease or spinal stenosis and carpal tunnel syndrome (CTS) or other entrapment syndromes. Enhanced imaging of normal and injured peripheral nerves can be obtained using a short-tau inversion recovery (STIR) excitation-emission sequence due to the increased sensitivity to free water content associated with tissue edema using this MR recording protocol. The increased signal generated by injured nerves using STIR pulse sequences probably reflects an increase in free water content of the nerve due to altered axoplasmic flow, axonal and/or myelin degeneration, and endoneurial or perineurial edema due to a breakdown in the blood-nerve barrier (4). MRI imaging of denervated skeletal muscle shows increased MR signal using the STIR protocol when there is significant muscular weakness and well-defined changes indicating muscle denervation on needle electromyography (5).

Osteomyelitis causes a reduction in bone marrow signal intensity on T1-weighted images because of the replacement of normal fatty marrow by inflammatory exudate. In T2-weighted images, these areas of active infection become hyperintense.

MRI has particular value in evaluating both bone and soft-tissue neoplasms. Most of them demonstrate moderately low signal intensity on T1-weighted images and very high signal intensity on T2-weighted images.

Emphasis will now be directed to the application of imaging modalities to common regional pathologic conditions of the musculoskeletal system. Particular focus will be given to MRI because of its superb soft-tissue imaging capabilities and its rapidly expanding diagnostic applications.

The important role played by radiology in the diagnosis of diseases has come at the expense of increased radiation exposure to the general population. With the advent of new technology, such as Positron Emission Tomography (PET) studies and MDCT technology, there has been a sharp increase in the number of radiographic examinations performed to the general population and as a consequence an increase in the cumulative radiation dose to the individual patient and the general population. It is an expected outcome for increase in radiation exposure to lead to an increased rate of malignancy. Therefore, increased awareness is needed in issues concerning radiation safety.

Ionizing radiation, especially at high doses, is known to increase the risk of developing cancer. It is estimated that medical exposure might be responsible for 1% of cancer diagnosis in the United States. This rate is expected to increase in the coming years due to the increased number of examinations performed today.

The scientific measurement for the effective dose of radiation is the millisievert (mSv). The background radiation dose for the average person in the United States is about 3 mSv. This is secondary to cosmic radiation and naturally occurring radioactive materials. By comparison, the effective radiation dose for a spinal CT is equivalent to 6 mSv or 2 years of natural background radiation. Radiation exposure is particularly important in pregnant women and pediatric patients due to the cumulative life effect of radiation exposure at a younger age. In nuclear medicine examinations, special precautions are needed. Some of the radiopharmaceuticals used in nuclear medicine can pass into the milk of lactating women (6).

The relative radiation level (RRL) is a radiation measurement used to calculate effective dose. This is the dose used to estimate population total radiation risk associated to an imaging procedure. This takes into account the sensitivity of different body organs and tissues. This estimate cannot assess the specific risk of an individual patient.

Every effort should be made to order examination, which is best indicated to address the clinical concern of the patient. To aid in this regard, the American College of Radiology (acr.org) has established guidelines for the appropriate use of imaging to answer specific clinical questions. The appropriateness criteria can be of help when deciding which imaging study to order to answer a clinical question.

Shoulder

Plain film radiographic evaluation of the shoulder should include frontal examinations with internal and external humeral rotation. If there is a question of instability or dislocation, an axillary view, a scapular Y view, or both should be obtained. There have been several reports that recommend the use of a 30-degree caudal-angled radiograph or a suprascapular outlet view for the assessment of the anterior acromion in cases of suspected shoulder impingement. Since these are special views, they must be ordered as routine shoulder radiographs do not include axillary or suprascapular outlet views. The RRL for plain film radiographic examinations of the shoulder is less than 0.1 mSv, which is considered minimal.

MRI has become valuable in evaluating a host of shoulder abnormalities very familiar to the physiatrist. These include impingement syndrome, other rotator cuff abnormalities, instability syndrome, and bicipital tendon abnormalities. It is
also useful in demonstrating arthritic changes, occult fractures, ischemic necrosis, and intra-articular bodies. MRI with intra-articular contrast is now considered the modality of choice for the evaluation of labral and capsular pathology. The use of MRI for shoulder evaluation avoids radiation exposure to the nearby thyroid gland, which can occur with CT examinations. The excellent visualization of marrow by MRI permits early diagnosis of ischemic necrosis, infection, and primary or metastatic tumors.

Because of the oblique orientation of the scapula on the chest wall and the consequent anterolateral facing direction of the glenoid, the direct multiplanar imaging capability of MRI provides optimal visualization of all the important shoulder structures. An oblique coronal image parallel to the plane of the scapula provide full-length views of the rotator cuff musculature, especially the supraspinatus and is the best plane for the evaluation of injuries to the biceps-labral complex (BLC) (Fig. 6-1A and B). Coronal oblique images can also provide information about the presence of impingement upon the supraspinatus by the acromion and osteophytes in the presence of acromioclavicular joint osteoarthritis. Oblique sagittal imaging planes parallel to the glenoid provide cross-sectional views of the rotator cuff apparatus and evaluates the anatomical configuration of the coracoacromial arch and the presence of impingement (Fig. 6-1D). Axial imaging planes provide good visualization of the anterior and posterior capsular apparatus, glenoid labrum, bony glenoid rim, and humeral head (Fig. 6-1C).

**SHOULDER IMPINGEMENT SYNDROME AND SUPRASPINATUS INJURY**

The MRI findings of shoulder impingement syndrome and its associated supraspinatus injury are best seen on oblique coronal MR images that visualize the full length of the supraspinatus muscle belly and tendon (Fig. 6-1B). The normal muscle belly displays moderately low signal intensity. The tendon is visualized as an intermediate—signal-intensity structure that blends with the low signal intensity of the superior capsule as it courses to its insertion on the greater tubercle of the humerus. The tendon demonstrates smooth tapering from medial to lateral into its insertion in the greater tuberosity. The inferior aspect of the tendon is delimited below by the moderate signal intensity of the hyaline cartilage on the superior aspect of the humeral head. The superior aspects of both the muscle belly and tendon are delimited by a high–signal-intensity subacromial and subdeltoid fat plane. The normal subacromial-subdeltoid bursa is not specifically visualized because its walls are separated only by monomolecular layers of a synovial-type fluid, but it is situated between the supraspinatus tendon and the fat plane. Above the fat plane, the clavicle, acromioclavicular joint, acromion, and deltoid muscle are demonstrated on different oblique coronal sections.

Although rotator cuff impingement is a clinical diagnosis, MRI can provide direct visualization of the constituents to the coracoacromial arch and their relationship to the supraspinatus (Fig. 6-2A and B). Downward slanting of the acromion in the coronal or the sagittal plane, a thickened coracoacromial ligament or inferior osteophytosis within the acromioclavicular joint can exert mass effect upon the supraspinatus. This has been implied as being in part responsible for chronic tears of the supraspinatus.

Neer stated that 95% of rotator cuff tears are associated with chronic impingement syndrome (7) and described three stages in the progression of rotator cuff injury. These can be visualized by MRI (6–10). Stage 1 is characterized by edema and hemorrhage within the supraspinatus tendon characteristic of an early tendinitis. On MRI, there is focal tendon

---

**FIGURE 6-2.** A: Coronal oblique T2-weighted pulse sequence with fat suppression demonstrates downward slanting to the acromion (long arrow), which is against the supraspinatus tendon (*). Note focal area of increased signal at the myotendinous junction of the SsT (short arrow). B: Sagittal oblique T2WI with fat suppression demonstrates to a better advantage the inferior slanting to the acromion against the SsT (short arrow). Note focal tendinosis (long arrow).
thickening and diffuse moderate increase in signal intensity within the tendon (Fig. 6-3A–C). In stage 2, Neer described both inflammation and fibrosis within the tendon. MRI shows this as thinning and irregularity of the tendon. Stage 3 is a frank tear of the supraspinatus tendon. On MRI, complete tears are noted by a discontinuity of the tendon with a well-defined focus of high signal intensity on T2-weighted images (Fig. 6-4). The most susceptible area is the critical zone of hypovascularity, located about 1 cm from the insertion (11). With small or partial tears, there is no retraction of the muscle-tendon junction, the subacromial-subdeltoid fat plane is commonly obliterated, and fluid may accumulate in the subacromial-subdeltoid bursa, which becomes hyperintense on T2-weighted images. There also may be effusion of the shoulder joint, which may extend inferiorly along the tendon sheath about the long head of the biceps. With a complete supraspinatus tendon tear, the muscle belly may retract medially, and atrophy may occur as the tear becomes chronic (Fig. 6-5A,B). Muscle atrophy appears as areas of high signal

**FIGURE 6-3.** A: T1-weighted MRI of focal supraspinatus tendinosis demonstrating focal thickening and slight increase signal intensity to the tendon (arrow). B: Coronal oblique T2-weighted fat suppressed sequence with increased signal intensity within the area of tendinosis (arrowhead). There is fluid within the adjacent subdeltoid bursa (long arrow). C: Sagittal oblique T2-weighted fat suppressed image demonstrates the area of increased signal to be within the anterior superior portion of the rotator cuff representing fibers of the supraspinatus tendon (short arrow).

**FIGURE 6-4.** Complete rupture of the supraspinatus tendon is seen in a T2-weighted MRI. There is fluid filling the gap (arrow) and there is retraction to the tendon fibers underneath the acromion.
intensity because of fatty replacement within the muscle belly and decreased muscle mass. Finally, the acromiohumeral interval narrows as the humeral head migrates superiorly, because of the loss of supraspinatus restraint to the deltoid’s tendency to sublux the humerus superiorly during abduction.

**SHOULDER INSTABILITY AND DISRUPTION OF THE ANTERIOR CAPSULAR MECHANISM**

For the assessment of shoulder instability and labral tears, it is imperative that intra-articular contrast medium be injected in order to be able to evaluate the entire articular labrum, glenoid fossa, and capsular mechanism (12). Axial MR images provide the best visualization of the anterior and posterior glenoid labra, capsule, and lower rotator cuff muscles (Fig. 6-1). Anteriorly, the moderate–signal-intensity subscapularis muscle belly and its low–signal-intensity tendon are visualized. The tendon fuses with the low–signal-intensity anterior capsule as it courses to its insertion on the lesser tubercle. The fibrocartilaginous anterior and posterior labra appear as low–signal-intensity triangular or rounded areas attached to the glenoid and humeral head. The posterior capsule is visualized as a low-intensity area blending with the deep surface of infraspinatus and teres minor muscles as they extend to their insertions on the greater tubercle of the humerus. The long tendon of the biceps is demonstrated as a round, low–signal-intensity area within the bicipital groove.

Shoulder instability and the associated disruption of the anterior capsular mechanism can cause chronic shoulder pain and disability. The instability may be caused by an acute traumatic episode or can occur with no history of a traumatic event. Both recurrent traumatic subluxation and nontraumatic instability are typically associated with disruption of the anterior capsular mechanism. Anteriorly, where most instability occurs, this mechanism includes the subscapularis muscle and tendon, the anterior joint capsule, three underlying glenohumeral ligaments, the synovial lining, and the anterior labrum. With instability, the labrum shows tears, separation from the glenoid rim, or degeneration (13). Also frequently present are medial stripping of the capsule from its normal attachment to the labrum and glenoid rim, an enlarged fluid-filled subscapular bursa secondary to joint effusion, attenuation of the glenohumeral ligaments, and injury or laxity of the subscapularis muscle or tendon.

By MRI, labral tears may be visualized as discrete linear areas of increased signal intensity within the normal signal void of the labrum (Figs. 6-6 and 6-7). These areas show moderate intensity on T1-weighted images and high intensity on T2-weighted images. With recurrent dislocation or subluxation, the labrum can become fragmented or attenuated.

Capsular detachment from the scapula (i.e., stripping) is visualized by T2-weighted MRI as an area of high–signal-intensity fluid dissecting medially from the glenoid rim. With trauma to the subscapularis tendon, there can be medial retraction of the muscle-tendon junction when the tendon is completely ruptured. Chronic atrophy of the subscapularis muscle belly is identified by high–signal-intensity fatty replacement. The glenoid marrow underlying a labral detachment may show pathologically decreased signal intensity even before the plain film radiograph shows an osseous Bankart lesion. MRI and CT can be used to visualize Bankart fractures of the anterior glenoid and the Hill-Sachs compression deformity of the posterolateral humeral head (14,15) and are both useful in the assessment
an increased signal intensity area within the tendon. This can progress to frank rupture of the tendon with a high-signal-intensity area on T2-weighted images and may be associated with joint effusion. A complete tear will eventually cause muscle retraction and later atrophy.

Calcific tendinosis (Fig. 6-8) of the supraspinatus is a common clinical entity most commonly affecting middle aged persons. It is slightly more common in females and can affect multiple tendons in the body. It is however far more frequently in the supraspinatus. Although the exact etiology is unknown, it is felt to be secondary to chronic ischemia of the tendon fibers.

of the extent of a Hill-Sachs defect in patients with engaging lesions. Patients with the rarer posterior instability show similar posterior labral, capsular, and muscular defects.

**TENDINITIS OR RUPTURE OF OTHER SHOULDER MUSCLES**

Tendinitis and rupture also can involve the subscapularis, infraspinatus, and teres minor, or biceps tendons, although far less commonly than the supraspinatus. Early tendinitis involves an increased signal intensity area within the tendon. This can progress to frank rupture of the tendon with a high-signal-intensity area at the site of the tear on T2-weighted images and may be associated with joint effusion. A complete tear will eventually cause muscle retraction and later atrophy.

Calcific tendinosis (Fig. 6-8) of the supraspinatus is a common clinical entity most commonly affecting middle aged persons. It is slightly more common in females and can affect multiple tendons in the body. It is however far more frequently in the supraspinatus. Although the exact etiology is unknown, it is felt to be secondary to chronic ischemia of the tendon fibers.
Ischemic necrosis is more fully described with the hip, where its incidence is higher.

**ISCHEMIC NECROSIS OF THE HUMERAL HEAD**

As in other joints, ischemic necrosis of the humeral head is depicted as an area of decreased signal intensity within the subarticular bone marrow in T1-weighted images. On T2-weighted images, curvilinear bright bands surrounding areas of decreased signal intensity, the so-called tram track sign, represent reactive marrow surrounding the core of dead bone. Ischemic necrosis is more fully described with the hip, where its incidence is higher.

**ELBOW**

Plain film radiographic examination of the elbows should be the initial evaluation for patients with chronic elbow pain. Radiographs can be useful for the assessment of calcium within the joint compartment or periarticular soft tissues. Standard frontal and lateral radiographs are used for the routine evaluation of the elbow joint. Radiographic examination of the elbow has a minimal RRL.

MRI has not been applied to the evaluation of elbow pathology as extensively as it has been to other large joints (17). However, improving imaging techniques and the use of surface coils permit superb visualization of the bony, ligamentous, muscular and neurovascular structures around the elbow. Common elbow injuries evaluated with MRI are usually related to sports (weightlifting, throwing, and racquet sports) or compartmental nerve entrapment.
Axial MRI views of the elbow region permit good visualization of the biceps, brachialis, triceps, and all the extensor and flexor muscles of the forearm (Fig. 6-9A). High–signal–intensity fat planes and low–signal–intensity intermuscular septa permit clear delineation of each muscle and their tendons of insertion or origin. Axial images clearly depict brachial, ulnar, and radial arteries and all the subcutaneous and deep veins. They also allow identification of the ulnar nerve within the cubital tunnel and the radial nerve in the brachioradialis–brachialis interval and under the supinator muscle’s arcade of Frohse, where it is commonly entrapped. The median nerve is visualized at all its common elbow entrapment sites, including under the bicipital aponeurosis, between the heads of the pronator teres, and under the fibrous arch of the flexor digitorum superficialis.

The humeroulnar, humeroradial, and proximal radioulnar joint spaces and articular cartilages are well visualized on both coronal and sagittal MR images (Fig. 6-9B and C). The low–signal–intensity ulnar collateral, radial collateral, and annular ligaments are depicted on both axial and coronal MR images. Sagittal images delineate the anterior and posterior subsynovial fat pads.

MRI has the capability of directly visualizing degenerative or traumatic abnormalities of the annular and the radial and ulnar collateral ligamentous complexes. A sprain appears in MRI as thickened or thinned ligament with surrounding high T2 signal intensity. The collateral ligaments may show degeneration in association with adjacent epicondylitis. The affected ligament commonly shows thickening and intermediate signal intensity. Full thickness of avulsion ligamentous tears appears as discontinuities of the low–signal–intensity ligament. The T2-weighted images disclose hyperintense edema and hemorrhage between the torn ends of the ligament extending into the joint interval and adjacent soft tissues. A partial thickness tear appears as high T2 fluid signal intensity within an uninterrupted ligament (Fig. 6-10).

MRI also provides good visualization of the sites of muscle injury and denervation about the elbow (18) (Fig. 6-11A,B). Acute muscle denervation is demonstrated by increased T2-weighted signal intensity within the specific muscle
thickening, increased signal intensity, and discontinuity of ten-don fibers are the findings commonly observed in tendon tears (Fig. 6-12A,B).

MRI has the ability to demonstrate tendinosis involving the common extensor and flexor tendon origins from the lateral and medial aspects of the humerus with findings similar to those described in tendinosis about the shoulder. It also can display abnormalities of the radial and ulnar collateral liga-

Wrist
Plain film radiography and CT provide good visualization of the osseous structures of the wrist. Most physicians agree that the imaging evaluation of a painful hand and wrist should begin with radiographs. This inexpensive study may establish a specific diagnosis in arthritis, injury, infection, and wrist insta-

CARPAL TUNNEL SYNDROME
MRI can serve as an adjunct diagnostic tool for CTS when the clinical or neurophysiologic findings are equivocal. The carpal tunnel is a fibro-osseous space with little fat that contains the flexor tendons and the median nerve. The flexor retinaculum composes the volar aspect of the tunnel and normally shows slight palmar bowing. The median nerve courses through the tunnel within its volar and radial aspect, and it can be different-

1. Distal radioulnar joint before the median nerves enter the tunnel.
2. Proximal tunnel, at the level of the pisiform.
3. Distal tunnel, at the level of the hook of the hamate.
**FIGURE 6.12.** Sagittal A and axial B FSE T2-weighted fat suppressed images of the distal arm in a patient with a complete biceps tear. A: The tendon free margin is retracted (long arrow in A). B: There is significant edema (arrowheads) surrounding the retracted tendon (long arrow). B, biceps muscle; Br, brachialis muscle.

**FIGURE 6.13.** Normal wrist anatomy as seen on T1-weighted MR images. Axial MR images are at the levels A: of the distal radioulnar joint, B: the proximal and C: the distal carpal tunnel. D: Longitudinal MRI through the median nerve within the carpal tunnel. C, capitate; FDP, flexor digitorum profundus; FDS, flexor digitorum superficialis; FR, flexor retinaculum; H, hamate; L, lunate; MN, median nerve; PDN, palmar digital branches of the median nerve; R, radius; T, trapezium; U, ulna; UA, ulnar artery; UN, ulnar nerve. Note fracture through the base of the hook of the hamate (arrow in B).
There are four universal findings of CTS visible by MRI regardless of etiology (21):

1. Swelling of the median nerve (i.e., pseudoganglion) in the proximal part of the carpal tunnel at the level of the pisiform. Best evaluated by comparing the size of the median nerve at the level of the distal radioulnar joint with its size at the proximal tunnel.

2. Increased signal intensity of the edematous median nerve on T2-weighted images.

3. Palmar bowing of the flexor retinaculum, determined by a bowing ratio of more than 15%. The bowing ratio is calculated by drawing a line from the trapezium to the hook of the hamate on the axial plane. The distance from this line to the flexor retinaculum is divided by the previously calculated length.

4. Flattening of the median nerve in the distal carpal tunnel at the level of the hamate (Fig. 6-14A–D).

MRI also has the potential to establish the cause of CTS. Some of the etiologies visualized by MRI include traumatic tenosynovitis, rheumatoid tenosynovitis, a ganglion cyst of a carpal joint, excessive fat within the carpal tunnel, a hypertrophied adductor pollicis muscle in the floor of the carpal tunnel, and a persistent median artery (23).

MRI also provides a means of postoperative evaluation of those patients in whom the symptoms persist, to ensure that the flexor retinaculum has been completely incised and that there are no other complicating postoperative factors producing continuing discomfort. When the flexor retinaculum has been completely incised, the incision site is well documented by MRI and the contents of the carpal tunnel are typically displaced forward (Fig. 6-15A). If the distal part of the flexor retinaculum has been incompletely incised, this can be demonstrated by MRI, and the preoperative MRI findings of CTS will persist (Fig. 6-15B and C).

FIGURE 6-14. Axial FSE T2-weighted fat suppressed images (A, B, C) and sagittal T1-weighted image in a patient with carpal tunnel syndrome. A, B: There is a normal size to the nerve within the tunnel (arrowheads). C: There is thickening, increased girth and increased signal intensity proximal to the flexor retinaculum (arrowheads). D: Note tapering to the nerve as it approaches the carpal tunnel in the sagittal view (arrowheads).
OTHER WRIST ABNORMALITIES

MRI can visualize postincisional neuromas as lobulated masses in the typical location of the palmar cutaneous branches of the median nerve. Other peripheral nerve tumors such as schwannomas (Fig. 6-16) and neurofibromas can be well recognized as well. It can also demonstrate tenosynovitis involving any of the tendons crossing the wrist. MRI also displays marrow abnormalities such as ischemic necrosis of the proximal fragment of a scaphoid fracture and avascular necrosis of the lunate, where the marrow shows reduced signal intensity (24). MRI has the ability to evaluate the integrity of the intrinsic/extrinsic ligaments of the wrist and the triangular fibrocartilage complex (TFCC) (25). The TFCC, scapholunate, and lunotriquetral ligaments are best evaluated with MR arthrography (Fig. 6-17).

Hip

Similar to the remaining joints in the body, plain film radiographic examinations should be the first diagnostic test in patients with chronic hip pain. Frontal and frog leg views

**FIGURE 6-15.** Postoperative MR of carpal tunnel syndrome. A: Axial T1-weighted image. There has been release to the flexor retinaculum (short arrow). The median nerve insinuates through the surgical defect. B: Axial T2-weighted fat suppressed image. The median nerve (arrow) has intermediate signal and is better delineated. C: Patient with failed carpal tunnel release. There is a linear area of decreased signal intensity (short arrow) which was found to represent a fibrous band at surgery. The median nerve (long arrow) is flattened underneath the fibrous band. (A and B; courtesy of Zehava Rosenberg, NY. C; courtesy of Mark Kransdorf, FL.)
with plain films. Special radiographic views are useful in the assessment of particular clinical problems. Patients suspected of femoroacetabular impingement can benefit from cross table lateral radiographs of the hip joint.

MRI of the hip is usually performed in standard axial, coronal, and sagittal planes; although oblique axial images are commonly used in routine protocols. The normal marrow of the femoral head epiphysis and the greater trochanter displays very high signal intensity and is surrounded by a thin layer of compact bone that appears as a signal void. However, in children and young adults, the marrow of the femoral neck and shaft normally shows lower signal intensity, because it contains some residual hematopoietic marrow. The acetabulum has imaging characteristics similar to those of the femoral neck. The periphery of the hip joint interval displays the moderate signal intensity of the apposed hyaline cartilage surfaces of the femoral head and acetabulum, whereas the centrally situated acetabular notch contains high-signal-intensity fat. The thick, very-low-signal-intensity hip joint capsule blends with the acetabular labrum proximally and the cortex of the femoral neck distally. All the muscles, nerves, and blood vessels crossing the hip are well visualized.

ISCHEMIC NECROSIS

One of the common indications for MRI of the hip is to determine the presence of ischemic necrosis. This is bone death produced by a compromised blood supply. It also has been called avascular necrosis, osteonecrosis, or aseptic necrosis. Predisposing factors that should raise the physician’s index of suspicion include corticosteroid therapy, alcoholism, known hip trauma, chronic pancreatitis, Gaucher’s disease, sickle cell disease, exposure to hypobaric conditions, subcapital fractures, childhood septic arthritis or osteomyelitis of the hip, and congenital hip dislocation (27). If undetected early, the disease can progress and finally undergo irreversible collapse of the femoral head. MRI has been demonstrated to be even more sensitive and specific than bone scintigraphy for the early diagnosis of ischemic necrosis of the femoral head (28–31).

On T1-weighted MRI, the foci of ischemic necrosis of the femoral head appear as homogeneous or heterogeneous well-delimited or diffuse areas of decreased signal intensity in the shape of rings, bands, wedges, or crescents, or in an irregular configuration (Fig. 6-18A–C) (32–33). The low signal intensity is caused by death of marrow fat and replacement of the marrow by a fibrous connective tissue. Some cases show a lower signal band surrounding the lesion, and this has been attributed to healing sclerotic bone at the interface between normal and necrotic bone. On T2-weighted images, many cases show a double-line sign with a high-signal-intensity zone just inside of a low-signal-intensity margin. This is thought to be produced by granulation tissue surrounded by sclerotic bone (31–33).
FEMOROACETABULAR IMPINGEMENT

Femoroacetabular impingement is another cause of hip pain that may initially have a non–specific clinical presentation; however, a thorough physical evaluation should be able to identify this condition. Pain can be elicited on physical exam by passive movement of the thigh into full flexion, adduction, and internal rotation (34). Radiographs often help to identify anatomical variations such as dysplasia of the femoral neck (Fig. 6-19) or acetabular overcoverage. MR arthrography (Fig. 6-20) can be a helpful imaging study for evaluating the consequences of femoroacetabular impingement as well as thoroughly evaluating the abnormal anatomy. Intra-articular contrast helps to better delineate hyaline cartilage defects as well as labral tears. These findings, along with the clinical evaluation, help guide the appropriate treatment plan.

OTHER HIP ABNORMALITIES

Transient regional osteoporosis presents with a low–signal-intensity lesion on T1-weighted images that is similar to ischemic necrosis, but it typically involves both femoral head and
Patients complaining of traumatic or nontraumatic knee pain should undergo plain film radiographic examination with AP and lateral (Lat) knee radiographs. Patients, who complain of anterior knee pain, should undergo an additional axial (Merchant) view to the patella. In elderly patient, where degenerative osteoarthritis is highly prevalent, standing views of the knees should be order. Standing views with the knee flexed can better predict the degree of cartilaginous loss within the posterior medial and lateral knee compartments (36,37).

Knee

Patients complaining of traumatic or nontraumatic knee pain should undergo plain film radiographic examination with AP and lateral (Lat) knee radiographs. Patients, who complain of anterior knee pain, should undergo an additional axial (Merchant) view to the patella. In elderly patient, where degenerative osteoarthritis is highly prevalent, standing views of the knees should be order. Standing views with the knee flexed can better predict the degree of cartilaginous loss within the posterior medial and lateral knee compartments (36,37).
MENISCAL INJURIES

All parts of both menisci are well visualized by MRI. Sagittal MR images provide good views of the anterior and posterior horns and a fair view of the body of both menisci. In more central sections, both horns of the menisci appear as wedge-shaped signal voids contrasted on their superior and inferior surfaces by the moderate signal intensity of the hyaline cartilage on the articular surfaces of the femur and tibia. In more peripheral sections, where the images are tangential to the circumference of the menisci, they appear bow tie shaped (Fig. 6-22). Coronal MR images provide the best visualization of the bodies of both menisci.

There are three types of meniscal findings visualized by MRI (16,39,40). One is the presence of small globular or irregular high–signal-intensity foci confined to the interior of the meniscus. This is considered to be an early type of mucoid degeneration. A second type of meniscal MRI finding is the presence of a linear region of increased signal intensity within the meniscus that does not extend to either the femoral or tibial articular surface of the meniscus but may extend to the meniscocapsular junction. Histologically, this represents fragmentation and separation of the fibrocartilage and is considered by many to be an intrameniscal tear. The significance of the globular or linear signals that do not extend to either articular surface of the meniscus is not fully agreed on (41).

Frank meniscal tears are demonstrated by MRI as linear or irregular areas of signal intensity that extend to one or both articular surfaces of the meniscus (Fig. 6-23). The high signal intensity is produced by synovial fluid in the crevices within the meniscus. These meniscal tears can be horizontal, vertical, or complex. Bucket-handle tears are vertical tears where the inner meniscal fragment is displaced toward the intercondylar notch (Fig. 6-23A,C,D). At times, repeated trauma or chronic degeneration may cause a gross distortion of meniscal shape, and the meniscus may then appear to have a truncated apex or to be grossly small with a free fragment.

Other meniscal abnormalities well visualized by MRI include discoid meniscus, meniscal cysts, and abnormalities involving the postoperative meniscus. In discoid meniscus, typically involving the lateral meniscus, there is a continuous bridge of meniscal tissue between the anterior and the posterior horns in the central part of the joint. Meniscal cysts are usually associated with underlying horizontal meniscal tears through which synovial fluidcollects at the meniscocapsular junction (42). They show high signal intensity on T2-weighted images. MRI also can be used to evaluate the postmeniscectomy patient with continuing or recurrent symptomatology (39). It can detect an incompletely excised meniscal tear, retained meniscal fragments, or a tear developing within the residual part of the meniscus. MR arthrography with gadolinium can be helpful to distinguish between retears and old healed tears that might still show increased signal intensity on T2-weighted images (43).

CRUCIATE LIGAMENT INJURIES

The cruciate ligaments are best visualized by sagittal or oblique sagittal MR images that display the full length of the ligaments (Fig. 6-24A). On straight sagittal images, the slender nature of the anterior cruciate ligament and its oblique course cause a volume-averaging effect that averages fat signal intensity about the ligaments with the normal low signal intensity of the ligament so that the anterior cruciate ligament frequently does not appear as a complete signal void. Furthermore, straight sagittal images typically fail to demonstrate the anterior cruciate ligament’s femoral attachment because of its oblique orientation in both sagittal and coronal planes. Oblique sagittal images that parallel the ligament show the full thickness and length of the anterior cruciate ligament without subjecting it to partial volume averaging (44).

In the extended position of the knee, which is typically used for MR images, the anterior cruciate ligament is normally taut. The posterior cruciate ligament is a thicker ligament and is therefore well visualized on straight sagittal MR images (Fig. 6-24B). It can be visualized as a signal void structure from its attachment to the posterior tibial intercondylar area to its attachment on the medial femoral condyle. With the knee extended, the posterior cruciate ligament is visualized as thick and posteriorly bowed. It straightens with knee flexion. Axial images can be very helpful to evaluate the femoral insertion of the anterior and the posterior cruciate ligaments.
The MRI appearance of an anterior cruciate ligament injury depends on the site and degree of disruption, as well as on the age of the tear. A complete tear may be visualized as a discontinuity of the ligament (Fig. 6-25A–D). In the acute complete tear, the interval between the torn ends of the ligament is often occupied by a mass of intermediate signal intensity on T1-weighted images that appears hyperintense on T2-weighted images (45). At other times, the torn ligament may present as a fusiform or irregular soft-tissue mass of intermediate signal intensity on T1-weighted images that appears hyperintense on T2-weighted images. These fluid masses are usually a combination of edema and hemorrhage, and there may be an associated joint effusion. If the ligament tears from its femoral attachment, the axial images will show fluid signal between the lateral femoral condyle and the expected ligament insertion. In partial tears there is no complete discontinuity, but the ligament that appears intact on a T1-weighted image may show a hyperintense signal on T2-weighted images, or the ligament may display an interrupted or concave anterior or posterior margin when the knee is extended (42). In chronic anterior cruciate ligament deficiency, there may be a complete absence of the ligament, or there may be only remnants remaining in its usual location. Some secondary signs of anterior cruciate ligament injury may be present. These include a forward shift of the tibia and an anterior bowing or buckling of the posterior cruciate ligament caused by the position of the
CHAPTER 6  |  IMAGING TECHNIQUES RELATIVE TO REHABILITATION

FIGURE 6-24. Normal cruciate ligaments. A: An oblique sagittal MRI parallel to the anterior cruciate ligament demonstrates excellent visualization of all borders and attachments of the anterior cruciate ligament (arrow). B: A T2-weighted MRI of the posterior cruciate ligament (arrow), which is normally posteriorly bowed when the knee is extended.

knee within the coil, which duplicates the knee position of an anterior drawer or Lachman test (39).

On T1-weighted MR images, partial tears of the posterior cruciate ligament typically appear as foci of increased signal intensity within the normal black signal void of the ligament. These appear hyperintense on T2-weighted images. With complete tears, a frank discontinuity is visualized with an intervening fluid mass that becomes hyperintense on T2-weighted images (Fig. 6-26). The gap between the ends of a completely torn posterior cruciate ligament can be exaggerated by imaging the knee in flexion, which tenses the posterior cruciate ligament.

COLLATERAL LIGAMENT INJURIES

The collateral ligaments are best visualized by coronal MR images (Fig. 6-27). The medial collateral ligament appears as a narrow low–signal-intensity band extending from the medial epicondyle of the femur to an attachment on the anteromedial aspect of the tibia 5 to 6 cm below the joint line. It is overlaid at its tibial attachment by the tendons of the pes anserinus, which are separated from it by an intervening anserine bursa that is not visualized unless it is inflamed. Deep to the tibial collateral ligament, the medial capsular ligament, sometimes called the deep portion of the tibial collateral ligament, has femoral and tibial attachments close to the joint interval and deep attachments to the medial meniscus, referred to as the meniscofemoral and meniscotibial or coronary ligaments. Valgus and rotary stresses can injure the medial capsular ligament or the tibial collateral ligament, usually in that order (39). In a complete rupture (i.e., grade III injury), MRI can show discontinuity, serpentine ligamentous borders, and edema within adjacent connective tissues. In a partial tear (i.e., grade II injury) or in the case of microtears confined to the ligament substance (i.e., grade I injury), the ligament may show no discontinuity, but the overlying subcutaneous fat typically demonstrates edema and hemorrhage, which is indicated by moderate signal intensity on T1-weighted images and high signal intensity on T2-weighted images. Injury to the tibial collateral ligament is commonly associated with injuries to the anterior cruciate ligament and medial meniscus.

The lateral collateral complex commonly refers to the lateral supporting structures of the knee, whose main components are the iliotibial tract, the lateral collateral ligament, the long head of the biceps femoris, and the popliteofibular ligament. These structures are best seen on axial and coronal MR images as a low–signal-intensity band extending somewhat obliquely from the lateral femoral epicondyle to the fibular head. The lateral collateral ligament is usually injured by varus and rotary stresses to the knee, although its frequency of injury is less than that of the tibial collateral ligament. The MRI findings of the injured fibular collateral ligament are similar to those for the tibial collateral ligament.

OTHER KNEE ABNORMALITIES

Patellar tendinitis ( jumper’s knee) is demonstrated by MRI as an area of edema within the patellar ligament (i.e., tendon) at its patellar (Fig. 6-28) or tibial tuberosity attachment. There is also associated edema in the adjacent subcutaneous fat or the infrapatellar fat pad.

Ischemic necrosis about the knee most commonly involves the weight-bearing surface of the medial femoral condyle, and its MRI findings are as described for the hip.
Osteochondritis dissecans occurs mainly in adolescents and involves a partial or total separation of a segment of articular cartilage and subchondral bone from the underlying bone (39). It commonly involves the intercondylar portion of the medial femoral condyle articular surface. It is visualized on T1-weighted MR images as a low-signal-intensity region in the subchondral bone with or without disruption of the overlying articular cartilage (Fig. 6-29A,B). If the involved osteochondral segment becomes completely separated from the underlying bone, it becomes an intra-articular loose body. The role of MRI in osteochondritis dissecans is mainly to determine the stability of the fragment, because the treatment hinges on that.
Popliteal (i.e., Baker's) cysts and other synovial cysts about the knee appear hyperintense on T2-weighted images (Fig. 6-31A–D). They can be visualized on axial, sagittal, or coronal images. Popliteal cysts are usually an enlargement of the semimembranosus-gastrocnemius bursa, which is located between the tendon of insertion of the semimembranosus and the tendon of origin of the medial head of the gastrocnemius. Popliteal cysts may communicate with the knee joint and therefore may be caused by chronic knee joint pathology that

Chondromalacia patella can be diagnosed and graded noninvasively by MRI (40). In stage I, the posterior patellar articular cartilage demonstrates local areas of cartilage swelling with decreased signal intensity on both T1- and T2-weighted images. Stage II is characterized by irregularity of the patellar articular cartilage with areas of thinning. Stage III demonstrates complete absence of the articular cartilage with synovial fluid extending through this cartilaginous ulcer to the subchondral bone (Fig. 6-30).

Popliteal (i.e., Baker's) cysts and other synovial cysts about the knee appear hyperintense on T2-weighted images (Fig. 6-31A–D). They can be visualized on axial, sagittal, or coronal images. Popliteal cysts are usually an enlargement of the semimembranosus-gastrocnemius bursa, which is located between the tendon of insertion of the semimembranosus and the tendon of origin of the medial head of the gastrocnemius. Popliteal cysts may communicate with the knee joint and therefore may be caused by chronic knee joint pathology that

Chondromalacia patella can be diagnosed and graded noninvasively by MRI (40). In stage I, the posterior patellar articular cartilage demonstrates local areas of cartilage swelling with decreased signal intensity on both T1- and T2-weighted images. Stage II is characterized by irregularity of the patellar articular cartilage with areas of thinning. Stage III demonstrates complete absence of the articular cartilage with synovial fluid extending through this cartilaginous ulcer to the subchondral bone (Fig. 6-30).
LIGAMENT INJURIES

Previously, arthrography and tenography were the primary means of imaging ankle ligament injuries. They had the limitations of being invasive, providing only an indirect depiction of ankle ligament disruption, and yielding potentially false-negative results. MRI provides a noninvasive means of directly imaging all the ligaments in the vicinity of the ankle as well as all the other osseous and soft tissues.

Axial MR images provide good visualization of the tibiofibular ligaments of the tibiofibular mortise. All the lateral collateral ligaments of the ankle have an oblique orientation, and to image these ligaments in full length, either an oblique imaging plane that parallels their length must be chosen or the foot must be placed in sufficient dorsiflexion or plantar flexion to bring the ligaments into one of the standard imaging planes. With the imaging plane parallel to the anterior talofibular ligament, it is displayed as a low-signal-intensity band extending anteromedially from the lateral malleolus to gain attachment to the talus just anterior to its fibular articular surface (Fig. 6-33A) (50). The calcaneofibular ligament is visualized as a low-signal-intensity structure extending from the lateral malleolus to the calcaneus with the peroneus longus and brevis tendons situated superficial to its fibular end (Fig. 6-33B). The posterior talofibular ligament is visualized as a wide low-signal-intensity structure extending from the deep surface of the lateral malleolus to a broad attachment on the talus from its fibular articular surface to its posterior process (Fig. 6-33C).

MRI of ankle ligament injuries offers promise for the noninvasive evaluation of the site and severity of both acute ankle ligament injuries and chronic ankle instability (51).

The mechanism of injury of the lateral collateral ligaments typically involves plantar flexion and inversion, and they are usually injured in a predictable sequence from anterior to posterior. The anterior talofibular ligament is the most commonly produces effusion. A previously undescribed bursa is now known to be consistently present between the tibial collateral ligament and a major slip of the semimembranosus tendon that extends beneath it, and may serve to clarify many cases of previously unexplained medial knee pain (46). Inflammation of this bursa is well demonstrated by MRI (Fig. 6-32A,B).

Ankle

MRI is valuable as a screening modality for assessing a variety of painful ankle disorders (47–51).

FIGURE 6-29. Osteochondral lesion (previously osteochondritis dissecans). T2-weighted fat suppressed (A) and PD (B) sequences. A: There is a hyperintense T2 signal at the interface of the OC lesion and the adjacent cortex compatible with an unstable fragment. B: The lesion is well demarcated by a hypointense rim (short arrows).

FIGURE 6-30. Grade IV chondromalacia. Sagittal PD sequence with grade IV chondromalacia. There is a full thickness defect (arrow) with subchondral bone sclerosis and early subchondral cyst formation within the proximal patella.
CHAPTER 6 | IMAGING TECHNIQUES RELATIVE TO REHABILITATION

FIGURE 6-31. Baker’s cyst. A: A T1-weighted axial MRI demonstrates a hypointense Baker’s cyst (arrowheads) in the interval between the semimembranosus (SM) and the medial head of the gastrocnemius (MG). T1-weighted (B) and T2-weighted (C) sagittal MR images through Baker’s cyst (arrowheads). Note that the hypointense fluid in the cyst in the T1-weighted image becomes hyperintense on the T2-weighted image. D: A coronal T1-weighted image locates the cyst between the SM and the MG.

injured, followed in sequence by injury to the calcaneofibular and posterior talofibular ligaments. The major MRI finding in a complete rupture (i.e., grade III sprain) of the anterior talofibular ligament is a complete discontinuity of the ligament visualized at all imaging levels (Fig. 6-34A,B). This is accompanied by periarticular edema or hemorrhage and joint effusion because this ligament is a thickening of the ankle joint capsule. The edema and effusion are visualized with moderate signal intensity on T1-weighted MR images and hyperintensity on T2-weighted images. A partial tear (i.e., grade II sprain) of the anterior talofibular ligament is visualized on MRI as a discontinuity of the upper part of the ligament, with the lower portion remaining intact. Again, there is periarticular edema, hemorrhage, and joint effusion. Grade II sprains of the
FIGURE 6-32. Axial (A) and coronal (B) T2-weighted PD fat suppressed sequences. A: There is minimal fluid anterior to the semimembranosus (long arrow). B: The fluid is deep to the semitendinosus (St) and superficial to the meniscocapsular junction of the medial meniscus (long arrow in B).

calcaneofibular ligament may appear as a longitudinal splitting or waviness of the ligament with fluid accumulation within the tendon sheath of the overlying peroneal tendons (Fig. 6-35).

In contrast to the three discrete lateral collateral ligaments, the medial collateral or deltoid ligament is a continuous ligamentous sheet with an apical attachment to the tibial

FIGURE 6-33. T1-weighted images of the normal lateral collateral ligaments of the ankle. A: The anterior talofibular ligament (ATAF) extends from the fibular malleolus to the neck of the talus. B: The calcaneofibular ligament (CFL) attaches to the calcaneus and is deep to the peroneal tendons. C: The strong talofibular ligament (between arrowheads).
malleolus, and a broad base attaching below to the navicular, talar neck, spring ligament, sustentaculum tali of the calcaneus, and posterior talus. The posterior tibiotalar part of the deltoid ligament is its thickest and strongest (52). The deltoid ligament can be visualized by either axial or coronal MRI. Axial images allow simultaneous visualization of all parts of the deltoid ligament, the overlying flexor retinaculum, and the walls and contents of the tarsal tunnel (Fig. 6-36A). The contents of the four compartments under the flexor retinaculum include, from anterior to posterior, the tibialis posterior tendon, flexor digitorum longus tendon, posterior tibial artery, tibial nerve, and flexor hallucis longus tendon. Coronal MR images through the deltoid ligament display the proximal and distal attachments of each part of the deltoid ligament (Fig. 6-36B).

MRI has the potential to visualize even grade I sprains, which are microtears confined to the interior of the ligament. The minute foci of edema and hemorrhage accompanying such tears become hyperintense on T2-weighted images. Findings compatible with such grade I tears have been identified in the posterior tibiotalar portion of the deltoid ligament. They are frequently accompanied by fluid within the tendon sheath of the overlying tibialis posterior.

In chronic ankle instability, MR images show thinned, lengthened, wavy ligaments in some locations and thickened, scarred ligaments in others.

**FIGURE 6-34.** Complete rupture of the anterior talofibular ligament. Axial T1-weighted (A) and T2-weighted fat suppressed sequences (B). There is discontinuity to the ligament fibers (arrows) and associated soft tissue swelling (arrowhead in B).

**FIGURE 6-35.** Partial tear to the calcaneofibular ligament. There is thickening and increased signal within the ligament fibers (long arrow) and edema within the soft tissues between the ligament and the calcaneus.
MRI is superior to any other modality in displaying tendon pathology (47,48). In tenosynovitis, MRI detects fluid within the tendon sheath as having moderate signal intensity on T1-weighted images and as hyperintense on T2-weighted images. Tendinosis is commonly observed in the Achilles, tibialis posterior, flexor hallucis longus, tibialis anterior, and peroneal tendons (Fig. 6-37). Tendinosis is visualized as a focal or diffuse thickening of the tendon that may show areas of

OTHER ANKLE ABNORMALITIES

Technetium-99 scintigraphy has been valuable for detecting stress fractures of metatarsal and tarsal bones, and CT has high accuracy for detecting osteochondral fracture. In foot pain of undetermined etiology, however, MRI is an excellent screening modality because it permits direct evaluation of all osseous and soft-tissue structures.

MRI is superior to any other modality in displaying tendon pathology (47,48). In tenosynovitis, MRI detects fluid within the tendon sheath as having moderate signal intensity on T1-weighted images and as hyperintense on T2-weighted images. Tendinosis is commonly observed in the Achilles, tibialis posterior, flexor hallucis longus, tibialis anterior, and peroneal tendons (Fig. 6-37). Tendinosis is visualized as a focal or diffuse thickening of the tendon that may show areas of

OTHER ANKLE ABNORMALITIES

Technetium-99 scintigraphy has been valuable for detecting stress fractures of metatarsal and tarsal bones, and CT has high accuracy for detecting osteochondral fracture. In foot pain of undetermined etiology, however, MRI is an excellent screening modality because it permits direct evaluation of all osseous and soft-tissue structures.

MRI is superior to any other modality in displaying tendon pathology (47,48). In tenosynovitis, MRI detects fluid within the tendon sheath as having moderate signal intensity on T1-weighted images and as hyperintense on T2-weighted images. Tendinosis is commonly observed in the Achilles, tibialis posterior, flexor hallucis longus, tibialis anterior, and peroneal tendons (Fig. 6-37). Tendinosis is visualized as a focal or diffuse thickening of the tendon that may show areas of
Osteoarthritis

Osteoarthritis or degenerative joint disease (DJD) is an asymmetric, usually bilateral mechanical degenerative process that involves joints significantly involved in weight bearing, such as the hip, knee, and spine, and those involved in frequent repetitive mechanical trauma, such as the distal interphalangeal joints of the fingers, trapezius–first metacarpal joint, trapezius–scaphoid joint, and metatarsophalangeal joint of the great toe. It is the most common arthritis, and it is estimated that 80% of the population with more than 50 years will show radiographic evidence of osteoarthritis. The most common radiographic findings include the following:

- Nonuniform loss of joint space caused by cartilage degeneration in high load areas (e.g., the superior aspect of the hip and medial knee).
- Sclerosis of the subchondral bone.
- Osteophyte formation at the margins of the articular surfaces.
- Cystlike rarefactions in the subchondral bone that may collapse to produce marked joint deformities.
- Adjacent soft-tissue swelling (e.g., that which occurs with Heberden’s nodes of the distal interphalangeal joints of the fingers) (Fig. 6-39) (53).

Rheumatoid Arthritis

Rheumatoid arthritis is a connective tissue disorder of unknown etiology that can affect any synovial joint in the body. It is a bilaterally symmetric inflammatory degenerative disease that involves the following joints in order of decreasing frequency:

- Small joints of the hands and feet, with the exception of the distal interphalangeal joints
- Knees
- Hips
- Cervical spine
- Shoulders
- Elbows

COMMON ARTHRITIDES

The usual mode for diagnosing the arthritides is plain film radiography. However, MRI is a very useful tool particularly for the assessment of acute inflammatory arthritides. Only the findings of the more common arthropathies will be described.
Seronegative Spondyloarthropathies

These disorders are all linked to the human leukocyte antigen (HLA)-B27 histocompatibility antigen. These groups of diseases include ankylosing spondylitis, inflammatory bowel disease, psoriatic arthritis, and reactive arthritis. They are characterized by osseous ankylosis, proliferative new bone formation (syndesmophytes), and predominantly axial (spinal and sacroiliac) involvement.

Gout

Gout is a metabolic disorder that most commonly involves the feet, especially the first metatarsophalangeal joint, as well as the

The major radiographic findings include the following:

Symmetric periarticular soft-tissue swelling
Juxta-articular osteoporosis proceeding to diffuse osteoporosis
Erosions of the intracapsular portions of the articulating bones not covered by cartilage, which can proceed to severe subchondral bone erosion
Uniform joint space narrowing
Synovial cysts (e.g., Baker’s cysts behind the knee)
Subluxations (e.g., boutonniere or swan-neck deformities of the fingers, and palmar and ulnar subluxation of the proximal phalanges on the metacarpal heads) (Fig. 6-40A,B) (54)
CHAPTER 6 | IMAGING TECHNIQUES RELATIVE TO REHABILITATION

ankles, knees, hands, and elbows in asymmetric fashion. It is produced by a deposition of monosodium urate crystals in tissues with a poor blood supply, such as cartilage, tendon sheaths, and bursae. The radiographic features of gout typically do not appear until after 4 to 6 years of episodic arthritis. Radiographic features characteristic of gout include the following:

- Tophi or periarticular soft-tissue nodules/masses created by the deposition of urate crystals that may contain calcium.
- Tophi-induced periarticular or intra-articular bone erosion.
- Prominent cortical edges overhanging the tophi and well-defined bone erosions (with sclerotic margins) (Fig. 6-41) (54).
- Random distribution, without marked osteoporosis.

**Calcium Pyrophosphate Dehydrate Deposition Disease**

It is also known as pseudogout and has the classic triad of pain, cartilage calcification, and joint destruction. Chondrocalcinosis at the knee, wrist, or symphysis pubis is virtually diagnostic of calcium pyrophosphate dehydrate deposition disease (CPPD) (Fig. 6-42).

**Diffuse Idiopathic Skeletal Hyperostosis**

Diffuse idiopathic skeletal hyperostosis (DISH) is not really an arthropathy because it spares synovium, articular cartilage, and articular osseous surfaces. It is a fairly common ossification process involving ligamentous and tendinous attachments to bones and occurs in 12% of the elderly (55). It most commonly affects the thoracic spine but also may involve the pelvis, foot, knee, and elbow. It can involve ossification of all the ligaments surrounding the vertebral bodies, particularly the anterior longitudinal ligament. Ossification of the posterior longitudinal ligament (OPLL) can also be seen. This is reported to be more common in orientals and can be responsible for significant spinal canal stenosis. By definition, DISH must involve a flowing ossification of at least four contiguous vertebral bodies (Fig. 6-43A,B). There must be normal disc spaces and facet joints, without joint sclerosis.

**SPINE AND SPINAL CORD IMAGING**

It is clear, after multiple research studies to assess the usefulness of imaging in low back pain, that uncomplicated acute low back pain is a benign, self-limited condition that does not warrant any imaging studies. The vast majority of patients are back to their usual activities within 30 days. Radiographic evaluation of the lumbar spine includes frontal and lateral radiographs. These are indicated in the evaluation of back pain and weight loss, after mild trauma in patients older than 50 y/o, in patients with unexplained fever, immunosuppression, history of cancer, prolonged use of steroids and focal neurologic, or disabling symptoms. Oblique views are useful for the assessment of defects to the pars interarticularis when suspecting spondylolysis and for the evaluation of the nerve root foramina. The relative radiation dose level for a routine
uncovertebral joints (i.e., Luschka) are well depicted with coronal reformatted and sagittal oblique images. Sagittal oblique images through the lumbar spine provide excellent anatomical reconstruction of the pars interarticularis for the assessment of spondylolysis. Soft-tissue windows typically permit visualization of the moderate radiodensity of the soft-tissue structures, such as the intervertebral disc, ligamentum flavum, and thecal sac. Sagittal images provide assessment of anatomical alignment, intersegmental instability and allows for adequate evaluation of foraminal stenosis. The epidural fat contains the internal vertebral venous plexus, which can be enhanced by a circulating bolus of contrast material to improve visualization of soft-tissue encroachments into the spinal canal, such as herniated discs. Introduction of contrast material into the subarachnoid space (i.e., CT myelography) delimits the contained spinal cord and the nerve roots (Fig. 6-46).

The introduction of MDCT technology allows for the acquisitions of multiple thin cut images in the axial plane that can be reconstructed in the sagittal and the coronal planes that allows for better depiction of pathologic processes and hardware evaluation in the post–operative spine (57). MRI permits noninvasive visualization of the spinal cord and subarachnoid space within the spinal canal and the nerve roots within the neural foramina. Discrimination of these structures by CT requires injection of intrathecal contrast agents. MRI has a superior ability to evaluate intramedullary abnormalities. It also offers direct multiplanar imaging capabilities.

Axial CT images of the normal cervical and lumbar spine (Figs. 6-44 and 6-45) provide good visualization of all osseous elements, including the facet joints. In the cervical spine, the uncovertebral joints (i.e., Luschka) are well depicted with coronal reformatted and sagittal oblique images. Sagittal oblique images through the lumbar spine provide excellent anatomical reconstruction of the pars interarticularis for the assessment of spondylolysis. Soft-tissue windows typically permit visualization of the moderate radiodensity of the soft-tissue structures, such as the intervertebral disc, ligamentum flavum, and thecal sac. Sagittal images provide assessment of anatomical alignment, intersegmental instability and allows for adequate evaluation of foraminal stenosis. The epidural fat contains the internal vertebral venous plexus, which can be enhanced by a circulating bolus of contrast material to improve visualization of soft-tissue encroachments into the spinal canal, such as herniated discs. Introduction of contrast material into the subarachnoid space (i.e., CT myelography) delimits the contained spinal cord and the nerve roots (Fig. 6-46).

Degenerative Spine Disorders

CT and MRI provide complementary information about degenerative diseases of the spine. MRI is often the modality of choice in assessing degenerative changes within the spine due to its superior soft-tissue contrast. CT has superior spatial resolution and provides better conspicuity of osseous and calcified structures. The advent of MDCT technology allows for superb reconstruction in the sagittal and the coronal planes that allows for adequate evaluation of the contained spinal cord and the nerve roots within the neural foramina. Discrimination of these structures by CT requires injection of intrathecal contrast agents. MRI has a superior ability to evaluate intramedullary abnormalities. It also offers direct multiplanar imaging capabilities.

Axial CT images of the normal cervical and lumbar spine (Figs. 6-44 and 6-45) provide good visualization of all osseous elements, including the facet joints. In the cervical spine, the uncovertebral joints (i.e., Luschka) are well depicted with coronal reformatted and sagittal oblique images. Sagittal oblique images through the lumbar spine provide excellent anatomical reconstruction of the pars interarticularis for the assessment of spondylolysis. Soft-tissue windows typically permit visualization of the moderate radiodensity of the soft-tissue structures, such as the intervertebral disc, ligamentum flavum, and thecal sac. Sagittal images provide assessment of anatomical alignment, intersegmental instability and allows for adequate evaluation of foraminal stenosis. The epidural fat contains the internal vertebral venous plexus, which can be enhanced by a circulating bolus of contrast material to improve visualization of soft-tissue encroachments into the spinal canal, such as herniated discs. Introduction of contrast material into the subarachnoid space (i.e., CT myelography) delimits the contained spinal cord and the nerve roots (Fig. 6-46).

The introduction of MDCT technology allows for the acquisitions of multiple thin cut images in the axial plane that can be reconstructed in the sagittal and the coronal planes. The high spatial resolution of the acquired data allows for near–perfect isometric reconstruction in different planes. In addition, computer-generated volume rendering images provide superb 3D images of the spine (Fig. 6-47A–C).

Sagittal T1-weighted MR images of the cervical, thoracic, or lumbar spine provide excellent noninvasive survey to evaluate patients with suspected regional spinal pathology. Midsagittal T1-weighted images display the high–signal-intensity marrow of the vertebrae bordered by low–signal-intensity cortical
FIGURE 6-44. MDCT of the normal cervical spine. A: Bone window CT of the cervical spine in the axial plane at the level of C5-6 displays normal facet joints (short arrow) and Luschka joints (long arrow). B: Coronal multiplanar reformatted images (MPR) through the level of the uncovertebral joints (arrow). C: Midplane sagittal reconstructed images of the cervical spine shows mild reverse of the normal lordosis that could be secondary to spasm or associated to positioning. Note adequate alignment and no intersegmental subluxations. D: Sagittal oblique MPR demonstrates the nerve root foramina, uncovertebral joints, and facet joints.

Degenerative Disc Disease

The intervertebral disc space is a cartilaginous joint with a central nucleus pulposus surrounded by an annulus fibrosus. Degenerative change in the nucleus pulposus is termed intervertebral osteochondrosis (Fig. 6-49A,B). Early signs of disc disease may include loss of fluid signal within the nucleus pulposus (dehydration), which results in decreased signal within the central portion of the disc on T2-weighted images, and blurring of the transition between the nucleus pulposus and the annulus fibrosus. This is followed later by narrowing of the intervertebral disc, which sometimes results in vertebral body end-plate degenerative changes. These end-plate changes

bone. Structures displaying very low signal intensity include the peripheral part of the annulus fibrosus of the intervertebral disc, all ligaments, the dura, and the cerebrospinal fluid (CSF), and these are usually indistinguishable from each other (Fig. 6-48A–C). The nucleus pulposus, and probably the inner portion of the annulus fibrosus, shows moderate signal intensity. The spinal cord and the nerve roots display moderate signal intensity, which is well contrasted against the low–signal-intensity CSF. Collections of epidural fat, which are largest at lumbar levels, produce high signal intensity on T1- and T2-weighted images. On T2-weighted MR images, CSF and the normal well-hydrated nucleus pulposus assume high signal intensity.
FIGURE 6-45. Normal lumbar spine CT. A: Axial soft-tissue window CT of L5 demonstrating thecal sac (T), L5 nerve root (NR) within the lateral recess, epidural fat (E), and ligamentum flavum (LF). B: Bone windows axial through the same level demonstrate the sclerotic margins of the facet joint and the right sacroiliac joint. C: Midplane sagittal reconstruction (MPR image). Normal homogeneous trabecular pattern. Note intervertebral space distance and anatomical alignment of the anterior and posterior surface to the vertebral bodies. D: The density of the nucleus pulposus (NP) and annulus fibrosus (AF) can be appreciated. Note decreased attenuation to the CSF in the thecal sac. E: Oblique reconstruction through the posterior elements demonstrating the normal facet joints and pars interarticularis.
are referred to as Modic changes (58). Modic type I changes are characterized by signal of edema, that is decreased signal on T1-weighted images and increased signal on T2-weighted images. Type II changes (Fig. 6-50A,B) follow the signal characteristic of fat, with intermediate to increased signal on T1-weighted images and increased signal on T2-weighted images. This is the most common type of reactive end-plate changes appreciated on degenerative osteochondrosis. Type III changes represent osseous sclerosis, characterized by decreased signal on both the T1- and the T2-weighted images.

When the disc becomes degenerated, it may undergo tearing of the annulus fibrosus collagen bundles, which is often a precursor to herniation of the nucleus pulposus, particularly at the posterolateral aspect of the disc. Protrusion is herniation of the nucleus pulposus that is contained by the annulus. In the axial plane, it usually demonstrates a base broader than the height against the parent disc (59). Extrusion is defined as a herniation of the nucleus pulposus beyond the fibers of the

FIGURE 6-46. Metrizamide CT myelogram at L1 level delimiting the spinal cord (SC), nerve roots (NR) arising from the cord, and the contrast-enhanced CSF.

FIGURE 6-47. Volume rendering images of the normal lumbar spine. A: Coronal (frontal) projection. Arrow points to a hypoplastic right transverse process, a normal anatomical variant. B: Sagittal (lateral) view, note relationship of pars interarticularis (PI) to the facet joint (FJ) and root foramina (RF). C: Midplane sagittal view through the central canal. Facet joint (FJ), pars interarticularis (PI), and pedicle (P).
FIGURE 6-48. MRI of the normal spine and spinal cord. **A:** A T2-weighted axial MRI at the L5-S1 disk level shows the thecal sac (T) as an area of high signal intensity, the nerve roots (S1, S2) within the posterior thecal sac are of low signal intensity. Note the ligamentum flavum (LF), and fluid within the facet joints (FJ) with high–signal-intensity fluid. **B:** T1-weighted midsagittal MRI of the lumbosacral spine displays the nucleus pulposus (NP) of the intervertebral discs, spinal cord (SC), conus medullaris (CM), and nerve roots of the cauda equina (CE) as areas of moderate signal intensity. Cerebrospinal fluid (CSF) is of low signal intensity, and epidural fat (E) is hyperintense. **C:** T2-weighted midsagittal MRI of the lumbosacral spine demonstrates the increased intensity of the nucleus pulposus (NP) and CSF.

Peripheral annulus. In the axial plane, it can demonstrate a narrower base in relationship to the height of the herniation. A sequestered or free fragment is an extruded disc without contiguity to the parent disc. The majority of disc herniation are central or paracentral (subarticular) in location.

The loss of the load-diffusing function of the normal disc also causes facet joint osteoarthritis and marginal osteophytosis of the vertebral body ends (spondylosis deformans) by virtue of the increased loads these joints must bear.

Cervical Disc Herniation

Disc herniation is typically preceded by degenerative changes in the mucopolysaccharides of the nucleus pulposus, which produce fibrillation of the collagen (60). This eventually causes dehydration and loss of disc volume. As a result, the nucleus pulposus no longer serves as a normal load-dispersing mechanism, and excessive stress is borne by the annulus fibrosus. This produces annular fissuring and tears that can culminate in herniation of the nucleus pulposus. The loss of the load-diffusing function of the normal disc also causes facet joint degeneration and marginal osteophytosis of the vertebral body ends by virtue of the increased loads these joints must bear.

Cervical disc herniation occurs with less frequency than lumbar disc herniation. About 90% of cervical disc herniations occur, in order of decreasing frequency, at C5-6, C6-7, and C4-5 (61,62). On CT examination, a herniated cervical disc appears as a dense soft-tissue mass protruding from the disc space centrally or paracentrally into the spinal canal or posterolaterally into the neural foramen (Fig. 6-51).

On T1-weighted MR images, the herniated cervical disc appears as a posterior extension of the moderate signal intensity of the disc into the low–signal-intensity region of the thecal sac (Fig. 6-52A,B). Because the spinal cord appears as a relatively high–signal-intensity structure outlined by the low–signal-intensity CSF, the relationship of the herniated disc to the spinal cord can be visualized directly by MRI. On T2-weighted MR images, the degenerated disc appears as a narrowed disc interval. The disc herniation appears as a moderate–to low–signal-intensity impingement on the now high–signal-intensity
CSF. The posterior margin of the herniated disc may have a very-low-signal-intensity margin interfacing with the CSF. This may be a posterior longitudinal ligament elevated by the herniated disc, or it may be fragments of the posterior part of the annulus fibrosus (61). T2-weighted images also permit evaluation of the relationship of the herniated disc to the spinal cord to determine its probability of causing a patient’s myelopathic findings (Fig. 6-53A,B). It is sometimes difficult to differentiate lateral herniations of the disc into the neural foramen from osteophytic encroachments by MRI because they may both demonstrate low signal intensity. In these circumstances, CT provides good differentiation between bone and soft-tissue density.
Cervical Spinal Stenosis and Foraminal Stenosis

Cervical spinal stenosis can be congenital or acquired. In the less common congenital stenosis, a small spinal canal is produced by short pedicles and thick laminae. It commonly remains asymptomatic until degenerative changes are superimposed on the congenital stenosis later in life.

Acquired stenosis can be produced by a host of hypertrophic degenerative changes often collectively referred to as cervical spondylosis. These include osteophytic lipping of the posterior margins of the vertebral body ends bordering the disc, hypertrophic degenerative changes involving Luschka’s joints or the facet joints, buckling or hypertrophy of the ligamenta flava, and OPLL. All these structures border the spinal canal; therefore, hypertrophic degenerative changes can produce spinal canal stenosis. Because Luschka’s joints, the facet joints, and the ligamenta flava also border the neural foramen, their involvement by degenerative processes can produce foraminal stenosis.

Although hypertrophic degenerative changes of any of the structures bordering the spinal canal or neural foramen can occur in isolation, they are commonly precipitated by intervertebral disc degeneration. As the disc degenerates and loses its normal load-dispersing ability, loads tend to become concentrated on the vertebral body margin toward which the spine is bent. This excessive loading can produce marginal osteophytes around the entire circumference of the vertebral body end plates. Those osteophytes developing on the posterior margin can encroach on the spinal canal to produce spinal stenosis (Fig. 6-54A, B). Luschka (i.e., uncovertebral) joints are situated between the uncinate processes that protrude from the lateral or posterolateral margins of the upper surface of the vertebral bodies and a reciprocal convexity on the lateral aspect of the inferior surface of the next higher vertebral body. Recent evidence indicates that they are not true joints. Rather, they are degenerative clefts within the lateral part of the intervertebral disc that begin in the second decade of life. The increased loading of Luschka joints produced by these degenerative changes produces bony spurs that can extend posteriorly into the lateral part of the spinal canal or posterolaterally into the neural foramen (Fig. 6-55A–C).

Disc degeneration is accompanied by dehydration and loss of disc height, with decreased space between vertebral bodies resulting in increased facet joint loads. The resultant...
facet joint degeneration involves cartilage erosion with joint space narrowing, subchondral bone sclerosis, and osteophyte formation. The osteophytes may encroach on the spinal canal or the neural foramen.

Loss of disc height results in decrease of the laminae interspace, which causes the ligamentum flavum to buckle and bulge into the spinal canal, contributing to the spinal stenosis. Because the ligamentum flavum continues laterally into the facet joint capsule, buckling of this part of the ligamentum flavum can cause foraminal stenosis.

OPLL occurs more commonly at cervical than at other vertebral levels. It is best visualized via CT, where it appears as an ossification extending over several vertebral levels, separated from the posterior margin of the vertebral bodies by a thin radiolucent interval (Fig. 6-56A,B).

When any of these potential causes of cervical stenosis sufficiently narrow the spinal canal, cord compression can produce myelopathic signs and symptoms. Spinal stenosis most frequently narrows the AP dimension of the spinal canal. Although the cross-sectional area of the spinal canal is smallest...
at the C4 and C7 levels, the smallest AP diameter is usually at the C3 through C5 levels (62). It has been stated that all spinal stenosis that reduces the AP dimension to less than 10 mm could produce quadriplegia (64).

Although the uppermost cervical cord segments are nearly round, at most cervical levels the cord has an elliptical outline with its major axis transversely oriented. With encroachment of the cord by spinal stenosis, it is usually first flattened anteriorly by an encroaching osteophyte. With progression, the anterior median fissure becomes indented and widened until the cord assumes a kidney bean shape (Fig. 6-57) (48). The lateral funiculi may become tapered anterolaterally because of tension on the denticulate ligaments. The cord may become notched dorsally because of posterior white column atrophy. It has been estimated that a 30% reduction in cord cross-sectional area may be required to produce signs of ascending and descending tract degeneration (65).

**Thoracic Spine Abnormalities**

Both thoracic disc herniation and thoracic spinal stenosis are rare compared with cervical and lumbar level disease. When thoracic disc herniation does occur, it most frequently involves discs below T8 (66). The CT findings are similar to those of cervical levels, except that calcification of the disc protrusion is more common at thoracic levels. The causes as well as CT and MRI findings of thoracic spinal stenosis are similar to those at cervical levels.

**Lumbar Disc Herniation**

The correlation of lumbar disc herniation with a patient’s complaints of low back pain or sciatica is not always clearly established. It has been estimated that as many as 20% of patients with radiologic findings of disc herniation are asymptomatic (67). Furthermore, when disc herniation occurs in asymptomatic patients, other findings are often present that also could explain the clinical findings.

Lumbar disc herniations occur most frequently posterolaterally because the annulus is thinnest in the posterior quadrants but reinforced in the midline by the posterior longitudinal ligament. Also, flexion is the most prevalent lumbar spine motion, which places greatest stress on the posterior part of the disc. When the disc herniates to the posterolateral direction, it frequently does not impinge on the spinal nerve roots emerging
lateral herniations occur outside of the neural foramina. When present, they usually impinge on the ventral ramus that has just emerged from that foramen.

Noncontrast CT has been described as being accurate in diagnosing disc herniations. On CT examination, the herniated disc appears as a focal protrusion of the disc that displaces the epidural fat (Fig. 6-58). The herniated disc material is typically slightly hyperdense relative to the non–contrast-enhanced

---

**FIGURE 6-56.** Ossification of the posterior longitudinal ligament OPLL. A: Axial CT at the level of C3 demonstrates the calcification to the fibers of the posterior longitudinal ligament. There is narrowing to the central canal by the mass effect exerted by the enlarged calcified ligament. B: Sagittal multi-planar reformatted image. The short arrows at the flowing calcifications within the anterior longitudinal ligament. The calcification extends from C2 up to C7. The OPLL (long arrow) extends from C2 up to the proximal border of C5.

**FIGURE 6-57.** Axial T2-weighted images in the same patient as in Figure 6-54. Note deformity to the cord that has assumed a bean shape in the transaxial plane. There is compression and increased signal intensity to the cord compatible with myelopathy (arrow).

**FIGURE 6-58.** Sagittal reconstructed image of a MDCT acquisition of the lumbar spine demonstrates a large disc protrusion.
dural sac and its adjacent nerve roots. The dural sac or adjacent nerve roots may be seen to be indented, displaced, or compressed. In more lateral herniations, the soft-tissue material of the disc can encroach on the neural foramen or the extraforaminal soft tissues, where it also displaces fat, and here it may encroach on the dorsal root ganglion, spinal nerve, or its ventral ramus. Herniated lumbar discs may calcify or contain gas. Extruded disc fragments can become separated from the disc and are thus able to migrate superiorly, inferiorly, or laterally. A herniated disc should be distinguished from a bulging annulus. A bulging annulus is produced by dehydration and volume loss within the nucleus pulposus. In contrast to the focal protrusion of a herniated disc, the bulging annulus typically has a symmetrical smooth contour, bulging beyond all margins of the vertebral body.

On T1-weighted sagittal and axial MR images, the herniated lumbar disc appears as a moderate–signal-intensity intrusion into the high–signal-intensity epidural fat or on the moderate–to-low–signal-intensity thecal sac or the lumbar nerve roots within their dural sleeves (Fig. 6-59A,B). Similarly, disc herniation into the neural foramen is visualized by a moderate–signal-intensity mass displacing the foraminal fat and encroaching on the dorsal root ganglion or nerve roots.

On T2-weighted sagittal MR images, the low signal intensity of a degenerated disc contrasts sharply with the high signal intensity of the nucleus pulposus of adjacent well-hydrated discs (Fig. 6-60). Any intrusion of the low–signal–intensity disc herniation on the thecal sac is well seen because of the high–signal–intensity myelographic effect of the CSF on T2-weighted images.

Discography remains a controversial diagnostic imaging modality. It appears that its major diagnostic value lies in the reproduction of the patient’s specific pain on contrast injection of a given disc, with controls demonstrating that injection of adjacent discs produces either no pain or foreign pain (68). Discography, especially when combined with CT, may provide information about degeneration and the extent of fissures and rupture.

**Lumbar Spinal Stenosis and Foraminal Stenosis**

Like cervical stenosis, lumbar spinal stenosis is frequently precipitated by disc degeneration with subsequent marginal osteophytosis of the vertebral body ends, hypertrophic degeneration of the facet joints, and bulging of the ligamenta flava. Lumbar stenosis may be lateral, central, or combined. The lower lumbar vertebrae normally have shorter pedicles that cause the superior articular processes to intrude into the spinal canal.
to cut off narrow lateral recesses (Fig. 6-61A–E). The lateral recesses are bordered by the pedicles laterally, the vertebral body anteriorly, and, most important, the superior articular processes posteriorly. The lateral recesses are occupied by the nerve roots of the next spinal nerve to exit as they descend within their dural sleeve. Osteophytes that develop on the anteromedial margin of the superior articular processes of the next lower vertebra are most likely to encroach on the lateral

FIGURE 6-61. MDCT of the lumbar spine in a patient with degenerative spondylosis and spinal canal stenosis. At the L4-5 level, there are hypertrophic degenerative changes of the facet joints and ligamentum flavum (long white arrow), which produce lateral stenosis, and hypertrophic changes of the articular process, which cause central stenosis (short white arrow). A: Bone windows. B: Soft-tissue windows that demonstrate to a better advantage the hypertrophied ligamentum flavum (short arrow). C: Soft-tissue windows and D: bone windows. There is congenital canal stenosis as well. Compression to the superior end plate of L1 is appreciated. E: Volume rendering reconstruction looking at the central canal demonstrates the significant narrowing to the central canal.
PART I | PRINCIPLES OF ASSESSMENT AND EVALUATION

recess to produce lateral stenosis. Because the inferior articular processes of the next higher vertebra are situated posteromedial to the superior articular processes, osteophytes developing on their anterior margin are more likely to produce central stenosis. In central stenosis, any or all of the rootlets of the cauda equina can be encroached on. Vertebral body margin osteophytes and buckling of the ligamentum flavum can contribute to lumbar spinal stenosis.

Hypertrophic degenerative changes involving the facet joints can also encroach on the posterior aspect of the neural foramen and produce foraminal stenosis with compression of the nerve roots exiting that foramen. Therefore, hypertrophic degenerative changes involving a single superior articular process can involve the roots of two closely adjacent nerves, with the possibility of producing both foraminal and lateral spinal stenosis. With disc degeneration and loss of disc height, the neural foramen can be further compromised by the upward and forward displacement of the superior articular process into the upper part of the neural foramen, where the nerve roots are situated. In addition, because of the obliquity of the facet joint, the accompanying downward displacement of the inferior articular process of the next higher vertebra can produce retrolisthesis (i.e., backward displacement) of its vertebral body into the upper portion of the neural foramen.

By standard myelography, the protruding disc anteriorly and the bulging ligamenta flava posteriorly can produce an hourglass appearance of the thecal sac (Fig. 6-62). By CT, all osteophytes are clearly visualized, and measurements of the AP dimension of lateral recesses that are less than 3 mm are strongly suggestive of lateral stenosis (69). The hypertrophic changes producing central and foraminal stenosis are also well visualized. Sagittal reformations are especially helpful in evaluating foraminal stenosis.

Facet anatomy is well seen by MRI, with subchondral bone appearing as a signal void. On T1-weighted images, articular cartilage is visualized as a moderate–signal-intensity interval between the subchondral bone of the two articular processes. This becomes more signal intense on T2-weighted images. Facet joint degeneration appears as an irregularity or reduction in the thickness of the articular cartilage. Osteophytes are usually displayed as signal voids encroaching into the foramen, lateral recess, or spinal canal. Occasionally, osteophytes show high–signal-intensity interior, indicating the presence of marrow.

Spondylolysis and Spondylolisthesis

Spondylolysis is a defect in the pars interarticularis, commonly involving the L5 and occasionally the L4 vertebrae. Most spondylolysis is thought to be produced by repetitive stress. The gravitational and muscular loads acting across the steep incline of the upper surface of the sacrum can be resolved into a shearing component, which tends to displace the L5 vertebral body forward on S1, and a compressive component at right angles to the superior surface of S1 (Fig. 6-63). In accordance with Newton’s third law, S1 will exert an equal and opposite force against the inferior aspect of the L5 vertebral body. The tendency of L5 to be displaced forward on S1 is primarily resisted by the impaction of the inferior articular processes of L5 on the superior articular processes of S1. Again, Newton’s third
law dictates that there will be an equal and opposite force exerted against the inferior articular process of L5. The upward and forward force of the sacral body on the L5 body and the upward and backward force of the superior articular process of the sacrum on the inferior articular process of L5 cause shearing stresses to be concentrated on the pars interarticularis, and this can produce a stress fracture.

Spondylolisthesis is an anterior subluxation of one vertebral body on another. It can occur at any vertebral level, but the mechanics of the lumbosacral junction cause a higher incidence at this level. The most common cause at this level is spondylolysis, where the impaction of the inferior articular process of L5 or L4 will no longer be able to resist forward displacement of the vertebral body. Whether or not a spondylolisthesis follows a spondylolysis is largely determined by the resistance of the other supporting structures of the lumbosacral junction, which include the intervertebral disc, the anterior longitudinal ligament, and the iliolumbar ligaments. When they fail, the lysis becomes a listhesis.

Other causes of spondylolisthesis include degenerative changes in the facet joints and disc that produce joint instability, fractures, dysplasia of the upper sacrum or the neural arch of L5, generalized pathology such as Paget’s disease, or iatrogenically induced laminectomy or facetectomy (62).

On oblique plain films, spondylolysis is visualized as a break in the neck of the “Scotty-dog” outline, which is produced by the ipsilateral transverse process forming a nose; the ipsilateral pedicle, an eye; the pars interarticularis, a neck; the ipsilateral inferior articular process, a forelimb; the lamina, a body; the contralateral inferior articular process, a hind limb; and the spinous process, a tail (Fig. 6-64). When spondylolysis is suspected clinically or on plain film studies, a volumetric CT with sagittal reconstruction, MR, or plain film examination with nuclear medicine scintigraphy can be useful for diagnosis (Fig. 6-65A,B) (70).

Spondylolisthesis is graded by the amount of subluxation, with grade I being a forward displacement of less than 25%; grade II, a forward displacement of 25% to 50%; grade III, a displacement of 50% to 75%; and grade IV, a displacement of greater than 75%. Grading the spondylolisthesis is usually accomplished by lateral plain films or sagittal reconstructed images in CT (Fig. 6-65C,D).

On CT, the defect of spondylolysis is differentiated from the facet joint interval by its location at the axial level of the pedicles rather than at the level of the neural foramen, as well as by the defect’s irregular margins and adjacent sclerosis. By MRI, the defect in the pars is visualized as a low-signal-intensity zone within the high-signal-intensity marrow of the pars.

Spinal Trauma

Although much spinal trauma is well visualized on plain films, CT has a number of advantages over this modality. These include the demonstration of fractures not seen in plain films, an accurate determination of the amount of spinal canal encroachment by fracture fragments (Fig. 6-66A,B), the identification of neural foramen impingement by fractures involving its boundaries, and a more precise evaluation of facet disruption.

MR can display impingement on the dural sac or the spinal cord by bone fragments, as well as any resultant cord contusion (Fig. 6-66C). It can demonstrate acute cord enlargement as a sign of cord edema or hemorrhage and cord atrophy. CT myelography can be used to diagnose post-traumatic cystic myelopathy because the cyst will take up the contrast and be displayed as a well-marginated, homogeneous, high-density region within the cord. When MRI is not available or contraindicated, CT myelography can be used to assess the degree of spinal canal stenosis and cord compression.

CT also can augment the interpretation of the signs of vertebral instability seen on plain radiographs (Fig. 6-67A–C) (71). These signs include the following:

Vertebral displacements involving the whole vertebra or fracture fragments.

Widening of the interspinous interval, which implies injury to the posterior spinal ligaments secondary to hyperflexion injury.

Increased dimensions of the vertebral canal in the sagittal or coronal plane often evaluated by an increased interpedicle distance, which implies a complete disruption of the vertebral body in the sagittal plane.

Widening of the facet joint interval, which implies ligamentous disruption.
Disruption of the alignment of the posterior aspect of the vertebral bodies, such as occurs in burst fractures or seatbelt fractures (Fig. 6-68).

T1-weighted sagittal and axial MR images provide the best evaluation of vertebral alignment and the bony and ligamentous boundaries of the spinal canal. They also allow the best delineation of the low signal intensity of a traumatic syringomyelia against the higher signal intensity of the surrounding spinal cord. T2-weighted sagittal MR images that produce a high–signal-intensity CSF provide the best estimate of the degree of encroachment of a bony fragment on the thecal sac or the spinal cord.

MRI has a number of advantages over other modalities for imaging spinal trauma. First, it permits evaluation of vertebral alignment at the cervicothoracic junction of the spine, which is relatively inaccessible by other modalities. Second, it provides a means to evaluate adjacent soft-tissue damage. For example, hemorrhage in the prevertebral space that can occur with hyperextension injuries is identified on T2-weighted images as a high–signal-intensity area. MRI also identifies high–signal-intensity hemorrhage in the posterior paraspinal muscles that can occur secondary to hyperflexion injuries. In addition, MRI is the most sensitive modality for the assessment of ligamentous injuries as it detects edema within the supporting ligaments, a finding not assessed by any other imaging modality (Figs. 6-66 and 6-67C). Of importance is the fact that MRI provides a noninvasive means of evaluating the relationship of retropulsed vertebral body fragments or anteriorly displaced neural arch fragments to the spinal cord (Fig. 6-66C). In most centers, MRI has replaced myelography.

**FIGURE 6-65.** Spondylolysis and spondylolisthesis. A: Right parasagittal reconstruction of an MDCT demonstrates a spondyloytic defect in the pars interarticularis of L5 (PD arrow). B: Parasagittal reconstructed image form a volumetric CT demonstrating the pars defect in a patient with stage II spondylolisthesis (PD arrow). C: Midplane sagittal reconstruction demonstrates the grade II spondylolisthesis. D: Volume rendering, oblique posterior view demonstrating the pars defect (PD) and a normal pars (NP) above.
as the procedure of choice for evaluating the effects of vertebral trauma on the spinal cord. Most important, MRI can evaluate the extent and type of spinal cord injury (71,72).

An acutely injured spinal cord tends to enlarge, thereby filling the spinal canal and displacing the epidural fat. This can be visualized by both CT and MRI. However, MRI provides the best means of evaluating the type of spinal cord trauma and its evolution. MRI is valuable in the early stages of spinal cord injury in determining the type of spinal cord injury and the prognosis for recovery. It can identify the level and completeness of cord transection by direct visualization of the transection site. In the nontransected cord, it can discriminate cord hemorrhage from cord contusion with edema. Spinal cord contusion with edema causes high signal intensity on T2-weighted images within the first 24 hours of injury. Acute hemorrhage of less than 24 hours’ duration appears as a low-signal-intensity area on T2-weighted images. Within a few days of the trauma, the subacute hemorrhage site becomes hyperintense on T2-weighted images as a result of the accumulation of paramagnetic methemoglobin (Fig. 6-69). Kulkarni et al. found that the type of injury visualized by MRI correlated with the patient’s recovery of neurologic function (73). Those

FIGURE 6-66.  MDCT visualization of spinal trauma. A: CT of a burst fracture of L1 shows displacement of a fracture fragment into the spinal canal (arrow). There is significant comminution to the superior articular surface of the vertebral body. B: Sagittal MRI reconstruction demonstrates the large fragment (arrow) extending into the spinal canal. There is mild compression of the superior end plate of L2. C: Sagittal T2-weighted MR image demonstrates increased signal intensity to the bone marrow of the L1 and L2 vertebral bodies due to the presence of edema and hemorrhage. There is posterior displacement of the conus and proximal cauda equina in the spinal canal. Increased signal to the conus (arrow) is compatible with a cord contusion.
It is thought that the myelomalacic areas become larger intramedullary cysts because the scar tissue about the injured cord tethers the cord to the dura so that the episodic changes in CSF pressure that occur during daily activities tend to be concentrated on the injured cord segment as stretching forces. It is hypothesized that these stresses cause coalescence of the myelomalacic microcysts into a progressively enlarging gross cyst. CSF is theorized to enter the cysts along enlarged perivascular Virchow-Robin spaces that connect the subarachnoid space to the cyst.

patients with cord contusion and edema exhibited significant functional recovery, whereas those with hemorrhage made little functional progress. Therefore, the MRI characteristics of the injury may provide the clinician with important prognostic data.

MRI is also invaluable for identifying late sequelae of spinal cord trauma, including myelomalacia and post-traumatic spinal cord cysts or syringomyelia. Myelomalacia is thought to develop within an injured segment of the spinal cord as a result of ischemia or the release of enzymes from damaged spinal cord tissues, or both (74). The myelomalacic area is made up of the products of neuronal degeneration, scar tissue, and microcysts. It is thought that the myelomalacic areas become larger intramedullary cysts because the scar tissue about the injured cord tethers the cord to the dura so that the episodic changes in CSF pressure that occur during daily activities tend to be concentrated on the injured cord segment as stretching forces. It is hypothesized that these stresses cause coalescence of the myelomalacic microcysts into a progressively enlarging gross cyst. CSF is theorized to enter the cysts along enlarged perivascular Virchow-Robin spaces that connect the subarachnoid space to the cyst.

**FIGURE 6-67. A:** Axial image at the C5 cervical segment. There is a burst comminuted fracture to the vertebral body. A fracture through the posterior arch is present as well. **B:** Sagittal MPR reconstruction demonstrates anterior collapse and posterior subluxation in relationship to the adjacent vertebral bodies. **C:** Sagittal T2-weighted MR shows increased signal to the vertebral body compatible with edema, prevertebral soft-tissue swelling (arrow) and edema (arrow) within the interspinous spaces as a secondary sign of ligamentous injury to the posterior supporting ligaments.
stabilized may cause the patient to develop progressive sensory and motor deficits. Although myelomalacia has no definitive treatment mode, a spinal cord cyst can be surgically decompressed with a shunt to achieve improvement or at least an arrest of the patient’s neurologic deterioration. Therefore, the MRI distinction between cysts and myelomalacia is important. MRI can also be used in postoperative follow-up to ensure that the cyst has been fully decompressed and that the catheter is continuing to function to prevent reaccumulation of fluid within the cyst.

**Atlantoaxial Instability**

Spine instability is the loss of the spine motion segment stiffness, where applied force produces greater displacement than normal, causing deformity and pain. Atlantoaxial instability can be produced by softening, laxity, or rupture of the transverse atlantal, alar, and apical ligaments of the dens (i.e., odontoid). These ligaments hold the odontoid in its proper position against the anterior arch of the atlas and below the level of the foramen magnum. Such ligamentous changes can be produced by rheumatoid arthritis, Down’s syndrome, or traumatic rupture. Rheumatoid changes that destroy the articular cartilage and bone of the atlantoaxial joints can further increase the instability. Atlantoaxial instability also can be caused by odontoid abnormalities, such as an unfused apical portion of the odontoid (i.e., os odontoideum), or by odontoid fractures. Normally, the cartilaginous radiolucent interval between the anterior arch of the atlas and the dens does not exceed 3 mm in the adult. With ligamentous abnormalities, lateral radiographs of the flexed cervical spine may show a posterior subluxation of the odontoid into the spinal canal that increases the atlas-dens interval to more than 3 mm. When posterior subluxation of the odontoid exceeds 9 mm, it is likely to compromise the spinal cord and produce neurologic abnormalities (75). There also can be a superior subluxation of the odontoid above the level of the foramen magnum that can cause death by impingement on the medulla or the vertebral arteries. These subluxations are well visualized by CT and MRI. MRI also can directly evaluate the ligaments. CT myelography and MR images in the axial or sagittal planes can assess the involvement of the spinal cord or medulla by the subluxation.

**BRAIN IMAGING**

The following section will be dedicated to brain imaging relevant to rehabilitation. Emphasis is placed on the imaging of ischemic and hemorrhagic strokes, head trauma, and common degenerative diseases. The imaging of brain neoplasms and infections will not be covered in this section, as it is beyond the scope of this text.

**Stroke**

The term *stroke* refers to both transient and permanent neurologic signs and symptoms of a nontraumatic vascular etiology. Imaging of acute stroke should help (a) exclude intracranial hemorrhage, (b) differentiate between irreversibly affected brain...
tissue (‘dead brain’) and reversibly impaired tissue (‘tissue at risk’), which can benefit from early treatment, and (c) identify arterial stenosis or occlusion. Tissue at risk, or penumbra, refers to an area of reduced perfusion and loss of function, yet whose neurons are still viable. Therefore, once hemorrhage has been ruled out, timely reperfusion of this tissue with thrombolytics may prevent neuronal cell death and help re-establish normal function (76). CT continues to be the initial imaging modality for most acute stroke patients for three reasons. First, CT detects intracerebral hemorrhage with great specificity and sensitivity because freshly extravasated blood is more radiodense than either gray or white matter. Second, MRI is unable to detect the oxyhemoglobin that predominates in the hemorrhage in the early hours after a stroke because it is a nonparamagnetic substance. Third, the uncooperativeness of many acute stroke patients during the long MRI scan times and the incompatibility of critical monitoring equipment with strong magnetic fields often preclude early MRI examination. Current stroke diagnostic protocols follow a multimodal approach that can include non–contrast-enhanced CT, CT perfusion, CT angiography (CTA), conventional MRI, MR angiography (MRA), and diffusion- and perfusion-weighted MR imaging techniques, in order to establish early diagnosis and subsequently select appropriate therapy.

**ISCHEMIC STROKE**

Cerebral ischemia can be produced by thrombosis of large extracranial or small intracerebral vessels, emboli originating from atherosclerotic plaques or thrombi within more proximal vessels or the heart. In addition, decreased perfusion of systemic origin, such as shock, decreased cardiac output, or respiratory failure can also cause cerebral ischemia with or without infarction.

Cerebral ischemia can be completely or partially reversible, or irreversible leading to neuronal cell death, commonly known as infarction. Once blood flow to the brain is decreased or interrupted for a sufficient period of time, usually in a well-defined vascular territory, the chemical pumps within the neuronal cell membranes cease to function adequately disturbing the normal electrolyte homeostasis. Extracellular water subsequently rushes into the affected neuronal cells. This cascade of events initially causes neurons to halt cellular function in an attempt to survive. This “stunned” cell population is potentially salvageable with prompt reperfusion. If adequate reperfusion does not occur in a timely fashion, irreversible neuronal cell deaths will occur. Edema related to infarction involves both gray and white matter and has certain CT and MRI findings.

Nonenhanced CT is the initial study of choice in patients with suspected stroke, as it is readily available, can be performed quickly, and is highly sensitive in the detection of cerebral hemorrhage. On nonenhanced CT, edema related to a cerebral infarction appears as a hypodense or low attenuation area, which means that it appears darker than expected. Early nonenhanced CT signs of ischemic cerebral infarction in the MCA territory are as follows: (1) Obscuration of the lentiform nucleus, some-

**FIGURE 6-70.** Obscuration of the lentiform nucleus. Nonenhanced CT shows effacement of the left lentiform nucleus (arrowhead). Compare with normal lentiform nucleus on the right (arrowhead).

**FIGURE 6-71.** Insular ribbon sign. CT changes of slight hypodensity, loss of normal gray-white matter differentiation, and effacement of overlying cortical sulci in the region of the right insula (arrow).
soon as 90 minutes after the event (Fig. 6-72). It is important to note that this sign implicates occlusion and not necessarily infarction. Nevertheless, nonenhanced CT is usually negative during the first few hours after an ischemic infarct, and it is only later that areas of hypoattenuation can be identified with associated effacement of the adjacent cortical sulci (Fig. 6-73). Edema reaches its peak at 3 to 5 days, and by this time non–contrast-enhanced CT typically demonstrates a well-defined hypodense area that usually corresponds to the vascular territory of one of the cerebral arteries or its branches. With large infarctions, brain swelling can eventually lead to brain herniation or obstructive hydrocephalus, which can be life threatening (Fig. 6-74A,B). With subsequent degeneration and phagocytosis of the infarcted brain tissue, there is volume loss that causes an increase in size of the overlying cortical sulci and underlying ventricles (78). When the infarction is caused by a systemically induced general reduction in brain perfusion, the infarcted areas correspond to the border zones between the territories of the major cerebral arteries because perfusion is most tenuous here (Fig. 6-75). Emboli can at times be directly visualized by noncontrast CT as hyperdensities within arteries. It is important to note that hemorrhage can not only occur de novo related to a hemorrhagic infarct, but that it can also occur within an ischemic infarct, as a consequence of reperfusion injury due to blood-brain barrier breakdown. When the latter occurs, it appears as a hyperdense mass within the hypodense edema of the infarct (Fig. 6-76).

Injection of intravenous (IV) contrast provides no brain enhancement in the first day or two after a stroke. Contrast enhancement must await sufficient damage to the blood-brain barrier. It reaches its peak at 1 to 2 weeks and usually ceases to occur after 2 or 3 months (79). The greatest vascular damage to intact vessels is at the periphery of the infarct. Therefore, contrast-enhanced CT frequently visualizes a contrast-enhanced ring about the infarcted area or in the immediately adjacent cortical gyri, a phenomenon known as luxury perfusion (Fig. 6-77A,B).

Conventional MRI is more sensitive and specific than CT for the detection of acute ischemic brain infarcts, within the first few hours after the onset of symptoms. On MRI, the edema of an early infarct is of low signal intensity on T1-weighted images with corresponding high signal intensity on FLAIR (fluid-attenuated inversion recovery) and T2-weighted images (Fig. 6-78A,B). In addition, there is loss of gray-white matter differentiation, sulcal effacement, and mass effect analogous to CT imaging findings. With the administration of IV gadolinium-diethylenetriamine pentaacetic acid (DTPA), a damaged blood-brain barrier can often be visualized as a hyperintense area on T1-weighted images. MRI is more sensitive than CT at detecting lacunar infarcts, which are small infarcts of less than 1.5 cm (78) typically located in the basal ganglia, periventricular areas, and at the brain stem (Fig. 6-79A,B). Lacunar infarcts are most commonly caused by hypertension or diabetes-induced arteriolar occlusive disease of the deeply penetrating arteries, such as the lenticulostrate branches of the middle cerebral arteries. MRI is also superior to CT in detecting ischemic infarcts of the posterior cranial fossa, because MR images are not degraded by osseous structures. Another powerful tool in the imaging stroke arsenal is diffusion-weighted MR imaging (DWI). The concept here is that water molecules normally move within tissues in a random fashion known as Brownian
motion. As was discussed earlier, acute stroke produces an electrolyte imbalance, which causes water molecules to rush into the intracellular compartment, where free random motion is no longer possible and therefore falling into a state of restricted diffusion. DW images reflect restricted diffusion as a signal increase, which corresponds with a signal drop in its accompanying sequence, the apparent diffusion coefficient (ADC) map. The combination of increased signal

![Figure 6-74](image)

**FIGURE 6-74.** A: Axial CT scan examination with large left subdural hematoma (arrowheads) and early subfalcine herniation (long arrow). B: Coronal multiplanar reformatted image (MPR) demonstrates to a better advantage the mass effect and the subfalcine herniation (long arrow). The subdural hematoma has a lentiform shape (short arrow).

![Figure 6-75](image)

**FIGURE 6-75.** Axial diffusion weighted images (DWI) of an acute watershed infarct at the left superior frontal lobe. Increase signal intensity (arrow) represents restricted diffusion.

![Figure 6-76](image)

**FIGURE 6-76.** There is a hyperdense region at the left basal ganglia and thalamus surrounded by hypodense rim of edema, compatible with hemorrhagic transformation of an ischemic stroke. Patient had suffered an ischemic stroke 12 days earlier.
in the DW images and decreased signal in the ADC map is compatible with an infarct in the appropriate clinical setting, as other entities such as viscous abscesses and dense masses such as lymphomas can have a similar restricted diffusion pattern (Fig. 6-80A,B). One of the key features of DWI of acute cerebral ischemia is that it becomes positive as soon as 30 minutes after the insult and can remain positive for 5 days or more (80).

Nonenhanced CT is highly sensitive in detecting intracranial bleeds, which, in the setting of an ischemic stroke, represents hemorrhagic transformation. In MR imaging, T2*-weighted gradient-echo sequences depict areas of hemorrhage as focal

**FIGURE 6-77.** Axial CT images of the brain of patient with small focal areas of low attenuation representing lacunar infarcts within the right inferior basal ganglia (A) and within the caudate nucleus (B).

**FIGURE 6-78.** Subacute ischemic infarct. A: Axial CT image demonstrates an area of decreased attenuation (arrow) within the head of the caudate nucleus. B: Axial proton density (PD) sequence. There is increased signal intensity (arrow) due to restricted diffusion characteristic of an infarct.
PART I | PRINCIPLES OF ASSESSMENT AND EVALUATION

regions of low signal intensity, secondary to a phenomenon known as blooming (Fig. 6-81).

As was stated before, cerebral ischemia can be reversible. Tissue that is potentially salvageable with prompt recanalization is referred to as penumbra. The goal of modern stroke imaging is not merely to document an infarct and exclude hemorrhage, but rather to differentiate infarcted from salvageable tissue (penumbra) in an effort to guide thrombolytic therapy and save as much brain tissue as possible. CT and MR imaging techniques that are currently being used with this purpose in mind will be briefly discussed. CT perfusion is a technique in which a bolus of contrast is injected into the patient with simultaneous imaging of a slice of tissue, usually chosen at the level of the basal ganglia, because it represents the three major vascular territories: Anterior, middle, and posterior cerebral arteries. The three main parameters obtained and compared throughout the slice are cerebral blood volume (CBV), cerebral blood flow (CBF), and mean transit time (MTT). In general terms, a mismatch between these parameters usually represents tissue suffering reversible ischemia or penumbra. MR perfusion is a contrast-dependent technique also utilized to determine the amount, if any, of salvageable brain tissue. In general terms, when a perfusion defect matches a diffusion defect, irreversible infarction has occurred. On the other hand, a perfusion-diffusion mismatch represents an area of reversible ischemia or penumbra, where infarction can possibly be avoided with timely thrombolytic treatment.

FIGURE 6-79. Lacunar infarcts.
A: CT shows multiple bilateral lacunar infarcts as small hypodense areas (arrows). B: By MRI, these infarcts are shown as multiple hyperintense areas (arrows).

FIGURE 6-80. Ischemic infarct.
Restricted diffusion is shown as increased signal intensity in the diffusion-weighted image (A) with corresponding signal drop in the ADC map (B) (large arrows).
CTA is a technique that uses IV contrast to image extracranial and intracranial blood vessels. Different methods are utilized to reconstruct the arterial system, in an attempt to identify the cause of the patient’s symptoms, usually an obstructing thrombus or embolus, which is seen as a cut-off in one or various vessels. CTA information is commonly used to guide intra-arterial or mechanical thrombolyis in stroke centers. Just as in CTA, MRA can also be performed following injection of IV contrast. Nevertheless, MRI has the added bonus of being able to perform angiograms without having to inject contrast material based on the MR properties of flowing blood; a useful proposition in patients with renal insufficiency. Contrast-enhanced MRA findings are analogous to CTA findings, nevertheless in non–contrast-enhanced (time-of-flight) MRAs, normal vessels are depicted as a flow void and intra-arterial thrombus is seen as an area of increased signal intensity (Fig. 6-82).

Cerebral venous thrombosis is caused by aseptic or septic etiologies, and can lead to infarction in a nonarterial distribution. This rare cause of infarction has characteristic imaging features. Whether the thrombosis involves a deep cerebral vein or a dural venous sinus, the thrombus can be detected on a noncontrast CT as a hyperdensity within the vein (81). The hyperdensity may have a hypodense center, implying a residual lumen. In a contrast-enhanced CT, a thrombus appears as a filling defect, with tortuous dilated collateral venous channels occasionally demonstrated around the thrombosed vein. By MRI, while the thrombus is still in the oxyhemoglobin stage, which is isodense to brain tissue, it can be suspected by the absence of the normal flow void in that vessel. In the deoxyhemoglobin stage the thrombus is hypointense on T1, and in the later methemoglobin stage it becomes hyperintense on T1-weighted images. The venous thrombus typically does not proceed to the hemosiderin phase because it usually lyases spontaneously and flow is reestablished. Contrast and non–contrast-enhanced MR venography techniques can also be utilized to diagnose venous thrombosis (Fig. 6-83).

A stroke-like clinical presentation frequently encountered in the ER is a transient ischemic attack (TIA). A TIA is a functional neurologic disturbance usually lasting a few minutes, which clears completely within 24 hours. TIAs typically produce no CT or MRI findings, yet one third of these patients eventually will suffer a cerebral infarction, 20% of them within the first month after the episode. Some stroke centers perform an MRI to all patients who suffered a TIA, as occasionally acute infarcts are actually found.

HEMORRHAGIC STROKE

A stroke is considered truly hemorrhagic if blood is found within the first 24 hours after initial symptoms. When blood is noted after this time, it is usually hemorrhagic transformation of an ischemic stroke, which is due to reperfusion injury. Hypertension is the most common cause of intraparenchymal hemorrhage, which can also be caused by ruptured aneurysm, arteriovenous malformation, and more rarely, by infarction, neoplasms, blood coagulation defects, and cerebral
arteritis (81). Common hemorrhage sites include the putamen and the thalamus, which receive their major blood supply from the lenticulostriate and the thalamogeniculate arteries, respectively.

Because freshly extravasated blood is more radiodense than gray or white matter, an acute hemorrhagic stroke is well visualized by CT as a hyperdense region usually conforming to an arterial distribution (Fig. 6-84A,B). The radiodensity of the blood clot increases over 3 days because of clot retraction, serum extrusion, and hemoglobin concentration. The extruded serum may form a hypodense rim around the hyperdense clot (Fig. 6-84C). As edema develops over 3 to 5 days, the hypodense rim may increase. Eventually, the hyperdensity of the clot gradually fades and usually disappears by 2 months, leaving only a narrow hypodense slit to mark the site where hemorrhage took place (Fig. 6-84D).

The appearance of hemorrhage by MRI depends on the state of the hemoglobin in the hemorrhage (81). The oxyhemoglobin present in a fresh hemorrhage is nonparamagnetic; therefore, very early hemorrhage is not detected by MRI. Within a few hours, the oxyhemoglobin will be converted to deoxyhemoglobin, which is a paramagnetic substance. Intracellular deoxyhemoglobin will cause acute hemorrhage to appear very hypointense on T2-weighted images and slightly hypointense or isointense on T1-weighted images (Fig. 6-85A). By 3 to 7 days, intracellular deoxyhemoglobin is oxidized to methemoglobin as the clot enters the subacute phase. Although a subacute hemorrhage has several subphases in which the signal intensity of methemoglobin varies, in general methemoglobin appears hyperintense on both T1- and T2-weighted images (Fig. 6-85B). Because the conversion to methemoglobin begins at the periphery of the clot, early in the subacute phase a hemorrhage can have a hyperintense margin and a central hypointense region still containing deoxyhemoglobin. Eventually the entire region of subacute hemorrhage becomes hyperintense. Over several months, the methemoglobin is gradually resorbed and the clot develops a rim of hemosiderin-containing macrophages. Hemosiderin is hypointense on both T1- and T2-weighted images. Therefore, a chronic hemorrhage of several months duration often has a hyperintense methemoglobin center and a hypointense hemosiderin rim. Because the hemosiderin deposits remain indefinitely, an old hemorrhage of several years duration shows up as a totally hypointense area. Gradient-echo sequences have recently been added to many brain MRI protocols, as they are very sensitive in the detection of degrading blood products, which appear as areas of hypointensity. As can be seen, CT provides the very earliest information about cerebral hemorrhage, whereas MRI is the better technique for determining hemorrhage age.

**SUBARACHNOID AND INTRAVENTRICULAR HEMORRHAGE**

Subarachnoid and intraventricular hemorrhage can be spontaneous, as in the case of a bleeding aneurysm or arteriovenous malformation, or secondary to trauma. CT is the imaging modality of choice for evaluating these types of hemorrhages because it detects the hemorrhage from its onset as a hyperdensity. However, subarachnoid hemorrhage is not as radiodense as epidural or subdural hemorrhage because the blood will be diluted by CSF. Unless blood replaces at least 70% of the CSF, the subarachnoid hemorrhage remains isodense to adjacent gray matter (82). When the volume of blood is sufficient to make the hemorrhage hyperdense, it accumulates in the extensions and expansions of the subarachnoid space. Subarachnoid hemorrhage appears as linear radiodensities within the sulci or fissures or as larger aggregations in the basal cisterns (Fig. 6-86). MRI will not visualize a very early hemorrhage when oxyhemoglobin, a nonparamagnetic substance, is the primary constituent, and thus CT is the study of choice in the very early stages. Subarachnoid and intraventricular hemorrhage can cause communicating hydrocephalus by virtue of red blood cells blocking the arachnoid granulations, the CSF resorption sites.

Aneurysms and arteriovenous malformations can be detected directly by contrast-enhanced CT and MRI, or by their flow void characteristics on non–contrast-enhanced MR images (Fig. 6-87A,B).

**Head Injuries**

Head injury can be produced by direct contact or impact loading, where impacts either set a resting head in motion or stop a moving head, or can be produced by impulse or inertial loading,
and epidural hematomas are produced only by impact loading, but other types of head injury can be produced by either type of loading. Head injury is typically categorized as focal or diffuse. Focal injuries include extracerebral hemorrhages such as...
PART I | PRINCIPLES OF ASSESSMENT AND EVALUATION

CT is typically the imaging modality of choice for patients suffering head trauma, because it is very accurate at detecting the depressed fractures and acute hematomas that require emergency surgery. Its other advantages include rapid scanning with continuation of close monitoring of critically injured patients. MRI has the disadvantages of longer scans, high susceptibility to motion artifacts by uncooperative patients, inability to detect very recent hemorrhage, and difficulty evaluating fractures because of the signal void characteristics of cortical bone.

Epidural hematoma is caused by tears of the middle meningeal artery or vein, or of a dural venous sinus. The blood accumulates in the interval between the inner table of the calvarium and the dura by gradually stripping the dura from its bony attachment. CT visualizes the epidural hematoma as a well-localized biconvex radiodense mass (83) (Fig. 6-88). It is commonly, though not invariably, associated with a skull fracture. It causes mass effect upon the adjacent brain parenchyma with effacement of the underlying sulci, compression of the brain and ventricles, and possible contralateral midline shift. It is important to note that midline shift is a secondary injury caused by subfalcine herniation, which is herniation of the cingulate gyrus under the falx cerebri, and can eventually lead to ipsilateral anterior cerebral artery infarction. When there is a question about whether the mass might be intraparenchymal,
contrast injection enhances the dura, establishing the epidural position of the clot. As the clot lyses over the next few weeks, it shrinks and changes to isodense and then hypodense relative to the brain. The inner aspect of the clot vascularizes, and this may produce a thicker rim of enhancement on late contrast studies. The overlying dura may calcify. Epidural hematoma may be associated with subdural, subarachnoid, or intraparenchymal hemorrhages.

**SUBDURAL HEMATOMA**

Subdural hematoma is most commonly caused by acceleration-deceleration shearing stresses that rupture the bridging veins that extend from the movable brain to the fixed dural venous sinuses. The blood accumulates in a pre-existing but essentially volumeless subdural space. Normally, the pressure of the CSF holds the arachnoid in contact with the dura, thereby creating a real interval that is without significant volume. Because the subdural space is a real space surrounding all external surfaces of the brain, subdural hemorrhage tends to spread extensively over many aspects of the brain surface.

On CT examination, the typical acute subdural hematoma appears as a diffuse crescent-shaped radiodensity that may extend onto many surfaces of the brain, including the cerebral convexity, skull base, interhemispheric fissure, upper or lower surface of the tentorium, and area around the brain stem (Fig. 6-89). One way to differentiate subdural from epidural hematomas is that subdural hematomas cross suture lines yet do not cross midline, whereas epidural hematomas do not cross suture lines yet can cross midline. There are two ways of classifying subdural hematomas based on their changing...
MRI has valuable unique imaging properties that make it very sensitive to the detection of some extracerebral hemorrhages. First, the high signal intensity that subacute hematomas display on T1- and T2-weighted images makes MRI more sensitive than CT for detecting hematomas that are isodense by CT (73). Even chronic subdural hematomas remain hyperintense to CSF and gray matter for several months, which is long after they have become isodense or hypodense on CT. Also, the ability of MRI to discern the displaced signal voids of cortical or dural vessels facilitates the identification of small extracerebral hemorrhages. In addition, when the hematoma collects around the obliquely placed tentorium, axial CT images may average it into adjacent tissues. In these cases, the multiplanar imaging properties of MRI can be very valuable. Also, small hematomas next to the calvarium can be better seen by MRI because they are contrasted against the osseous signal void.

CONTUSIONS AND INTRAPARENCHYMAL HEMORRHAGE

Focal parenchymal injuries such as contusions and intraparenchymal hemorrhage usually develop as a result of contact of the brain with the osseous walls of the cranial cavity. The coup-type injuries occur at the point of contact, and the contrecoup injuries occur on the opposite side of the brain. Contusions often occur in areas where the walls of the cranial cavity are irregular, such as the anterior and middle cranial fossae. Therefore, frontal and temporal lobe contusions are common as the brain glides along these irregular surfaces (85) (Fig. 6-90A,B).

Cerebral contusions are heterogeneous lesions containing edema, hemorrhage, and necrosis, with any element predominating. When blood makes a major contribution, the contusion appears on CT as a poorly delimited irregular area of hyperdensity. A contusion with mostly edema or necrosis may not be detectable immediately, but after a few days it appears as a hypodense region. Where there is a general admixture of elements, contusions may have a heterogeneous density. Old contusions appear as hypodense areas. By MRI, the edematous and necrotic areas have low signal intensity on T1-weighted images and high signal intensity on T2-weighted images, and thus MRI is more sensitive than CT in identifying these nonhemorrhagic contusions. The areas of hemorrhage in a contusion older than a few days will be hyperintense on both T1- and T2-weighted images.

Intraparenchymal hemorrhage differs from contusions by having better demarcated areas of more homogeneous hemorrhage. The CT and MRI characteristics of acute and evolving intraparenchymal hemorrhage are the same as for hemorrhagic stroke.

DIFFUSE BRAIN INJURIES

Diffuse brain injuries include diffuse axonal injury, diffuse cerebral swelling, and edema. Diffuse axonal injury is produced by high shearing stresses that occur at different parts of the brain,
including at the gray matter-white matter interface. These shearing stresses cause axonal stretching commonly involving the corpus callosum, anterior commissure, and upper brain stem. Blood vessels may or may not be disrupted. When vessels are uninterrupted, the scattered small areas of edema are best demonstrated by T1-weighted MR images as slightly hypointense or isointense regions that become hyperintense on T2-weighted images. When vessel disruption produces hemorrhages, they appear early on CT as multiple sites of hyperdensity (Fig. 6-91).

Diffuse cerebral swelling occurs with many types of head injury. It is thought to be produced by a rapidly increased volume of circulating blood. By MRI and CT, the general brain enlargement is visualized by an obliteration or encroachment of the normal CSF spaces: the cortical sulci, the perimesencephalic and basal cisterns, and the ventricles (85). By CT, the enlarged brain may show slightly increased density.

In generalized cerebral edema, the enlarged brain also encroaches on the CSF spaces, but by CT the edema produces a generalized hypodensity that usually takes longer to develop than diffuse cerebral swelling (Fig. 6-92). The edema may obscure gray matter-white matter boundaries.

Both diffuse brain swelling and generalized cerebral edema are emergencies, because if not treated promptly they may lead to brain herniation sometimes with fatal outcomes.

**FIGURE 6-91.** Diffuse axonal injury. Nonenhanced CT scan shows hemorrhagic foci at the genu of the corpus callosum (arrows).
198

PART I | PRINCIPLES OF ASSESSMENT AND EVALUATION

cause CSF leaks that show up as CSF rhinorrhea or otorrhea. Local or diffuse brain swelling can compress the cerebral aqueduct or fourth ventricle, producing obstructive hydrocephalus. Subarachnoid hemorrhage may obstruct CSF resorption and cause a late-developing communicating hydrocephalus. Focal cerebral atrophy can occur at sites of infarction, hemorrhage, or trauma. Generalized atrophy can follow diffuse injuries and can be demonstrated by an increased size of sulci, fissures, cisterns, and ventricles.

Degenerative Diseases of the CNS
Degenerative diseases of the CNS include the wide spectrum of gray and white matter diseases, general degenerative changes of aging, and the dementias.

WHITE MATTER DISEASES
White matter diseases can be divided into demyelinating diseases, in which the white matter is normally formed and then pathologically destroyed, and dysmyelinating diseases, in which there is usually a genetically determined enzymatic disorder that interferes with the normal production or maintenance of myelin (86). The enzymatic disturbances are relatively rare; therefore, their imaging characteristics will not be described.

FIGURE 6-92. Diffuse brain edema. Nonenhanced CT scan shows diffuse hypodensity with sulci effacement and loss of gray/white matter differentiation. Mass effect is causing almost complete obliteration of the ventricular system. Compare low parenchymal attenuation with normal cerebellar density.

FIGURE 6-93. Nonenhanced CT scan shows a left MCA infarct with mass effect causing contralateral midline shift (arrow) corresponding to subfalcine herniation.

PENETRATING TRAUMA

Bullets and other types of penetrating objects will cause brain laceration by both the penetrating objects and the fragments of subcutaneous tissues, bone, and dura driven into the brain. The imbedded fragments of the foreign object and bone are well visualized by CT, as are the accompanying cerebral edema and various types of intracerebral or extracerebral hemorrhage.

COMPLICATIONS OF BRAIN INJURY

Brain injuries may be accompanied by a number of late or long-term complications. These secondary brain injuries include cerebral herniations, which may occur under the falx cerebri or through the tentorium. Herniations can cause compression of adjacent brain substance or vessels, with the production of secondary signs and symptoms (Fig. 6-93). Penetrating injuries or fractures can injure nearby large or small vessels, producing thrombosis, embolism, traumatic aneurysm formation, or internal carotid–cavernous sinus fistula. Basal skull fractures involving the dura and arachnoid can
The most common of the demyelinating disorders is multiple sclerosis (MS). The demyelinating plaques of MS are better visualized by MRI than by CT. In fact, MRI has become the primary complementary test to confirm a clinical diagnosis of MS. It also provides a quantitative means of evaluating the present state of a patient's disease and a mode of following its progress (87). Although the T1-weighted MR images are usually normal, the FLAIR and T2-weighted images demonstrate MS plaques as high-signal-intensity areas. These are most frequently seen in the periventricular white matter, especially around the atrium and the tips of the anterior and posterior horns of the lateral ventricles (Fig. 6-94A,B). The high-signal-intensity plaques also can be seen in other white matter areas of the cerebral hemispheres, the brainstem, and even the upper spinal cord. When these lesions are seen in patients younger than 40 years of age, they tend to be relatively specific for MS (86). In patients more than 50 years of age, the MRI findings of MS are similar to findings in some aging brains, and correlation with the clinical findings helps establish the diagnosis. Recent MS plaques that involve damage to the blood-brain barrier frequently enhance with the use of IV gadolinium-DTPA (Fig. 6-95).

CT demonstrates MS plaques with less reliability than does MRI. On CT, these plaques appear as areas of hypodensity. Recent plaques in the acute phase of an exacerbation of the disease will have damage to the blood-brain barrier, and IV contrast will then enhance the periphery of the lesion. In the chronic plaque, no contrast enhancement occurs on CT or MRI. Other demyelinating diseases, although numerous, are of relatively low incidence and therefore are not described.

**FIGURE 6-94.** T2-weighted (A) and FLAIR (B) MRI demonstrates the periventricular demyelinating plaques of MS as hyper-intense areas (arrows) adjacent to the anterior horns and atria of the lateral ventricles.

**FIGURE 6-95.** Active MS. T1 gadolinium enhanced MR image shows periventricular enhancing MS plaques (arrows).
GRAY MATTER DISEASES

At present, MRI is being used clinically to discriminate a number of movement disorders that are characterized by changes in the size or iron content of a number of deep gray matter nuclei (88). Normal nuclei that contain high iron levels, such as the globus pallidus, reticular part of the substantia nigra, red nucleus, and dentate nucleus of the cerebellum, appear hypointense on T2-weighted images. In Parkinson’s disease, T2-weighted MRI shows a hypointensity in the putamen that may exceed the normal hypointensity of the globus pallidus. In Huntington’s chorea, MRI consistently shows atrophy of the head of the caudate with associated dilation of the adjacent frontal horn of the lateral ventricle. Some patients with Huntington’s chorea also show a hypointensity of the caudate or putamen on T2-weighted images and atrophy predominating in the frontal lobe. Some forms of secondary dystonia show increased signal intensity of the putamen and caudate in T2-weighted images.

Age-Related Changes and Dementing Disorders
The aging brain is characterized on CT or MRI as demonstrating volume increases in both cortical sulci and ventricles (Fig. 6-96). T2-weighted MR images also frequently display small areas of hyperintense signal along the anterolateral margins of the anterior horns of the lateral ventricles. These changes may or may not be associated with neurologic findings.

Patients with Alzheimer’s disease (AD) and other dementing disorders consistently show these age changes, but because many normal elderly also do, these changes cannot be used to diagnose AD. However, the absence of these findings typically excludes AD. Findings more specifically related to AD are those involving the temporal lobe. The earliest findings in AD involve atrophy of the temporal lobe with dilation of the temporal horn of the lateral ventricle, as well as dilation of the choroidal and hippocampal fissures caused by atrophy of the hippocampus, subiculum, and parahippocampal gyrus (89).

EMERGING IMAGING TECHNOLOGIES OF INTEREST TO THE PHYSIATRIST

Advanced MR Imaging Techniques
Since the advent of medical imaging using nuclear magnetic resonance (NMR), there has been a continuous advance in this technology with improvements in spatial resolution, as well as the introduction of new methods that allow for the rapid acquisition of images and the ability to image different aspects of dynamic function within the imaged tissues. Most of the initially introduced techniques involved the imaging of static structural characteristics at different points in parenchymal tissue. Proton density imaging basically reflects the concentration of water and the electrochemical environment of the water-based protons in the tissues. When water-based protons are actively in motion during the standard MRI imaging sequence, a “flow artifact” is created. However, if one is actually interested in the imaging of proton movement, then the flow artifact becomes a flow signal to be recorded. This is the basis for both magnetic resonance angiography (MRA), which is actually the imaging of proton flow within the vascular spaces, and for diffusion-weighted imaging (DWI) and diffusion-tensor MR imaging (DTI), which is the imaging of the molecular diffusion of water molecules through tissue. When techniques are employed to image the diffusion of specific molecular species related to tissue metabolism, one can develop MR-based imaging techniques that allow one to image the active perfusion of tissues as well as their metabolic activity. This is the basis for perfusion-weighted and functional MRI (fMRI). Functional changes in the brain occur at relatively rapid rates during performance, and in order to image these changes, images must be acquired at a very rapid rate. Echo planar imaging (EPI) is a method for MR image acquisition that optimizes the image acquisition algorithm so that an entire image is acquired within a single TR period. EPI permits the acquisition of tomographic images at “video” rates of 15 to 30 images a second or down to as low as 20 ms per image, depending on the pulse sequence employed. EPI has been particularly important for the advancement of fMRI of the brain but is also used in DWI to reduce motion artifact.
MAGNETIC RESONANCE ANGIOGRAPHY

Conventional angiography uses x-ray methods that involve introduction of an iodine-based radio-opaque dye into the vascular space through an intravascular catheter, followed by imaging of the distribution of the dye within different vascular compartments. This technique involves the invasive introduction of the dye via intra-arterial catheterization and potential problems resulting from the physical introduction and threading of the catheter, as well as allergic reactions to the dye. The technique can be used to produce a structural image of a vascular tree but does not provide direct information regarding the relative flow of the blood within the tree. MRA, on the other hand, involves no introduction of dye, except in the contrast-enhanced technique described later, and no need for invasive catheter introduction. The technique takes advantage of the principle of signal intensity attenuation from a region in which the imaged protons are actively in motion, producing a dark-signal “flow void” that marks a vascular structure containing moving blood seen on a standard T1-weighted image. MRA produces images of dynamically flowing blood, with the intensity of the image signal being proportional to the actual blood flow velocity. MRA, therefore, provides a dynamic image of circulatory flow rather than a static image of a vascular structure as is obtained with conventional x-ray angiography. Generally, there are three different methods used to produce MRA images: time-of-flight, phase contrast angiography, and contrast-enhanced angiography.

As was described above, time-of-flight imaging involves utilizing the normal loss of signal produced by moving intravascular protons, seen as a signal drop or “void,” to create images of the vascular tree.

Phase-contrast angiography involves the application of a bipolar magnetic field gradient pulse. Two imaging sequences are performed in which the first has a positive bipolar gradient pulse and the second has a negative bipolar gradient pulse and the raw data from the two images are then subtracted. The signals emitted by stationary spin sources cancel out, whereas the signals emitted from flowing sources will add up. The net effect is to produce an image of proton spins that are flowing.

In contrast-enhanced MRA, the imaging of the vessels relies on the difference between the T1 relaxation time of the blood and that of the surrounding tissues after a paramagnetic contrast agent has been injected into the blood. The agent reduces the T1 relaxation time in the blood relative to the surrounding tissues. This technique generally provides a higher quality imaging of structural vascular anomalies comparable to that obtained with x-ray angiography. In addition, MRA techniques are normally obtained in conjunction with conventional MRI sequences, and therefore information regarding surrounding brain parenchyma also becomes available. An example of a normal contrast enhanced MRA is shown in (Fig. 6-97A,B).

FIGURE 6-97. (A and B) Normal magnetic resonance angiography (MRA).
evaluate white matter lesions in severe traumatic brain injury (TBI) suggest that the white matter lesion “load” associated with diffuse axonal injury, as detected by this sequence, correlates with the severity of the clinical presentation and the eventual clinical outcome (91–93). FLAIR images are generally more sensitive in the identification of white matter lesions than conventional T2-weighted images or T2-fast spin echo techniques. However, the technique is also subject to CSF pulsation artifact and to vascular flow artifact, which can limit its utility in certain situations. A FLAIR image of a left middle cerebral artery territory stroke is shown in Figure 6-99 compared to a corresponding T2-weighted image (Fig. 6-100).

**DIFFUSION-WEIGHTED IMAGING**

Diffusion is a physical property of molecules whereby they move randomly via Brownian motion and spread out through a medium in accordance with their thermal energy. The rate of diffusion for a particular molecule in a particular environment is measured by the diffusion coefficient. Molecular diffusion is influenced by concentration (i.e., chemical) gradients, as well as by mechanical tissue structure that can confer a directional (“anisotropic”) component. For example, in muscle or in cerebral white matter, anisotropic tissue structure creates a preferred pathway for water diffusion parallel to muscle or nerve fibers and the direction of diffusion is largely uniform across each imaged voxel. Diffusion of polarized molecules may also be influenced by an electrical potential gradient. Diffusion imaging is performed in the same general manner as phase-contrast angiography, described earlier, except that the amplitudes of
CHAPTER 6 | IMAGING TECHNIQUES RELATIVE TO REHABILITATION

203

pump. This membrane-based pump, which appears exquisitely sensitive to ATP availability, normally maintains water equilibrium across cell membranes and thereby regulates cell volume by controlling intracellular and extracellular cation concentrations. The resulting decrease in average diffusion of water-dependent protons in the ischemic region is detected as an area of increased signal intensity on the diffusion-weighted MR image. Signal intensity rapidly rises with tissue ischemia with the diffusion coefficient for water reduced by as much as 50% within minutes after onset of ischemia. With the emergence of new early interventions for acute cerebrovascular thrombosis, improved patient selection and triage at a very early point in the evolution of cerebral ischemia could be achieved with DWI. Furthermore, the changes noted during acute ischemia on DWI, particularly when compared with MR perfusion images described earlier, may indicate reversible tissue dysfunction and may be very helpful in selecting patients for early thrombolytic intervention for acute cerebrovascular thrombosis (95,96).

Additional applications of DWI involve the assessment of restricted diffusion due to the anisotropic structure of surrounding tissue such as the myelinated fiber tracts in cerebral white matter and in the spinal cord. The assessment of restricted diffusion involves computing an ADC for each of six different directions for each voxel in the image and using this to derive a diffusion tensor. This technique has been called diffusion tensor imaging (DTI). Using geometric methods to determine the degree of loss of anisotropic diffusion in subcortical white matter, it is possible to document white matter disruption such as focal dislocation, tearing, swelling, and infiltration. It is also possible to map out the structure and direction of major subcortical white matter tracts using a technique called diffusion tensor tractography (97). DWI and fMRI have been used together to map out altered white matter anatomy and to identify eloquent cortical regions for movement and language generation in process of planning brain tumor excisions. Imaging information can be used so that displaced but intact fiber tracts can be preserved and critical cortical regions can be left unharmed. There is now also ongoing active research to evaluate the application of DWI to the assessment and recovery from diffuse axonal injury due to TBI (98,99). Recent studies with limited numbers of patients suggest that DWI may be a helpful technique for detecting the presence and localizing the zones of damage in TBI due to diffuse axonal injury when standard imaging methods show no abnormality (100–102). This is clearly an evolving area with some currently used imaging applications and important future research.

IN VIVO MAGNETIC RESONANCE SPECTROSCOPY IMAGING

Although most of the MR signal that is typically examined in clinical imaging comes from water-based protons, it is possible to study specific resonances from other sources in
order to image other metabolically important molecules in the tissue. Although it is relatively easy to obtain a detailed NMR spectrum from a bulk sample as is typically done in NMR-based analytical chemistry, it is much more difficult to obtain the full spectrum from a spatially restricted volume of tissue. A number of different approaches have been taken to develop spectroscopic imaging techniques that would permit the recording of the NMR spectrum for each voxel in an image. However, because the concentration of most metabolic molecular species of interest (e.g., lactic acid) is typically several orders of magnitude less than that of water or fat in different tissues, the proton-NMR signals from water and fat must be suppressed when performing proton spectroscopy of metabolites, and large magnetic fields are also required to permit improved signal-to-noise ratio so that these much smaller spectral peaks can be registered clearly. At present, in vivo MR spectroscopy imaging is being used in a wide range of clinical as well as research applications.

PERFUSION-WEIGHTED IMAGING AND FMRI

PWI of the brain is obtained by rapidly injecting a paramagnetic dye (e.g., gadolinium) intravenously and imaging the appearance of the dye in perfused brain tissue. Its utility in stroke imaging, particularly when used in conjunction with DWI, was discussed earlier in the section on ischemic stroke. Functional MRI depends on the coupling of neuronal firing to local changes in metabolic rate that are, in turn, linked to changes in local perfusion of normal brain tissue. In a locally active region of brain tissue, there is an increase in oxygen extraction to support the increased metabolic rate of the tissue. However, there is also an increase in local CBF to support the active region, which increases oxygen delivery to the area to the point where it exceeds that which is actually required. There is therefore a relative decrease in the local concentration of deoxyhemoglobin at the postcapillary level in activated tissue. Deoxyhemoglobin contains iron and is paramagnetic, thus diminishing signal intensity on a T2-weighted image. Since deoxyhemoglobin is decreased in activated tissue, there is a net increase in signal intensity in this region on the T2-weighted image. This is a small signal but can be recovered from the background by averaging together several images obtained by rapid-acquisition EPI methods. The images are obtained while the subject is engaged in performing a specific continuous activating task such as repetitively moving the fingers of one hand or performing an extended cognitive task. This fMRI technique is referred to as blood-oxygen-level-dependent imaging, or BOLD. A number of other techniques have also been developed for use in fMRI applications (103).

USI of Muscle, Nerve, Connective Tissues, and Joints

USI involves the creation of visual images from the reflection of high-frequency mechanical waves back from different interfaces within the imaged tissues. Medical USI uses the ability of high-frequency sound to penetrate through multiple layers of tissue so that images can be constructed of the structures lying under the skin. Surfaces at which sound conductance changes generate an echo back to the transducer, depending on how much the sound conductance changes at the interface. Various different probes are available for USI. A transverse linear array probe is able to image from relatively flat surfaces and can be used to study soft tissues of the limbs or trunk. New machines that offer probe frequencies of up to 16 MHz offer the best-quality imaging of anatomic details for soft-tissue neuromusculoskeletal examination (Fig. 6-101).

USI OF MUSCLE

The appearance of muscle with USI is determined by the complex interweaving of fibrous tissue and myocytes in the stromal architecture of the muscle. Reflections arise at interfaces between relatively dense fibrous tissue and relatively soft muscle tissue, where fibrous tissue wraps around individual muscle fascicles. In sagittal section, these fascicular interfaces form longitudinal striations along the length of the muscle. With an axial section of the muscle, the end-on view of the fascicular structure...
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>T1 weighted</td>
<td>TR: short (&lt;1,000 ms) TE: short (&lt;30 ms) Spin echo technique</td>
<td>T1 contrast. Fluid is dark. Fat is white. Brain parenchyma is gray. Tumor and edema are gray or dark. Ligament and tendon are dark. Methemoglobin (hemoglobin breakdown product &gt;7 days) has high signal intensity.</td>
<td>Best spatial resolution for connective-tissue anatomy and bone marrow and trabecular anatomy. Myelin distribution and mucoid degeneration of tendons/menisci. Used for all post-contrast imaging—e.g., with gadolinium. Identify subacute to chronic hemorrhage.</td>
<td>Poor tissue edema contrast resolution. Poor sensitivity to pathologic lesions since most lesions involve increased tissue or edema fluid and T1 is not sensitive to the appearance of tissue fluid associated with acute inflammatory reaction or pathologic tissue changes.</td>
</tr>
<tr>
<td>T2 weighted</td>
<td>TR: long (&gt;1,500 ms) TE: long (&gt;60 ms) E.g., FSE technique</td>
<td>T2 contrast. Fluid is white. Fat is variable. Brain parenchyma is gray. Tumor and edema are white (hyperintense). Ligament and tendon are dark.</td>
<td>Regions with high free water content have high signal intensity—e.g., CSF, edema, nucleus pulposus, synovial fluid, abscess, and hyperacute hemorrhage (&lt;1 day). Sensitive to pathologic appearance of fluid in parenchymal tissue.</td>
<td>Difficult-to-see high-intensity lesions adjacent to CSF or other free fluid regions—CSF has high signal intensity, as do T2 lesions; therefore tissue lesions at CSF interface are obscured.</td>
</tr>
<tr>
<td>STIR</td>
<td>TR: long (&gt;4,000 ms) TE: short (&lt;50 ms) TI: short values to null out fatty tissue</td>
<td>Very sensitive to tissue edema. Water content in soft tissue (e.g., nerve and muscle) produces a high signal intensity. Fat is suppressed.</td>
<td>MR neurography Muscle pathology imaging Pathologic fracture, bone edema, ligamentous or tendinous injury imaging</td>
<td>Poor spatial resolution but suitable to large body parts such as limbs and trunk or pelvis.</td>
</tr>
<tr>
<td>FLAIR</td>
<td>TR: long (e.g., 9,000 ms) TE: long (e.g., 130 ms) Heavily T2-weighted Ti adjusted to null out free fluid signal (e.g., 2,200 ms)</td>
<td>Heavily T2-weighted with bulk water suppression, but without extracellular fluid/tissue edema suppression</td>
<td>White-matter T2 lesions seen more clearly than with T2-weighted Applied to MS, DAI/TBI, Lyme, HIV, brain infarction/ischemia</td>
<td>CSF pulsation artifact Blood flow artifact Requires careful adjustment of TI, especially when CSF is abnormal</td>
</tr>
<tr>
<td>MRA</td>
<td>Phase-contrast imaging; bipolar MFG.</td>
<td>Flowing spins</td>
<td>Blood flow; angiography Noninvasive screening for cerebrovascular anomalies (e.g., cerebral aneurysms) Ischemia; acute stroke; TBI; DAI Held to be exquisitely and nearly immediately sensitive to the effects of ischemia on brain tissue due to focal reduction in diffusion of extracellular water in the ischemic region. H2O moves into intracellular space due to metabolic failure of ATP-dependent membrane-based Na-K pump Tissue perfusion assessment</td>
<td>Longer study time Lower spatial resolution compared with x-ray angiography Motion artifact; MR measurement cannot differentiate diffusion from local blood flow or tissue pulsation; T2-weighted due to long probe time. Can correct for effect of T2 weighting by calculating map of ADC values.</td>
</tr>
<tr>
<td>DWI</td>
<td>EPIA; TR: long (e.g., 10,000 ms) Heavily T2-weighted Phase-contrast imaging with large-amplitude bipolar MFGs Spin echo</td>
<td>Diffusing spins. To look at “restricted” or “anisotropic” diffusion, a diffusion “tensor” or directionally specific diffusion rates is computed for each voxel. DTI can be used to compute pathways of white fiber tracts (i.e., diffusion tensor tractography) and focal white-matter disruption</td>
<td>Ischemia; acute stroke; TBI; DAI Held to be exquisitely and nearly immediately sensitive to the effects of ischemia on brain tissue due to focal reduction in diffusion of extracellular water in the ischemic region. H2O moves into intracellular space due to metabolic failure of ATP-dependent membrane-based Na-K pump</td>
<td>Ischemia; acute stroke; TBI; DAI Held to be exquisitely and nearly immediately sensitive to the effects of ischemia on brain tissue due to focal reduction in diffusion of extracellular water in the ischemic region. H2O moves into intracellular space due to metabolic failure of ATP-dependent membrane-based Na-K pump Tissue perfusion assessment</td>
</tr>
</tbody>
</table>
TABLE 6.1 A Comparison of MR Imaging Methods (Continued)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>PWI</td>
<td>TR: short TE: short Images taken while contrast medium (e.g., gadolinium) power-injected intravenously at a fixed rate</td>
<td>Appearance of blood-conducted paramagnetic contrast medium in different brain regions</td>
<td>Involves dye injection using specialized equipment to control rate of infusion</td>
<td></td>
</tr>
<tr>
<td>fMRI—BOLD</td>
<td>EPIA; T2-weighted; Multiple image averaging to extract deoxyhemoglobin signal</td>
<td>Spatially localized tissue decrease in deoxyhemoglobin in regions of functionally related metabolic activation</td>
<td>Local tissue activation in functional brain activation/physiologic studies</td>
<td>Must average images to extract small deoxyhemoglobin signal Limited temporal resolution Subject must be cooperative and able to continuously perform the activation task</td>
</tr>
</tbody>
</table>

NOTE: The exact choices of parameter settings for TR, TE, and TI will vary according to the size of the static external magnetic field and various design details of the MR instrument.

ADC, apparent diffusion coefficient; DTI, diffusion tensor imaging; DWI, diffusion-weighted imaging; EPIA, echo planar image acquisition; FLAIR, fluid-attenuated inversion recovery; fMRI—BOLD, functional magnetic resonance imaging—blood-oxygen level-dependent technique; FSE, fast spin echo; MFG, magnetic field gradient; MRA, magnetic resonance angiography; PWI, perfusion-weighted imaging; STIR, short-tau inversion recovery; T1, longitudinal relaxation time (aka “spin-lattice relaxation time”—magnetization relaxation parallel to the static external magnetic field); T2, transverse relaxation time (aka “spin–spin relaxation time”—magnetization relaxation perpendicular to static external magnetic field—“precession”); TE, echo record time between RF pulse and signal recording window; TI, inversion recovery time; TR, repetition time between each new excitation–emission sequence.

produces a stippled cellular mosaic pattern (Fig. 6-102). The echogenicity of the muscle depends on the relative density of the fibrous matrix. In the presence of disuse atrophy, the echogenicity of the muscle increases as a result of volume loss in the myocytes and increased fibrofatty tissue content of the muscle. The normal bone surfaces are smooth and echogenic with posterior shadowing. The subcortical bone and underlying marrow are not evaluated with ultrasound, a disadvantage when compared to MRI. Tendons tend to be intermediate between bone and muscle in terms of echogenicity and have a dense nonpulsatile fibrillar pattern that helps to separate their appearance from vessels and nerves (107). Furthermore, they can be readily recognized by having the subject contract the muscle connected to the visualized tendon so that its active movement can be directly observed. This highlights one of the very unique and valuable features of USI: its ability to generate real-time images of structures that are dynamically active. The ultrasound image is constantly being updated on the screen at video rates to produce a “live-action” dynamic imaging of the tissue structure. Thus, it is possible to evaluate the relative motility of different structures and the “slide” that occurs at tissue interfaces as active and passive movement occurs during imaging. When a normal contracting muscle is imaged, the muscle can be seen to enlarge in cross section with irregular jerky motions seen within the muscle bulk. These appear to be contractions of individual fascicles during voluntary contraction of motor units. Fasciculations can be readily identified on real-time USI as a sudden isolated limited contraction within the muscle. USI allows for specific indications as to the intramuscular location of the motor unit generating the fasciculation. Also, dynamic USI can evaluate intrafascial or interfascial muscle herniation.

FIGURE 6-101. MRI diffusion-weighted image of an acute left middle cerebral artery stroke.
USI OF PERIPHERAL NERVE

USI can also be applied to the imaging of peripheral nerves. Nerves can have either an echogenic cross section when surrounded by relatively echolucent tissues or a relatively echolucent appearance in surrounding tissues that are relatively hyperechoic. With high-resolution USI, a stippled transverse cross-sectional appearance can be seen reflecting the fascicular internal structure of the nerve. Nerves can be differentiated from the three other major tubular structures: tendons/ligaments, arteries, and veins. Tendons/ligaments have well-defined anatomic locations, are densely hyperechoic, and can be made to move in a specific manner with appropriate patient movements. Arteries tend to pulsate regularly and veins collapse when compressed.

The median nerve can be readily identified in a transverse cross section of the wrist just proximal to the transverse carpal ligament. It lies just above the tendons of the flexor sublimis and deep to the palmaris tendon at this point and has a flattened oval appearance. In sagittal section, it is seen as a relatively echolucent structure with tendons above and below it. When the fingers are flexed, the flexor digitorum sublimis tendons can be seen moving back and forth under the static image of the nerve. In the axial transverse image, the nerve is normally very motile and changes its shape readily with wrist and finger motion. With full flexion of the fingers into the palm of the hand, the median nerve initially rises and then dives down into a deep pocket formed by the first sublimis tendon above, the flexor pollicis longus tendon medially and the first profundus tendon below. When the nerve becomes pathologically enlarged, some of this dynamic motion of the nerve is lost. It is also possible to generate images of the ulnar nerve as it passes around the elbow (Fig. 6-103).

In patients presenting with CTS, some changes are consistently noted on USI. Some doming of the flexor retinaculum may be present, reflecting a generalized enlargement of the contents of the tunnel. There may be some observable compression of the nerve, usually at the distal edge of the flexor retinaculum. The nerve appears less deformable with compression and movement and may have increased blood flows associated with it, which can be detected with Doppler ultrasound. Finally, the nerve often is noted to be significantly enlarged (Fig. 6-104). The reason for this enlargement may be a combination of local edema, increased intrafascicular fibrous tissue development within the nerve (related to chronic inflammation), or a backup of axoplasmic fluid flow caused by compression. Furthermore, the nerve appears to be less mobile within the carpal segment, particularly with rapid repetitive motions of the wrist and fingers. USI can therefore provide some useful information about anatomic changes in the median nerve associated with CTS (Fig. 6-105). Imaging of the contents of the carpal tunnel can also detect other problems that could be contributing to CTS symptoms, such as nerve and ganglion cysts, synovial cysts, and osteophytes (108). USI has also been applied to the study of enlarged nerves in the hereditary motor
and sensory neuropathies, to studies of ulnar nerve entrapment in the cubital tunnel, and to evaluation of the brachial plexus.

**USI OF JOINTS**

Ultrasound is currently used for the evaluation of joint internal derangement. The shoulder, elbow, wrist, hip, knee, and ankle articulations are routinely evaluated with ultrasound for ligament, tendon, cartilage and fibrocartilage abnormalities. Current ultrasound technology allows imaging of structures not only during rest but also in real time during controlled limb movements, so that movement-dependent changes in anatomic relationship (e.g., shoulder impingement and snapping syndromes) can be visualized and evaluated. Different USI findings can be seen in muscle tear and rupture, tendon tear, tendonitis, ligament tears, and various soft-tissue overuse syndromes. Evaluation with USI, in the hands of a well-trained individual who understands the underlying anatomy and pathophysiology, as well as the advantages and limitations of ultrasound technology, can be very helpful in the differential diagnosis of musculoskeletal complaints and the assessment of extent and severity of a musculoskeletal injury or condition. A detailed review of the application of musculoskeletal ultrasound is beyond the scope of this chapter, but interested readers are referred to available sources (109,110).

**REFERENCES**

INTRODUCTION

With recent advances in computer technology, equipment miniaturization, reduced cost, and ease of use, the clinical applications of diagnostic ultrasound (U/S) have spread across various medical specialties, including musculoskeletal medicine. In this chapter, we review several common pathologies of the shoulder, the elbow, and the knee to demonstrate the utility of diagnostic U/S in musculoskeletal medicine. Detailed procedures for examining the upper and lower extremities have been described in two recent publications (1,2) and are only briefly mentioned in this chapter.

Two major imaging modalities for detection of soft tissue injuries are magnetic resonance imaging (MRI) and (U/S). Musculoskeletal U/S is, by no means, a replacement of MRI. Instead, it should be viewed as an extension of our physical examination. Typically, MRI can reveal static features of muscles, tendons, nerves, and bones. However, when compared with U/S, the disadvantages of MRI are (a) its limited accessibility in the clinic, (b) longer examination time, and (c) higher cost. On the other hand, modern musculoskeletal U/S can provide high-resolution and real-time imaging of the nerves, tendons, muscles, and joint recesses, provided that the structures are not too deep or obscured by hyperechoic body parts (3).

Musculoskeletal U/S plays a role not only in the assessment of soft tissue injuries but also as an adjunct to a number of common interventional procedures. When used appropriately, U/S guidance improves the accuracy of steroid injection into joint cavities, bursa, and tendon sheaths, thus improving its therapeutic efficacy (4,5) and thereby reducing the risk of iatrogenic complications. By the same token, the application of U/S for regional nerve blocks is also gaining popularity (6,7). In a recent study, ultrasonography has been successfully used to locate the sacral hiatus for caudal epidural injections (8). Moreover, U/S-guided sacroiliac joint injection, facet joint injection, and medial branch block have been advocated as viable options over fluoroscopy- and computed tomography–guided techniques (9–11).

U/S images vary with the reflection of the U/S waves, the amount of which defines the echogenicity. Thus, common terms used to describe the anatomic structures in the areas of interest include “hyperechoic,” “isoechoic,” “hypoechoic,” and “anechoic.” In ultrasonography reports, the images are also described in terms of the plane with which the scanning was performed, either longitudinal or transverse. Typically, a linear array probe is used in musculoskeletal examination, as its wider view and higher near-field resolution provide good images of superficial structures. Figure 7-1 shows an example of a U/S machine and the transducer probes. The degree of U/S penetration also depends on its frequency. Probes with higher frequency ranges (7 to 12 MHz) are commonly used to assess very superficial structures. Probes with lower frequency ranges (5 to 7.5 Hz) are often used to assess structures that are deeper, since they allow greater tissue penetration (3). In addition, power Doppler, a technique that takes into account the amount of red blood cells being scanned, can be used to indirectly demonstrate blood flow within the scanned area. In the three sections below, we review the imaging of several common musculoskeletal pathologies, (a) the shoulder, (b) the elbow, and (c) the knee.

EXAMPLES OF SHOULDER PATHOLOGY

Supraspinatus Tendon Tear

Supraspinatus tendon tears are commonly seen in individuals participating in sports such as baseball, tennis, or swimming. Pain on resisted shoulder abduction suggests pathology of the supraspinatus tendon. On longitudinal scan, a normal supraspinatus tendon should appear as a beak-shaped, echogenic fibrillar structure extending under the acromion, between the humeral head and the subacromial/subdeltoid bursa. On transverse scan, the tendon appears as a band of medium-level echogenic structures, deep to the subdeltoid bursa and superficial to the hypoechoic hyaline cartilage on the humeral head. A partial-thickness tear is demonstrated as a focal anechoic lesion (Fig. 7-2A) or as a mixed “hyperechoic and hypoechoic” focus in the poorly vascularized critical zone of the supraspinatus tendon. The main sonographic features of a full-thickness tear (Fig. 7-2B) include focal nonvisualization through the width of the tendon, as demonstrated in the transverse view (12).

There is inherent interobserver variability in the detection and characterization of supraspinatus tendon tears (12,13). However, with standardized diagnostic criteria, high-quality scanning equipment, and well-trained sonographers, several studies have reported sensitivities and specificities exceeding 90% in the detection of full- and partial-thickness rotator cuff tears (14–16).
FIGURE 7-2. Supraspinatus tendon tear. A: A longitudinal scan of supraspinatus tendon (SS) with a partial-thickness tear reveals an anechoic defect (white arrow) inside the tendon. B: A transverse scan of SS with a full-thickness tear reveals a sonoluent defect inside the tendon and a hyperechoic line (white arrow) between the defect and the hypoechoic cartilage of the humeral head (cartilage interface sign).
Subacromial-Subdeltoid Bursitis
Subacromial-subdeltoid bursitis is often associated with repeated trauma; and in middle-aged or older individuals, it tends to be linked with overuse or degenerative changes in the rotator cuff (17). A normal bursa should appear on U/S as a thin hypoechoic stripe, covered by a narrow layer of echogenic peribursal fat, located between the underlying supraspinatus tendon and the overlying deltoid muscle. Typically, it is less than 2 mm in thickness, even counting the hypoechoic layer of fluid located between the two sides of the bursa (18).

A small effusion in the subacromial/subdeltoid bursa may be identified lateral to the greater tuberosity, especially with the arm extended and internally rotated (Fig. 7-3A). The examiner should be careful not to compress and displace the small amount of fluid. Fluid accumulation within the subacromial/subdeltoid bursa is often noted in patients with infectious or inflammatory bursitis (Fig. 7-3B). However, it can also be observed in patients with full-thickness tear of rotator cuff tendon or in individuals with shoulder impingement syndrome (19). If a needle intervention is deemed necessary, U/S can be utilized to guide bursal fluid aspiration or steroid injection (20).

Bicipital Tenosynovitis
Bicipital tenosynovitis is an inflammation of the long head of the biceps where the tendon passes through the bicipital groove. When the biceps tendon is inflamed, local tenderness in the bicipital groove and an increasing painful arc (painful sensation when the shoulder is flexed from 30 to 120 degrees) are often present. When a clinician is uncertain about the accuracy of the Yergason’s supination test or Speed’s test (21,22), he or she may consider the use of U/S to improve its diagnosis and treatment.

The long head of the biceps brachii originates at the supraglenoid tubercle and glenoid labrum in the most superior portion of the glenoid. It lies in the bony groove between the greater tuberosity and the lesser tuberosity of the humeral head. The transverse bicipital ligament keeps the tendon confined within the groove. On transverse scan, with the arm in neutral position, the long head of biceps tendon of a normal subject is visualized as an oval-shaped, echogenic structure within the bicipital groove. In ultrasonography, the tendon is typically seen as an array of echogenic fibrillar lines, emerging from beneath the acromion and traversing distally to the musculoskeletal junction, where it becomes indistinguishable from the muscle belly.

In biceps tendonitis, effusion in the tendon sheath is associated with focal tenderness and often with heterogeneity of the tendon (Fig. 7-4A,B). Bicipital sheath fluid accumulation indicates intra-articular pathology rather than biceps tendon pathology per se, because of the connection of the sheath with the glenohumeral joint (23). Thus, bicipital sheath effusion may also be seen in supraspinatus tendon tear (24).

**EXAMPLES OF ELBOW PATHOLOGY**

**Lateral Epicondylitis**
Lateral epicondylitis is usually caused by repetitive traction of the common extensor tendon (CET) at its osteotendinous attachment to the lateral humeral epicondyle (27). Diagnosis is usually made clinically, without any need for imaging. U/S provides direct visualization of the CET for both the patient and the clinician and can provide confirmatory signs in the
evaluation process (25,26). In the longitudinal view via U/S examination, the CET should appear as an echogenic beak-shaped structure (Fig. 7-5A). Main U/S features of a “tennis elbow” may include swelling of the tendon, partial tear, calcification, and hyperemia. In Figure 7-5B, the diffuse, thickened hypoechoic area with loss of normal fibrillar pattern is compatible with tendinopathy. A focal anechoic area on ultrasonography indicates a complete or partial rupture of the tendon.

**FIGURE 7-4.** Biceps tendinitis. **A:** A transverse scan at the level of the bicipital groove shows anechoic fluid (white arrow) accumulation surrounding a tender, thickened biceps tendon in a patient with biceps tenosynovitis. **B:** On longitudinal scan, fluid accumulation (white arrow) around the tendon is noted.

**FIGURE 7-5.** Lateral epicondylitis. **A:** Normal sonographic image of the lateral elbow. The CET (T) presents as an echogenic beak-shaped structure, while the extensor muscle (EM) is relatively hypoechoic. LE, lateral epicondyle; RH, radial head. **B:** Swelling of the tendon is seen as hypoechoic thickening with loss of normal fibrillar pattern (between the arrowheads) near its insertion. **C:** Linear shape calcification (arrows) is seen inside the tendon. **D:** Power Doppler indicates a hypervascular status inside the tendon (arrows). This is frequently seen in an actively inflamed tendon.
CET, which may require surgical intervention. As expected, calcification of the tendon can be seen as echogenic foci within the CET with or without acoustic shadow beneath the lesion (Fig. 7-5C). Cystic lesions can occasionally be seen around the CET. With power Doppler imaging, hypervascularity of the CET (Fig. 7-5D) is suggestive of focal hyperemia and active inflammation.

**Ulnar Collateral Ligament Rupture**
An ulnar collateral ligament (UCL) tear may result from either acute or chronic valgus stress to the elbow. Repeated overhead throwing or pitching is a common cause of UCL injury (28,29). Traditionally, MRI has been used for assessment of UCL injury. However, U/S not only can visualize the UCL in a static view (Fig. 7-6) but also can evaluate the related structures during motion. A dynamic U/S study can assess joint laxity by comparing the degree of joint widening of both arms during valgus stress. Both U/S and MRI have been used to assess UCL injury, and their diagnostic accuracies are quite comparable (30). In another study, 26 asymptomatic major league professional baseball pitchers were assessed by U/S. When compared to the non-pitching arms, the pitching arms were found to have thicker UCL anterior bands and wider joint spaces (31).

**EXAMPLES OF KNEE PATHOLOGY**

**Knee Effusion**
For U/S detection of knee effusion, it is best to place the patient in a supine position, with a pillow below the knee, and the knee either extended or slightly bent. The transducer is placed longitudinally along the axis of quadriceps tendon (Fig. 7-7A). The suprapatellar recess is located between the quadriceps tendon and the prefemoral fat (Fig. 7-7B). The normal suprapatellar recess is slit-like, with the thickness no more than 2 mm (32). Knee effusion is demonstrated as a distended suprapatellar recess with anechoic space (Fig. 7-7C), which is easily compressed, and without vascularity on power Doppler imaging. In repeated injuries, synovium hypertrophy can occur within the suprapatellar recess. The enlarged synovium often appears as a hyperechoic mass, either attached to the wall of the suprapatellar recess or floating within it (Fig. 7-7D). As in other scenarios, increased synovial vascularity indicates active inflammation. If a large amount of effusion is noted after an injury, one should suspect internal knee derangement, which may be better diagnosed with MRI and arthroscopy (33).

**Patellar Tendonitis**
Patients with patellar tendinitis (also known as “jumper knee”) may complain of anterior knee pain while jumping or going down stairs. Local tenderness over the patellar tendon is a common manifestation. When scanning the patellar tendon, the patient should be placed in supine position, with the patellar tendon tightened by flexing the knee to 60 to 80 degrees (Fig. 7-8A). The normal longitudinal patellar tendon should appear as parallel, fibrillar, echogenic structures (Fig. 7-8B). Commonly observed sonographic findings of patellar tendinitis include loss of fibrillar pattern, reduced echogenicity, and tendon hypertrophy in comparison with the contralateral knee (33). This is demonstrated in Figure 7-8C. Increased vascularity on power Doppler examination (Fig. 7-8D) suggests an acute inflammatory process. Studies have also shown that the presence of abnormal sonographic findings in the patellar tendon, even in the absence of subjective symptomatology, is associated with the higher risk of future patellar tendon injury (34).

**Muscular Injury**
Muscular injury is quite common in athletes. It may result from direct trauma or may sometimes be caused by overstretching, especially with insufficient warm-up. On U/S, normal muscle should appear as feather-like, longitudinal fibrils. Muscle tear will manifest as a disruption of muscular fibrils. An anechoic area within an injured muscle suggests hematoma
or effusion (Fig. 7-9A). On U/S, the severity of muscle injury can be divided into three grades. In grade I injury, no obvious muscular fibril rupture is observed, indicating a minor injury. Echogenicity of the injured muscle may be reduced due to swelling or bleeding. In grade II injury, fibrillary tear can be observed, with disruption of the normal muscular texture. An anechoic area due to either local hematoma or effusion is often apparent. Sometimes, the ruptured end of muscular fibrils floats within the local effusion and is referred to as the “bell clapper sign.” In grade III injury, the muscle tear is unfortunately complete. The rounded end of the ruptured “stump” can be seen, and the muscle gap will widen during stretching. An example of a complete quadriceps tear is seen in Figure 7-9B.

**Posterior Cruciate Ligament Injury**

Unlike the anterior cruciate ligament, the posterior cruciate ligament (PCL) can be clearly visualized on U/S examination (35). When scanning the PCL, the patient should be placed in prone position, with the knee extended. The transducer is located between the lateral margin of the medial femoral condyle and the midportion of the tibial intercondylar (Fig. 7-10A). The normal PCL should appear as a hypoechoic, fan-shaped structure (Fig. 7-10B). Rupture is suspected when the thickness of PCL at the tibial spine is greater than 10 mm and with a waxy posterior margin (36) (Fig. 7-10C,D).

The above-mentioned examples present a limited selection of common musculoskeletal pathologies encountered by practicing physiatrists. Together they serve as a brief introduction of the clinical applications of musculoskeletal U/S. There are numerous other conditions that U/S can be used for diagnostic, therapeutic, and research purposes, such as for evaluation of the median nerve in carpal tunnel syndrome (37,38). The integrity of ankle ligaments and tendons can also be evaluated with U/S (39,40). Moreover, U/S characteristics of plantar fasciitis have been well established, and injection of steroid under U/S guidance had been shown to be effective (41). There are many other applications of U/S that are beyond the scope of this chapter. Readers are encouraged to refer to musculoskeletal U/S textbooks for more thorough descriptions (42).
FIGURE 7-8. Patellar tendinitis. A: Examination of patellar tendon: the knee is flexed to 60 to 80 degrees to stretch the patellar tendon. B: Normal sonogram of patellar tendon shows the typical hyperechoic fibrillar pattern. C: Sonogram of patellar tendon tendinopathy reveals reduced echogenicity and increased thickness. D: Increased vascularity of patellar tendon is revealed, indicating active inflammation of the patellar tendon.

FIGURE 7-9. Muscle tear. A: Anechoic gap is seen within the quadriceps tendon, representing a tear. The round end of the ruptured muscle is well-defined on sonogram (arrows). B: A wider view of the ruptured quadriceps muscle reveals the extent of the injury.
CONCLUSION

The ability to perform dynamic examinations with real-time visualization and rapid side-to-side comparisons makes U/S eminently suitable for the diagnosis of musculoskeletal pathologies. In addition to assisting with initial diagnosis, U/S images can also be used in guiding needle injections and in quantifying the recovery process. While helpful, musculoskeletal U/S is, by no means, a replacement for other diagnostic methods such as MRI and arthroscopy. Instead, it should be viewed as an extension of our physical examination.

Limitations of U/S include its dependence on the experience of the operator and a relatively long learning curve. Adequate training and supervision may minimize these limitations; however, training opportunities and certification for musculoskeletal U/S are currently not standardized. The paucity of widely accepted standard examination protocols accentuates the variability of results attributable to operator dependence. Despite these limitations, U/S, when used appropriately, provides an important adjunct to physical examination to aid in the diagnosis and thus in the treatment of common musculoskeletal pathologies.
REFERENCES

Principles and Applications of Measurement Methods

Objective measurement provides a scientific basis for communication between professionals, documentation of treatment efficacy, and scientific credibility within the medical community. Federal, state, private third-party payer, and consumer organizations increasingly are requiring objective evidence of improvement as an outcome of treatment. Empirical clinical observation is no longer an acceptable method without objective data to support clinical decision making. The lack of reliability of clinicians’ unaided measurement capabilities is documented in the literature (1–7), further supporting the importance of objective measures. In addition, comparison of alternative evaluation or treatment methods, when more than one possible choice is available, requires appropriate use of measurement principles (8–11).

Clinicians and clinical researchers use measurements to assess characteristics, functions, or behaviors thought to be present or absent in specific groups of people. The application of objective measures uses structured observations to compare performances or characteristics across individuals (i.e., to discriminate), or within individuals over time (i.e., to evaluate), or for prognostication based on current status (i.e., to predict) (12,13). It is important to understand the principles of measurement and the characteristics of good measures to be an effective user of the tools. Standards for implementation of tests and measures have been established within physical therapy (14,15), psychology (16), and medical rehabilitation (17) to address quality improvement and ethical issues for the use of clinical measures.

The purpose of this chapter is to discuss the basic principles of tests and measurements and to provide the reader with an understanding of the rationale for assessing and selecting measures that will provide the information required to interpret test results properly. A standardized test is a test administered and scored in a consistent manner. The tests are designed in such a way that the questions, conditions for administering, scoring procedures, and interpretations are consistent and are administered and scored in a predetermined, standard manner. A critical starting point is to define what is to be measured, for what purpose, and at what cost. Standardized measurements meeting these criteria should then be assessed for reliability and validity pertinent to answering the question or questions posed by the user. Measurements that are shown not to be valid or reliable provide misleading information that is ultimately useless (18).

The initial section of this chapter discusses the psychometric parameters used to evaluate tests and measures. Principles of evaluation, testing, and interpretation are detailed in the second section. The fourth section provides guidelines for objective measurement when a standardized test is not available to measure the behavior, function, or characteristic of interest.

The complexity and diversity of the tests and measures used in rehabilitation medicine clinical practice and research preclude itemized description in a single chapter. A prerequisite for a measure to be objective requires that adequate levels of reliability have been demonstrated (18). Measures that have acceptable levels of reliability must also be shown to have appropriate types of validity to ultimately be labeled as objective. It is, therefore, imperative that the user be able to recognize the limitations of tests he or she employs to avoid inadvertent misuse or misinterpretation of test results.

PSYCHOMETRIC PARAMETERS USED TO EVALUATE TESTS AND MEASURES

The methods developed primarily in the psychology literature to evaluate objective measures generally are applicable to the standardized tests and instruments used in rehabilitation medicine. The topics discussed in this section are the foundation for all useful measures. Measurement tools must have defined levels of measurements for the trait or traits to be assessed and a purpose for obtaining the measurements. Additionally, tests and measures need to be practical, reliable, and valid.

Levels of Measurement

Tests and measures come in multiple forms because of the variety of parameters measured in clinical practice and research.

Despite the seemingly overwhelming number of measures, there are classified levels of measurement that determine how test results should be analyzed and interpreted (19). The four basic levels of measurement data are nominal, ordinal, interval, and ratio. Nominal and ordinal scales are used to classify discrete measures because the scores produced fall into discrete categories. Interval and ratio scales are used to classify continuous
measures because the scores produced can fall anywhere along a continuum within the range of possible scores.

A nominal scale is used to classify data that do not have a rank order. The purpose of a nominal scale is to categorize people or objects into different groups based on a specific variable. An example of nominal data is diagnosis.

Ordinal data are operationally defined to assign individuals to categories that are mutually exclusive and discrete. The categories have a logical hierarchy, but it cannot be assumed that the intervals are equal between each category, even if the scale appears to have equal increments. Ordinal scales are the most commonly used level of measurement in clinical practice. Examples of ordinal scales are the manual muscle test scale (20–24) and functional outcome measures (e.g., functional independence measure [FIM]) (25).

Interval data, unlike nominal and ordinal scales, are continuous. An interval scale has sequential units with numerically equal distances between them. Interval data often are generated from quantitative instrumentation as opposed to clinical observation. It is important to note that it is statistically possible to transform ordinal data into interval data using Rasch logit scale methods. This has most notably been utilized in medical rehabilitation for analysis of FIM data. Detailed information regarding how to perform this transformation can be found elsewhere (26).

Examples of interval measurements are range-of-motion scores reported in degrees and the visual analogue pain scale (continuum from 0 to 10).

A ratio scale is an interval scale on which the zero point represents a total absence of the quantity being measured. An example is force scores obtained from a quantitative muscle strength–testing device.

Analysis of nominal and ordinal scales requires special consideration to avoid misinference from test results (27,28). The major controversies surrounding the use of these scales are the problems of unidimensionality and whether scores of items and subtests can be summed to provide an overall score. Continuous scales have a higher sensitivity of measurement and allow more rigorous statistical analyses to be performed. Sensitivity can be defined as the proportion of people with a condition/trait who test positive for that condition/trait.

Purpose of Testing
After the level of the measurement has been selected, the purpose of testing must be examined. Tests generally serve one of two purposes: screening or in-depth assessment of specific traits, behaviors, functions, or prognosis.

Screening Tests
Screening tests have three possible applications:

1. To discriminate between “suspect” and “normal” patients
2. To identify people needing further assessment
3. To assess a number of broad categories superficially

One example of a screening test is the Test of Orientation for Rehabilitation Patients, administered to individuals who are confused or disoriented secondary to traumatic brain injury, cerebrovascular accident, seizure disorder, brain tumor, or other neurologic events (29–32). This test screens for orientation to person and personal situation, place, time, schedule, and temporal continuity. Another well-developed screening test is the Miller Assessment for Preschoolers (MAP) (33). This test screens preschoolers for problems in the following areas: sensory and motor, speech and language, cognition, behaviors, and visual-motor integration.

The advantages of screening tests are that they are brief and sample a broad range of behaviors, traits, or characteristics. They are limited, however, because of an increased frequency of false-positive results that is due to the small sample of behaviors obtained. Screening tests should be used cautiously for diagnosis, placement, or treatment planning. They are used most effectively to indicate the need for more extensive testing and treatment of specific problem areas identified by the screening assessment.

Assessment Tests
Assessment tests have five possible applications:

1. To evaluate specific behaviors in greater depth
2. To provide information for planning interventions
3. To determine placement into specialized programs
4. To provide measurements to monitor progress
5. To provide information regarding prognosis

An example of an assessment measure is the Boston Diagnostic Aphasia Examination (34). The advantages of assessment measures are that they have a lower frequency of false-positive results; they assess a representative set of behaviors; they can be used for diagnosis, placement, or treatment planning; and they provide information regarding the functional level of the individual tested. The limitations are that an extended amount of time is needed for testing, and they generally require specially trained personnel to administer, score, and interpret the results.

Criterion-Referenced Versus Norm-Referenced Tests
Proper interpretation of test results requires comparison with a set of standards or expectations for performance. There are two basic types of standardized measures: criterion-referenced and norm-referenced tests.

Criterion-Referenced Tests
Criterion-referenced tests are those for which the test score is interpreted in terms of performance on the test relative to the continuum of possible scores attainable (18). The focus is on what the person can do or what he or she knows rather than how he or she compares with others (35). Individual performance is compared with a fixed expected standard rather than a reference group. Scores are interpreted based on absolute criteria, for example, the total number of items successfully completed. Criterion-referenced tests are useful to discriminate between successive performances of one person. They are
conducted to measure a specific set of behavioral objectives. The Tufts Assessment of Motor Performance (which has undergone further validation work and has been renamed the Michigan Modified Performance Assessment) is an example of a criterion-referenced test (36–40). This assessment battery measures a broad range of physical skills in the areas of mobility, activities of daily living, and physical aspects of communication.

**Norm-Referenced Tests**

Norm-referenced tests use a representative sample of people who are measured relative to a variable of interest. Norm referencing permits comparison of a single person’s measurement with those scores expected for the rest of the population. The normal values reported should be obtained from, and reported for, clearly described populations. The normal population should be the same as those for whom the test was designed to detect abnormalities (35). Reports of norm-referenced test results should use scoring procedures that reflect the person’s position relative to the normal distribution (e.g., percentiles, standard scores). Measures of central tendency (e.g., mean, median, mode) and variability (e.g., standard deviation, standard error of the mean) also should be reported to provide information on the range of normal scores, assisting with determination of the clinical relevance of test results. An example of a norm-referenced test is the Peabody Developmental Motor Scale (41). This developmental test assesses fine and gross motor domains. Test items are classified into the following categories: grasp, hand use, eye-hand coordination, manual dexterity, reflexes, balance, nonlocomotor, locomotor, and receipt and propulsion of objects.

**Practicality**

A test or instrument should ideally be practical, easy to use, insensitive to outside influences, inexpensive, and designed to allow efficient administration (42). For example, it is not efficient to begin testing in a supine position, switch to a prone position, and then return to supine. Test administration should be organized to complete all testing in one position before switching to another. Instructions for administering the test should be clear and concise, and scoring criteria should be clearly defined. If equipment is required, it must be durable and of good quality. Qualifications of the tester and additional training required to become proficient in test administration should be specified. The time to administer the test should be indicated in the test manual. The duration of the test and level of difficulty need to be appropriate relative to the attention span and perceived capabilities of the patient being tested. Finally, the test manual should provide summary statistics and detailed guidelines for appropriate use and interpretation of test scores based on the method of test development.

**Reliability and Agreement**

A general definition of reliability is the extent to which a measurement provides consistent information (i.e., is free from random error). Granger et al. (43) provide the analogy “it may be thought of as the extent to which the data contain relevant information with a high signal-to-noise ratio versus irrelevant static confusion.” By contrast, agreement is defined as the extent to which identical measurements are made. Reliability and agreement are distinctly different concepts and are estimated using different statistical techniques (44). Unfortunately, these concepts and their respective statistics often are treated synonymously in the literature.

The level of reliability is not necessarily congruent with the degree of agreement. It is possible for ratings to cluster consistently toward the same end of the scale, resulting in high-reliability coefficients, and yet these judgments may or may not be equivalent. High reliability does not indicate whether the raters absolutely agree. It can occur concurrently with low agreement when each rater scores patients differently, but the relative differences in the scores are consistent for all patients rated. Conversely, low reliability does not necessarily indicate that raters disagree. Low-reliability coefficients can occur with high agreement when the range of scores assigned by the raters is restricted or when the variability of the ratings is small (i.e., in a homogeneous population). In instances in which the scores are fairly homogeneous, reliability coefficients lack the power to detect relationships and are often depressed, even though agreement between ratings may be relatively high.

The reader is referred to Tinsley and Weiss for examples of these concepts (45). Both reliability and agreement must be established on the target population or populations to which the measure will be applied, using typical examiners. There are five types of reliability and agreement:

1. Interrater
2. Test-retest
3. Intertrial
4. Alternate form
5. Population specific

Each type will be discussed below, along with indications for calculating reliability versus agreement and their respective statistics.

**Interrater Reliability and Agreement**

Interrater or interobserver agreement is the extent to which independent examiners agree exactly on a patient’s performance. By contrast, interrater reliability is defined as the degree to which the ratings of different observers are proportional when expressed as deviations from their means; that is, the relationship of one rated person to other rated people is the same, although the absolute numbers used to express the relationship may vary from rater to rater (45). The independence of the examiners in the training they receive and the observations they make is critical in determining interrater agreement and reliability. When examiners have trained together or confer when performing a test, the interrater reliability or agreement coefficient calculated from their observations may be artificially inflated.

An interrater agreement or reliability coefficient provides an estimate of how much measurement error can be expected in scores obtained by two or more examiners who have...
Intertrial reliability assesses the consistency of one examiner's performance across different sessions. An acceptable level of intertrial reliability or agreement is essential for comparison of test results obtained from different clinical centers. Intertrial agreement or reliability is a basic criterion for a measure to be called objective. If multiple examiners consistently obtain the same absolute or relative scores, then it is much more likely that the score is a function of the measure, rather than of the collective subjective bias of the examiners (18).

Pure intertrial agreement and reliability are determined by having one examiner administer the test while the other examiner or examiners observe and independently score the patient's performance at the same point in time. When assessing some parameters, when the skill of the examiner administering the test plays a vital role (e.g., sensory testing, range-of-motion testing) or when direct observation of each examiner is required (e.g., strength), it is impossible to assess pure interrater agreement and reliability. In these instances, each examiner must test the individual independently. Consequently, these interrater measures are confounded by factors of time and variation in patient performance.

Test-Retest Reliability and Agreement
Test-retest agreement is defined as the extent to which a patient receives identical scores during two different test sessions when rated by the same examiner. By contrast, test-retest reliability assesses the degree of consistency in how a person's score is rank ordered relative to other people tested by the same examiner during different test sessions. Test-retest reliability is the most basic and essential form of reliability. It provides an estimate of the variation in patient performance on a different test day, when retested by the same examiner. Some of the errors in a test-retest situation also may be attributed to variations in the examiner's performance. It is important to determine the magnitude of day-to-day fluctuations in performance so that true changes in the parameters of interest can be determined. Variability of the test or how it is administered should not be the source of observed changes over time. Additionally, with quantitative measuring instruments, the examiner must be knowledgeable in the method of and frequency required for instrument calibration.

The suggested test-retest interval is 1 to 3 days for most physical measures and 7 days for maximal effort tests in which muscle fatigue is involved (46). The test-retest interval should not exceed the expected time for change to occur naturally. The purpose of an adequate but relatively short interval is to minimize the effects of memory, practice, and maturation or deterioration on test performance (47).

Intertrial Reliability and Agreement
Intertrial agreement provides an estimate of the stability of repeated scores obtained by one examiner within a test session. Intertrial reliability assesses the consistency of one examiner rank-ordering repeated trials obtained from patients using the same measurement tool and standardized method for testing and scoring results within a test session. Intertrial agreement and reliability also are influenced by individual performance factors such as fatigue, motor learning, motivation, and consistency of effort. Intertrial agreement and reliability should not be confused with test-retest agreement and reliability. The latter involves test sessions usually separated by days or weeks as opposed to seconds or minutes for intertrial agreement and reliability. A higher level of association is expected for results obtained from trials within a test session than those from different sessions.

Alternate Form Reliability and Agreement
Alternate form agreement refers to the consistency of scores obtained from two forms of the same test. Equivalent or parallel forms are different test versions intended to measure the same traits at a comparable level of difficulty. Alternate form reliability refers to whether the parallel forms of a test rank order people's scores consistently relative to each other. A high level of alternate form agreement or reliability may be required if a person must be tested more than once and a learning or practice effect is expected. This is particularly important when one form of the test will be used as a pretest and a second as a posttest.

Population-Specific Reliability and Agreement
Population-specific agreement and reliability assess the degree of absolute and relative reproducibility, respectively, that a test has for a specific group being measured (e.g., Ashworth scale scores for rating severity of spasticity from spinal cord injury). A variation of this type of agreement and reliability refers to the population of examiners administering the test (18).

Interpretation of Reliability and Agreement Statistics
Because measures of reliability and agreement are concerned with the degree of consistency or concordance between two or more independently derived sets of scores, they can be expressed in terms of correlation coefficients (35). The reliability coefficient is usually expressed as a value between 0 and 1, with higher values indicating higher reliability. Agreement statistics can range from −1 to +1, with +1 indicating perfect agreement, 0 indicating chance agreement, and negative values indicating less than chance agreement. The coefficient of choice varies, depending on the data type analyzed. The reader is referred to Bartko and Carpenter (48), Hartmann (49), Hollenbeck (50), Liebertau (51), and Tinsley and Weiss (45) for discussions of how to select appropriate statistical measures of reliability and agreement. Table 8-1 provides information on appropriate statistical procedures for calculating interrater and test-retest reliability and agreement for discrete and continuous data types. No definitive standards for minimum acceptable levels of the different types of reliability and agreement statistics have been established; however, guidelines for minimum levels are
A test score actually consists of two different components: the true score and the error score (35,54). A person’s true score is a hypothetical construct, indicating a test score that is unaffected by chance factors. The error score refers to the unwanted variation in the test score (55). All continuous scale measurements have a component of error, and no test is completely reliable. Consequently, reliability is a matter of degree. Any reliability coefficient may be interpreted directly in terms of percentage of score variance attributable to different sources (18). A reliability coefficient of 0.85 signifies that 85% of the variance in test scores depends on true variance in the trait measured and 15% depends on error variance.

### Specific Reliability and Agreement Statistics

There are several statistical measures for estimating inter-rater agreement and reliability. Four statistics commonly used to determine agreement are the frequency ratio, point-by-point agreement ratio, kappa (\(\kappa\)) coefficients, and Lawlis and Lu’s \(\chi^2\) and \(T\)-index statistics. For reliability calculations, the most frequently used correlation statistics are the Pearson product-moment (Pearson \(r\)) and intraclass correlation coefficients (ICCs). When determining reliability for dichotomous or ordinal data, specific ICC formulas have been developed. These nonparametric ICC statistics have been shown to be the equivalent of the weighted kappa (\(\kappa_w\)) (56–59). Consequently, the \(\kappa_w\) also can be used as an index of reliability for discrete data, and the values obtained can be directly compared with equivalent forms of ICCs (57). The method of choice for reliability and agreement analyses partially depends on the assessment strategy used (45,48,51,60). In addition to agreement and reliability statistics, standard errors of measurement (SEM) provide a clinically relevant index of reliability expressed in test score units. Each statistic is described below.

#### Frequency Ratio

This agreement statistic is indicated for frequency count data (47). A frequency ratio of the two examiners’ scores is calculated by dividing the smaller total by the larger total and multiplying by 100. This statistic is appealing because of its computational and interpretive simplicity. There are a variety of limitations, however. It only reflects agreement of the total number of behaviors scored by each observer; there is no way to determine whether there is agreement for individual responses using a frequency ratio. The value of this statistic may be inflated if the observed behavior occurs at high rates (60). There is no meaningful lower bound of acceptability (49).

#### Point-by-Point Agreement Ratio

This statistic is used to determine if there is agreement on each occurrence of the observed behavior. It is appropriate when there are discrete opportunities for the behavior to occur or for distinct response categories (47,61,62). To calculate this ratio, the number of agreements is totaled by determining the concurrence between observers regarding the

---

**TABLE 8.1**  
**Interrater Reliability, Test-Retest Reliability, and Agreement Analysis: Appropriate Statistics and Minimum Acceptable Levels**

<table>
<thead>
<tr>
<th>Data Type</th>
<th>Reliability Analysis</th>
<th>Agreement Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Appropriate Statistic</td>
<td>Level</td>
</tr>
<tr>
<td>Discrete</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nominal</td>
<td>ICC or (\kappa_w)</td>
<td>&gt;0.75</td>
</tr>
<tr>
<td>Ordinal</td>
<td>ICC</td>
<td>&gt;0.75</td>
</tr>
<tr>
<td>Continuous</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interval</td>
<td>ICC</td>
<td>&gt;0.75</td>
</tr>
<tr>
<td>Ratio</td>
<td>ICC</td>
<td>&gt;0.75</td>
</tr>
</tbody>
</table>

References: ICC: discrete (47,55), ordinal (47), continuous (44,47), minimal acceptable level (56); Cohen’s \(\kappa = \kappa\) (44,47,57,58), \(\kappa_w\) (47,59,60), \(\kappa_w\) equivalence with ICC for reliability analysis of minimal data (61–64), minimal acceptable level (65); Lawlis and Lu’s \(\chi^2\) and \(T\)-statistical and minimal level (43,44). ICC, intraclass correlation; \(\kappa\), kappa; \(\kappa_w\), weighted kappa; \(T\), \(T\) index.

---

As discussed earlier, these are distinctly different concepts and require separate statistical analysis. Several factors must be considered to determine the relative importance of each. Decisions that carry greater weight or impact for the people being assessed may require more exact agreement. If the primary need is to assess the relative consistency between raters, and exact agreement is less critical, then a reliability measure alone is a satisfactory index. By contrast, whenever the major interest is either the absolute value of the score or the meaning of the scores as defined by the points on the scale (e.g., criterion-referenced tests), agreement should be reported in addition to the reliability (45). Scores generated from instrumentation are expected to have a higher level of reliability or agreement than scores obtained from behavioral observations.
presence or absence of observable responses during a given trial, recording interval, or for a particular behavior category. Disagreements are defined as instances in which one observer records a response and the other observer does not. The point-by-point agreement percentage is calculated by dividing the number of agreements by the number of agreements plus disagreements and multiplying by 100 (62). Agreement generally is considered to be acceptable at a level of 0.80 or above (62).

The extent to which observers are found to agree is partially a function of the frequency of occurrence of the target behavior and of whether occurrence and/or nonoccurrence agreements are counted (61). When the rate of the target behavior is either very high or very low, high levels of interobserver agreement are likely for occurrences or nonoccurrences, respectively. Consequently, if the frequency of either occurrences or nonoccurrences is high, a certain level of agreement is expected simply owing to chance. In such cases, it is often recommended that agreements be included in the calculation only if at least one observer recorded the occurrence of the target behavior. In this case, intervals during which none of the observers records a response are excluded from the analysis. It is important to identify clearly what constitutes an agreement when reporting point-by-point percentage agreement ratios because the level of reliability is affected by this definition.

**Kappa Coefficients**

The κ coefficient provides an estimate of agreement between observers, corrected for chance agreement. This statistic is preferred for discrete categorical (nominal and ordinal) data because, unlike the two statistics discussed above, it corrects for chance agreements. In addition, percentage agreement ratios often are inflated when there is an unequal distribution of scores between rating categories. This often is the case in rehabilitation medicine, in which the frequency of normal characteristics is much higher than abnormal characteristics (63,64). By contrast, κ coefficients provide accurate estimates of agreement, even when scores are unequally distributed between rating categories (64).

Kappa coefficients are used to summarize observer agreement and accuracy, determine rater consistency, and evaluate scaled consistency among raters (60). Three conditions must be met to use κ:

1. The patients or research subjects must be independent.
2. The raters must independently score the patients or research subjects.
3. The rating categories must be mutually exclusive and exhaustive (63,64).

The general form of κ is a coefficient of agreement for nominal scales in which all disagreements are treated equally (45,48,51,65–68). The κw statistic was developed for ordinal data (48,51,67,70), in which some disagreements have greater gravity than others (e.g., the manual muscle testing scale, in which the difference between a score of 2 and 5 is of more concern than the difference between a score of 4 and 5). Refer to the references cited above for formulas used to calculate κ and κw.

Several other variations of κ have been developed for specific applications. The kappa statistic κm provides an overall measure of agreement, as well as separate indices for each subject and rating category (71). This form of κ can be applied in situations in which subjects are not all rated by the same set of examiners. The variation of κ described by Fleiss et al. is useful when there are more than two ratings per patient (58); a computer program is available to calculate this statistic (63). When multiple examiners rate patients and a measure of overall conjoint agreement is desired, the kappa statistic κm is indicated (72). Standard κ statistics treat all raters or units symmetrically (58). When one or more of the ratings are considered to be a standard (e.g., scores from an experienced rater), alternate analysis procedures should be used (72–74).

**Lawlis and Lu χ^2 and T Index**

These measures of agreement are recommended for continuous data (45). They permit the option of defining seriousness of disagreements among raters. A statistically significant χ^2 indicates that the observed agreement is greater than that expected owing to chance. The T index is used to determine whether agreement is low, moderate, or high. The reader is referred to Tinsley and Weiss (45) for a discussion of the indications for, calculation of, and interpretation of these statistics.

**Pearson Product-Moment Correlation Coefficient**

Historically, the Pearson r has been used commonly as an index of reliability. It has limited application, however, because it is a parametric statistic intended for use with continuous bivariate data. The generally accepted minimum level of this coefficient is 0.80; however, levels above 0.90 often are considered more desirable (35,52). The Pearson r provides only an index of the strength of the relationship between scores and is insensitive to consistent differences between scores. Consequently, a linear regression equation must be reported in addition to the Pearson r to indicate the nature of the relationship between the scores (18). Because the Pearson r is limited to the analysis of bivariate data, it is preferable to use an ICC to assess reliability because ICC can be used for either bivariate or multivariate data. The Pearson r and ICC will yield the same result for bivariate data (75).

**Intraclass Correlation Coefficients**

ICCs provide an index of variability resulting from comparing rating score error with other sources of true score variability (43,53,76). As indicated above, it is the coefficient of choice for reliability analyses. The ICC is based on the variance components from an analysis of variance (ANOVA), which includes not only the between-subject variance, as does the Pearson r, but also other situation-specific variance components, such as alternate test forms, maturation of subjects between ratings, and other sources of true mean differences.
in the obtained ratings (77). The individual sources of error can be analyzed to determine their percentage contribution to the overall error variance using generalizability analysis (54,55). For further information regarding the use of generalizability theory to distinguish between sources of error, the reader is referred to Brennan (78) and Cronbach et al. (79).

There are six different ICC formulas (55). The correct ICC formula is selected based on three factors:

1. The use of a one-way versus two-way ANOVA
2. The importance of differences between examiners’ mean ratings
3. The analysis of an individual rating versus the mean of several ratings (45,55)

Selection of the proper formula is critical and is based on the reliability study design (55,77,80). It is important to report which type of ICC is used to compute reliability because the calculations are not equivalent. Variations of the ICC formulas also exist for calculating ICCs using dichotomous (81) and ordinal (44) nonparametric data. The marginal distributions do not have to be equal, as was originally proposed for non-parametric ICCs (57). These nonparametric ICC formulas have been demonstrated to be equivalent to weighted κ coefficients, provided that the mean difference between raters is included as a component of variability and the rating categories can be ordered (57).

**Standard Error of Measurement**

It has been suggested that measurement error estimates are the most desirable index of reliability (18,35,76). The SEM is an estimate, in test score units, of the random variation of a person’s performance across repeated measures. The SEM is an expression of the margin of error between a person’s observed score and his or her true ability (47). The SEM is an important indicator of the sensitivity of the test to detect changes in a person’s performance over time.

The formula for the SEM is

$$SD \sqrt{1 - r_{uu}}$$

where $SD$ is the standard deviation of the test scores and $r_{uu}$ is the reliability coefficient for the test scores (35,46,76). Correlating scores from two forms of a test is one of several ways to estimate the reliability coefficient (76) and often is used in psychology when parallel forms of a test are available. In rehabilitation medicine, however, equivalent forms of a test often are not available. The test-retest reliability coefficient, therefore, is the coefficient of choice for calculating the SEM in most rehabilitation applications because the primary interest is in the variation of subject performance. The SEM is a relatively conservative statistic, requiring larger data samples (approximately 300 to 400 observations) in order to not overestimate the error (15).

It is best to report a test score as a range rather than as an absolute score. The SEM is used to calculate the range of scores (i.e., confidence interval) for a given person; that is, the person’s true performance ability is expected to fall within the range of scores defined by the confidence interval. A person’s score must fall outside of this range to indicate with confidence that a true change in performance has occurred. Based on a normal distribution, a 95% confidence interval would be approximately equal to the mean ± 2 SEM. A 95% confidence interval is considered best to use when looking for change over time. This rigorous level of confidence minimizes the likelihood of a type I error (i.e., there is only a 5% chance that differences between scores obtained from a given person during different test sessions will not fall within the 95% confidence interval upper and lower values). Consequently, there is less than a 5% chance that differences between scores exceeding the upper end of the confidence interval are due to measurement error (i.e., they have a 95% chance of representing a true change in performance).

**Factors Affecting Reliability**

There are four sources of measurement error for intrarater reliability (18,46):

1. Lack of agreement among scorers
2. Lack of consistent performance by the individual tested
3. Failure of the instrument to measure consistently
4. Failure of the examiner to follow the standardized procedures to administer the test

Threats to test-retest reliability similarly are caused by four factors:

1. The instrument
2. The examiner
3. The patient
4. The testing protocol

Sources and prevention of examiner error will be discussed in the section on “Principles of Evaluation, Testing, and Interpretation.”

There are several factors conducive to good reliability of a measure (46). These factors are the power to discriminate among ability groups; sufficient time allotted so that each patient can show his or her best performance without being penalized for an unrepresentative poor trial; test organization to optimize examinee performance; and test administration and scoring instructions that are clear and precise. Additionally, the testing environment should support good performance, and the examiner must be competent in administering the test. For tests designed to be appropriate for a wide age range, reliability should be examined for each age level rather than for the group as a whole (54).

In summary, reliability and agreement are essential components to any objective measurement. Measurements lacking test-retest reliability contain sufficient error as to be useless because the data obtained do not reflect the variable measured (18). Reliability is an important component of validity, but good reliability or agreement does not guarantee that a measure is valid. A reliable measurement is consistent, but not
necessarily correct. However, a measurement that is unreliable cannot be valid.

Validity
Validity is defined as the accuracy with which a test measures that which it is intended to measure. Application of the concept of validity refers to the appropriateness, meaningfulness, and usefulness of a test for a particular situation (18). Validity is initially investigated while a test or instrument is being developed and confirmed through subsequent use. Four basic aspects of validity will be discussed: content, construct, criterion-related, and face validity.

Content Validity
Content validity is the systematic examination of the test content to determine if it covers a representative sample of the behavior domain to be measured. It should be reported in the test manual as descriptive information on the skills covered by the test, number of items in each category, and rationale for item selection. Content validity generally is evidenced by the opinion of experts that the domain sampled is adequate. There are two primary methods that the developer of a test can use for obtaining professional opinions about the content validity of an instrument (82). The first is to provide a panel of experts with the items from the test and request a determination of what the battery of items is measuring. The second method requires providing not only the test items but also a list of test objectives so that experts can determine the relationship between the two. For statistical analysis of content validity, the reader is referred to Thorn and Deitz (83).

Construct Validity
Construct validity refers to the extent to which a test measures the theoretical construct underlying the test. Construct validity should be obtained whenever a test purports to measure an abstract trait or theoretical characteristics about the nature of human behavior such as intelligence, self-concept, anxiety, school or work readiness, or perceptual organization. The following five areas must be considered with regard to construct validity in test instruments (35, 82).

Age Differentiation
Any developmental changes in children or changes in performance due to aging must be addressed as part of the test development.

Factor Analysis
Factor analysis is a statistical procedure that can be performed on data obtained from testing. The purpose of factor analysis is to simplify the description of behavior by reducing an initial multiplicity of variables to a few common underlying factors or traits that may or may not be pertinent to the construct or constructs that the test was originally designed to measure. The reader is referred to Cronbach (76), Wilson et al. (84), Wright and Masters (85), and Wright and Stone (86) for in-depth discussions of factor analysis. The more recent development of confirmatory factor analysis (87, 88) overcomes the relative arbitrariness of traditional factor analysis methods. Confirmatory factor analysis differs from traditional factor analysis in that the investigator specifies, before analysis, the measures that are determined by each factor and which factors are correlated. The specified relationships are then statistically tested for goodness of fit of the proposed model compared with the actual data collected. Confirmatory factor analysis is therefore a more direct assessment of construct validity than is traditional factor analysis. Rasch modeling (89) is a further expansion on confirmatory factor analysis methods for the purpose of establishing construct validity of a measurement tool. The family of Rasch measurement models provides the means for constructing interval measures from raw data (26). Rasch models start with a carefully thought-out and systematically implemented analogy used to facilitate the construction of the concepts of the measurement tool in concrete terms and then use a developmental pathway analogy to develop the Rasch concepts of unidimensionality, fit, difficulty/ability estimation and error, locations for item difficulties, and locations for person abilities.

Internal Consistency
In assessing the attributes of a test, it is helpful to examine the relationship of subscales and individual items to the total score. This is especially important when the test instrument has many components. If a subtest or item has a very low correlation with the total score, the test developer must question the subtest’s validity in relation to the total score. This technique is most useful for providing confirmation of the validity of a homogeneous test. A test that measures several constructs would not be expected to have a high degree of internal consistency. For dichotomous data, the Kuder-Richardson statistic is used to calculate internal consistency (35). Cronbach’s coefficient alpha (\( \alpha \)) is recommended when the measure has more than two levels of response (35). The minimum acceptable level generally is set at 0.70 (90).

Convergent and Divergent Validity
Construct validity is evidenced further by high correlations with other tests that purport to measure the same constructs (i.e., convergent validity) and low correlations with measures that are designed to measure different attributes (i.e., divergent validity). It is desirable to obtain moderate levels of convergent validity, indicating that the two measures are not measuring identical constructs. If the new test correlates too highly with another test, it is questionable whether the new test is necessary because either test would suffice to answer the same questions. Moderately high but significant correlations indicate good convergent validity, but with each test still having unique components. Good divergent validity is demonstrated by low and insignificant correlations between
two tests that measure theoretically unrelated parameters, such as an activity of daily living assessment and a test of expressive language ability.

**Discriminant Validity**
If two groups known to have different characteristics can be identified and assessed by the test, and if a significant difference between the performances of the two groups is found, then incisive evidence of discriminant validity is present.

**Criterion-Related Validity**
Criterion-related validity includes two subclasses of validity: concurrent validity and predictive validity (35,47). The commonality between these subclasses of validity is that they refer to multiple measurement of the same construct. In other words, the measure in question is compared with other variables or measures that are considered to be accurate measures of the characteristics or behaviors being tested. The purpose is to use the second measure as a criterion to validate the first measure.

Criterion-related validity can be assessed statistically, providing clear guidelines as to whether a measure is valid. Frequently, the paired measurements from the tests under comparison have different values. The nature of the relationship is less important than the strength of the relationship (18). Ottenbacher and Tomchek (91) showed that the limits of agreement technique provided the most accurate measurement error when comparing test results versus other statistics frequently used for such comparisons.

**Concurrent Validity**
Concurrent validity deals with whether an inference is justifiable at the present time. This is typically done by comparing results of one measure against some criteria (e.g., another measure or related phenomenon). If the correlation is high, the measure is said to have good concurrent validity. Concurrent validity is relevant to tests used for diagnosis of existing status, rather than predicting future outcome.

**Predictive Validity**
Predictive validity involves a measure’s ability to predict or forecast some future criterion. Examples include performance on another measure in the future, prognostic reaction to an intervention program, or performance in some task of daily living. Predictive validity is difficult to establish and often requires collection of data over an extended period of time after the test has been developed. Hence, very few measures used in rehabilitation medicine have established predictive validity. A specific subset of predictive validity that is important to rehabilitation medicine practice is ecological validity. This concept involves the ability to identify impairments, functional limitations, and performance deficits within the context of the person's own environment. Measures with good concurrent validity sometimes are presumed to have good predictive validity, but this may not be a correct assumption.

Unless predictive validity information exists for a test, extreme caution should be exercised in interpreting test results as predictors of future behavior or function.

**Face Validity**
Face validity is not considered to be an essential component of the validity of a test or measure. It reflects only whether a test appears to measure what it is supposed to, based on the personal opinions of those either taking or giving the test (92). A test with high face validity has a greater likelihood of being more rigorously and carefully administered by the examiner, and the person being tested is more likely to give his or her best effort. Although it is not essential, in most instances, face validity is still an important component of test development and selection. Exceptions include personality and interest tests when the purpose of testing is concealed to prevent patient responses from being biased.

**Summary**
The information discussed in this section provides the basis for critically assessing available tests and measures. The scale of the test or instrument should be sufficiently sophisticated to discriminate adequately between different levels of the behavior or function being tested. The purpose for testing must be identified, and the test chosen should have been developed for this purpose. The measure selected should be practical from the standpoint of time, efficiency, budget, equipment, and the population being tested. Above all, the measure must have adequate reliability, agreement, and validity for the specific application it is selected. Reliability, agreement, and validity are important for both clinical and research applications. The power of statistical tests depends on adequate levels of reliability, agreement, and validity of the dependent measures (93). Consequently, it is essential that adequate levels of reliability, agreement, and validity be assessed and reported for dependent measures used in research studies.

For additional information on the test development process, the reader is referred to Miller (94). For information on the principles of tests and measurements, the reader is referred to Anastasi and Urbina (35), Baumgartner and Jackson (46), Cronbach (76), Safrir (52), Rothstein (18), Rothstein and Echternach (15), and Verducci (53).

Identification of the most appropriate test for a given application, based on the psychometric criteria discussed above, does not guarantee that the desired information will be obtained. Principles of evaluation, testing, and interpretation must be followed to optimize objective data acquisition.

**PRINCIPLES OF EVALUATION, TESTING, AND INTERPRETATION**
Systematic testing using standardized techniques is essential to quantify a patient’s status objectively. Standardized testing is defined as using specified test administration and scoring
Procedures, under the same environmental conditions, with consistent directions (35,47). Standardized testing is essential to permit comparison of test results for a given person over time and to compare test scores between patients (92). In addition, consistent testing techniques facilitate interdisciplinary interpretation of clinical findings among rehabilitation professionals and minimize duplication of evaluation procedures.

**Examiner Qualifications**  
Assessments using objective instrumentation or standardized tests must be conducted by examiners who have appropriate training and qualifications (14,16,17,35,92,95). The necessary training and expertise vary with the type of instrument or test used. The characteristics common to most rehabilitation medicine applications will be discussed. Examiners must be thoroughly familiar with standardized test administration, scoring, and interpretation procedures. Training guidelines specified in the published test manual must be strictly adhered to. A skilled examiner is aware of factors that might affect test performance and takes the necessary steps to ensure that the effects of these factors are minimized. Interrater reliability needs to be attained at acceptable levels with examiners who are experienced in administering the test to ensure consistency of test administration and scoring.

Examiners also must be knowledgeable about the instruments and standardized tests available to assess parameters of interest. They need to be familiar with relevant research literature, test reviews, and the technical merits of the appropriate tests and measures (14,16,17,35). From this information, examiners should be able to discern the advantages, disadvantages, and limitations of using a particular test or device. Based on the purpose of testing and characteristics of the person being assessed, examiners need to be able to select and justify the most appropriate assessment method from the available options.

When interpreting test results, examiners must be sensitive to factors that may have affected test performance (35). Conclusions and recommendations should be based on a synthesis of the person’s scores, the expected measurement error, any factors that might have influenced test performance, the characteristics of the given person compared with those of the normative population, and the purpose of testing versus the recommended applications of the test or instrument. Written documentation of test results and interpretation should include comments on any potential influence of the above factors.

**Examiner Training**  
Proper training of examiners is critical to attaining an acceptable level of interrater reliability for test administration and scoring (92). Examiners should be trained to minimize later decrements in performance (60). Training methods should be documented carefully so that they can be replicated by future examiners.

**Training Procedures**  
As part of their training, examiners should read the test manual and instructions carefully. Operational definitions and rating criteria need to be memorized verbatim (96). A written examination should be administered to document the examiners’ assimilation of test administration and scoring procedures (60). This information should be periodically reviewed to produce close adherence to the standardized protocol. It is helpful for examiners to view a videotape of an experienced examiner conducting the test. If test administration and scoring techniques need to be adjusted for the varying abilities of the target population (e.g., children of different age levels), the experienced examiner should be observed testing a representative sample from the target population to demonstrate the various testing, scoring, and interpretation procedures.

Videotapes also are useful to clarify scoring procedures and establish consistency of scoring between and within raters (67,92). Once scoring procedures have been reviewed adequately, interrater reliability can be established by having trainees view several patients on videotape and then compare their scores with those from an experienced examiner. Scoring discrepancies should be discussed, and trainees should continue to score videotaped segments until 100% agreement is established with an experienced examiner (96). Intrarater consistency of scoring also can be established by having an individual examiner score the same videotape on multiple occasions. Sufficient time should elapse between multiple viewings so examiners do not recall previous ratings.

For assessments that involve multiple trials (e.g., strength assessments), intertrial reliability can be calculated to provide a measure of the examiner’s consistency of administering multiple trials within a given test session. As was mentioned previously, intertrial reliability also is influenced by factors such as fatigue, motor learning, motivation, and the stability of performance over a short period of time. Multiple trials administered during a given session generally are highly correlated; thus, intertrial reliability coefficients are expected to be very high. Although this measure provides feedback on consistency in administering multiple trials, it should not be considered a substitute for establishing other types of reliability during the training phase.

**Establishing Procedural Reliability**  
Procedural reliability is defined as the reliability with which standardized testing and scoring procedures are applied. As part of training, examiners should be observed administering and scoring the test on a variety of people with characteristics similar to those of the target population (92). Procedural reliability should be established by having an experienced examiner observe trainees to determine if the test is being administered and scored according to the standardized protocol. Establishing procedural reliability greatly increases the likelihood that the observed changes in performance reflect true changes in status and not alterations in examiner testing or scoring methods. Unfortunately, this type of reliability often is neglected. According to Billingsley et al. (97), failure to assess procedural reliability poses a threat to both the internal and the external validity of assessments.
Procedural reliability is assessed by having an independent observer check off whether each component of an assessment is completed according to the standardized protocol while viewing a live or videotaped assessment. Specific antecedent conditions, commands, timing of execution, and positioning are monitored, and any deviations are noted. Procedural reliability is calculated as a percentage of correct behaviors (97). Checklists should include all essential components of the standardized protocol. An example of a procedural reliability checklist for selected items on the MAP is provided in Figure 8-1 (92). In this example, the checklist varies for each item administered. Another example of procedural reliability is referenced for strength testing using a myometer (98). In this case, the protocol was standardized across muscle groups, including the command sequence, tactile input, myometer placement, start and end positions, and contraction duration.

Deviation from the standardized protocol can be minimized by conducting periodic procedural reliability checks (97). Procedural reliability should be assessed on an ongoing basis at random intervals in clinical or research settings, in addition to the training period. Assessments should be conducted at least once per phase during a research study. Examiners should be informed that procedural reliability checks will occur randomly and ideally should be unaware of when specific assessments are conducted to avoid examiner reactivity. A minimum acceptable level of procedural reliability should be established for clinical or research use (generally, 90% to 100%). During the training phase, a 100% level should be attained. Feedback on procedural reliability assessments should be provided to examiners. If an examiner's score decreases below the acceptable level, pertinent sections of the standardized protocol should be reviewed.

Establishing Interrater Reliability and Agreement

Once an examiner has demonstrated consistency in scoring by viewing videotaped assessments and reliability in test administration through procedural reliability checks, then interrater reliability and agreement should be established with an experienced examiner (60,92). Both examiners should independently rate people with characteristics similar to the target population. Reliability and agreement assessments should be conducted under conditions similar to those of the actual data collection procedures (61). As with procedural reliability, interrater assessments should be conducted periodically in both clinical and research settings. It is essential to establish interrater reliability and agreement at least once per phase in a research study to determine the potential influence of examiner rating differences on the data recorded (60,61). When calculating interrater agreement when the experienced examiner's scores are considered to be a standard, specific statistical procedures are indicated (72–74).

For assessments where the person's performance can be observed directly (e.g., developmental or activities of daily living assessments), it is preferable to establish interrater reliability and agreement with the examiner in training administering the test while the experienced examiner simultaneously observes and independently scores the person, so that pure interrater reliability and agreement can be assessed. When measuring parameters such as range of motion, sensation, or strength, it is imperative that both examiners independently conduct the tests because the measurement error depends to a large extent on the examiner's skill and body mechanics in administering the test. In addition, direct observation of these parameters by each examiner is required. In these instances, interrater reliability and agreement are confounded by factors of time and variation in patient performance, as discussed above in the section on “Interrater Reliability and Agreement.”

If examiners are aware that interrater reliability and agreement is being assessed, the situation is potentially reactive (61). Reactivity refers to the possibility that behavior may change if the examiners realize they are being monitored. Examiners demonstrate higher levels of reliability and agreement when they are aware that they are being observed. It is difficult, however, to conduct reliability and agreement assessments without examiner awareness; consequently, during a research study, it might be best to lead examiners to believe that all of their observations are being monitored throughout the investigation (61). It is important to note that levels of reliability and agreement attained when examiners are aware that they are being monitored are potentially inflated compared with examiner performance in a typical clinical setting where monitoring occurs infrequently.

---

Detecting Examiner Errors

When training examiners in the use of rating scales, interrater reliability and agreement data should be examined to determine if there are any consistent trends indicative of examiner rating errors. These data should be obtained from testing patients who represent a broad-range sample of pertinent characteristics of the population, so that a relatively normal score distribution is expected. In many circumstances, a representative group of patients can be observed efficiently on videotape by multiple examiners. The distribution of examiners’ scores across patients is then compared for error trends (53). If only one examiner is using a given rating scale, so that multiple examiners’ scores cannot be compared for rating errors, rating errors still can be detected by examining the distribution of one examiner’s ratings across multiple patients. Rasch analysis is another useful method for detecting examiner errors on specific items or as an overall trend. Rating errors can be classified into five categories:

1. Error of central tendency
2. Error of standards
3. Halo effect error
4. Logical error
5. Examiner drift error

An indication of an error of central tendency is when one rater’s scores are clustered around the center of the scale and the other rater’s scores are spread more evenly over the entire scale. Errors of standards occur when one rater awards either all low or all high scores, indicating that his standards are set either too high (i.e., error of severity) or too low (i.e., error of leniency), respectively. Leniency errors are the most common type of rating error (53). Halo effect errors can be detected if several experienced examiners rate a number of people under identical conditions and the score distributions are examined. There should be little variability between well-trained examiners’ scores. If one examiner’s scores fall outside of this limited range of variability, a halo rating error may have occurred as a result of preset examiner impressions or expectations. A logical error occurs when multiple traits are rated and an examiner awards similar ratings to traits that are not necessarily related.

A fifth type of rating error is examiner drift. Examiner drift refers to the tendency of examiners to alter the manner in which they apply rating criteria over time (61). Examiner drift is not easily detected. Interrater agreement may remain high even though examiners are deviating from the standardized rating criteria (60,61). This occurs when examiners who work together discuss rating criteria to clarify rating definitions. They may inadvertently alter the criteria, diminishing rating accuracy, and yet high levels of interrater agreement are maintained. If examiners alter rating criteria over time, data obtained from serial examinations may not be comparable. Examiner drift can be detected by assessing interrater agreement between examiners who have not worked together or by comparing ratings from examiners who have been conducting assessments for an extended period of time with scores obtained from a newly trained examiner (61). Presumably, recently trained examiners adhere more closely to the original criteria than examiners who have had the opportunity to drift. Comparing videotaped samples of patient performance from selected evaluation sessions with actual examiner ratings obtained over time is another method of detecting examiner drift.

Reducing Examiner Errors

Examiner ratings can be improved in several ways (53,60,61). Operational definitions of the behavior or trait must be clearly stated, and examiners must understand the rating criteria. If examiners periodically review rating criteria, receive feedback on their adherence to the test protocol through procedural reliability checks, and are informed of the accuracy of their observations through interrater agreement checks, examiner drift can be minimized. Examiners should be aware of common rating errors and how these errors may influence their scoring. Adequate time needs to be provided to observe and rate behaviors. If the observation period is too brief for the number of behaviors or people to be observed, rating accuracy is adversely affected. The reliability of ratings also can be improved by averaging ratings from multiple observers because the effects of individual rater biases tend to be balanced. Averaging multiple scores obtained from one rater is not advantageous for reducing rating error, however, because a given rater’s errors tend to be relatively constant.

The complexity of observations negatively affects interrater reliability and agreement because observers may have difficulty discriminating between rating criteria (61). With more complex observations, examiners need to attain higher levels of agreement for each behavior during the training phase. These high levels of interrater agreement need to be achieved under the exact conditions that will be used for data collection (61). If multiple behaviors are observed on several patients, it is best to rate all patients on one behavior before rating the next behavior. This practice facilitates more consistent application of operational definitions and rating criteria for the individual behavior. It also tends to reduce the incidence of logical errors.

Another method for improving scoring is to make raters aware of examiner idiosyncrasies or expectations that can affect ratings. According to Verducci (53), there are five patient-rater characteristics that may affect scoring:

1. If an examiner knows the person being evaluated, ratings can be either positively or negatively influenced. The longer the prior relationship has existed, the more likely the ratings will be influenced.
2. The rater tends to rate more leniently if the rater is required to disclose ratings directly to the person or if the person confronts the examiner about the ratings.
3. Examiner gender also can influence ratings. In general, male examiners tend to rate more leniently than female examiners.
4. There is a tendency to rate members of one’s sex higher than those of the opposite sex.

5. Knowledge of previous ratings may bias examiners to rate similarly. Consequently, examiners should remain blind to previous scores until current ratings have been assigned.

Other potential sources of rater bias are the examiner’s expectations about the patient’s outcome and feedback received regarding ratings (60, 61). If examiners expect improvement, their ratings are more likely to show improvement. This is especially true when examiners are reinforced for patient improvement. In a research setting, examiner bias can be minimized if the observers remain blind to the purposes and hypotheses of the study. In a clinical setting, the baseline, intervention, and follow-up sessions can often be videotaped. Blind, independent observers can then rate the behaviors when shown the videotaped sessions in a random order.

**Test Administration Strategies**

Consistency in test administration is essential to permit comparison of test results from one session to another or between people. Multiple factors that might influence performance must be held constant during testing. These factors include test materials and instrumentation, the testing environment, test procedures, and scoring, state of the person being assessed, observers present in the room, and time of day. Examiners must be aware of the potential influence of these factors and document any conditions that might affect test performance. Examiners ideally should remain blind to previous test results until after conducting the evaluation to avoid potential bias.

If more than one method is acceptable for testing, it is important to document which protocol is used so that the same method can be used during future evaluations. If it is necessary to alter the method of measurement as a result of a change in status or the development of an improved measurement technique, measurements should be taken using both the new and old methods so there is overlap of at least one evaluation. This overlap permits comparison with previous and future test results so that trends over time can be monitored.

Multiple trials should be administered when assessing traits, such as muscle strength, which require consistent efforts on the part of the patient. An average score of multiple trials is more stable over time than a single effort (98). A measure of central tendency and the range of scores both should be reported.

Standardized test positions always should be used unless a medical condition prevents proper positioning (e.g., joint contractures). In this event, the patient should be positioned as closely as possible to the standardized position, and the altered position should be documented. It is important to make sure that patients are posturally secure and comfortable during the evaluation. For patients with neurologic involvement, the head should be positioned in neutral to avoid subtle influences of tonic neck reflexes. An exception occurs when testing is conducted in the prone position. In this case, the head should be turned consistently toward the side being tested.

A key to obtaining reliable and valid test results is providing clear directions and demonstrations to the patient. Standardized instructions always must be provided verbatim and may not be modified or repeated unless specifically permitted in the test manual. Verbal directions are enhanced by tactile, kinesthetic, and visual cues, if permitted. If confusion about the task is detected, this should be documented. If the examiner believes that a given patient could complete a task successfully with further instructions that are not specified in the standardized protocol, this item can be readministered at the end of the test session. The person’s test score should be based solely on performance exhibited when given standardized instructions. Test performance with augmented instruction can be documented in the clinical note but should not be considered when scoring.

When conducting tests that do not have standardized instructions (e.g., strength testing), it is important to use short, simple, and consistent commands. If repetitive or sustained efforts are required, the examiner’s voice volume needs to be consistent and adequate to heighten the arousal state and motivate patients to give their best effort.

Verbal reinforcement and feedback regarding performance can influence performance levels (99). Consequently, it must be provided consistently, according to the procedures specified in the test manual. For tests in which reinforcement and feedback intervals are not specified and are permitted as needed, the frequency and type of feedback provided should be documented.

**Test Scoring, Reporting, and Interpretation of Scores**

Examiners should be thoroughly familiar with scoring criteria so that scores can be assigned accurately and efficiently during evaluation sessions. It is not appropriate for examiners to look up scoring criteria during or after the evaluation. Uncertainty about the criteria prolongs the evaluation and leads to scoring errors. It is helpful to include abbreviated scoring criteria on the test form to assist the examiner during the evaluation. Test forms should be well organized and clearly written to facilitate efficient and accurate recording of test results. If multiple types of equipment and test positions are required, it is useful if the equipment and position are identified on the score sheet using situation codes for each item. Such a coding system expedites test administration by assisting the examiner in grouping test items with similar positioning and equipment requirements. Examples of well-organized test forms that use situation codes are the Bayley Scales of Infant Development (100), the MAP (34), and the revised version of the Peabody Developmental Motor Scales test forms (41).

If the scoring criteria for a test are not well-defined, it may be necessary for examiners within a given center or referral region to clarify the criteria. This was the case for many items on the Peabody Developmental Motor Scales. Intratrer
reliability levels of highly trained examiners were low for several items, so therapists at the Child Development and Mental Retardation Center in Seattle, Washington, clarified the scoring criteria to improve reliability. Examiners in the surrounding referral area were educated about the clarified criteria by means of in-services and videotapes to ensure that all examining centers in the area would be using identical criteria (41). If scoring criteria are augmented to improve reliability, it is imperative to document that the test was administered with altered criteria. Future results are comparable only if administered using identical scoring criteria. Additionally, if scores are compared with normative data, it is important to document that the test scores obtained may not be directly comparable because altered scoring criteria were used.

Raw scores obtained from testing are meaningless in the absence of additional interpretive data. To meaningfully compare a person’s current test results to previous scores, the SEM of the test must be known. To determine how a person’s performance compares with that of other people, normative data must come from a representative standardized sample of people with similar characteristics. In the latter case, the raw score must be converted into a derived or relative score to permit direct comparison with the normative group’s performance. These concepts are discussed in detail below.

Raw scores may be compared with previous scores obtained from a given person to monitor changes in status. However, the SEM of the test must be known to determine if a change in a score is clinically significant. A change in a test score exceeding the SEM is indicative of a meaningful change in test performance. As was discussed earlier, in the section on “Reliability and Agreement,” it is best to report test scores as a range, based on confidence intervals, rather than as an absolute score. This is because a person’s score is expected to vary as a result of random fluctuations in performance. It is only when a score changes beyond the range of random fluctuation that we can be confident that a true change in performance has occurred. This true score range usually is based on the 95% confidence interval. This rigorous level of confidence minimizes the likelihood of a type I error (i.e., believing a change occurred when actually there was no change) and is considered the confidence level of choice when looking for improvement in performance, resulting from a specific treatment regimen or improved physical status. A lower level of confidence (e.g., 75%, 50%) may be desirable when monitoring the status of

### TABLE 8.2 Descriptive and Standard Scores Commonly Reported in Rehabilitation Medicine

<table>
<thead>
<tr>
<th>Summary Statistic</th>
<th>Definition and Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Descriptive score</td>
<td></td>
</tr>
<tr>
<td>Raw scores</td>
<td>Expressed as number of correct items, time to complete a task, number of errors, or some other objective measure of performance.</td>
</tr>
<tr>
<td>Percentage scores</td>
<td>Raw scores expressed as percent correct.</td>
</tr>
<tr>
<td>Percentile scores</td>
<td>Expressed in terms of the percentage of people in the normative group who scored lower than the client’s score (e.g., a client scoring in the 75th percentile on a norm-referenced test has performed better than 75% of the people in the normative group). Often stratified for age, gender, or other pertinent modifying varieties.</td>
</tr>
<tr>
<td>Age-equivalent score</td>
<td>Average score for a given age group.</td>
</tr>
<tr>
<td>Grade-equivalent score</td>
<td>Average score for a given grade level.</td>
</tr>
<tr>
<td>Developmental age</td>
<td>The basal age score, plus credit for all items earned at higher age levels (up to the ceiling level of the test). The basal age level is defined as the highest age at and below which all test items are passed.</td>
</tr>
<tr>
<td>Scaled score</td>
<td>The client’s total score, summed across all sections of the test. Used for comparison to previous and future scores.</td>
</tr>
<tr>
<td>Standard scores</td>
<td></td>
</tr>
<tr>
<td>z score</td>
<td>The client’s raw score minus the mean score of normative group, divided by the standard deviation of the normative group. The mean of a z score is 0 with a standard deviation of 1. Scores may be plus or minus. Reported to two significant digits.</td>
</tr>
<tr>
<td>T score</td>
<td>z score times 10 plus 50. The mean of a T score is 50 with a standard deviation of 10.</td>
</tr>
<tr>
<td>Stanine</td>
<td>Standard scores which range from 1 to 9. A stanine of 5 indicates average performance and the standard deviation is 2. Often used to minimize the likelihood of overinterpreting small differences between individual scores.</td>
</tr>
<tr>
<td>DMQ</td>
<td>The ratio of the client’s actual score on the test (expressed as developmental age) and the client’s chronologic age, DMQ = DA/CA. The DMQ equals the z score times 15, plus 100. The mean DMQ is 100, with a standard deviation of 15.</td>
</tr>
<tr>
<td>Deviation IQ</td>
<td>A standard score deviation of the ratio between the client’s actual score on the test, expressed as a mental age and the client’s chronological age. The mean deviation IQ is 100, with a standard deviation of 15, based on the Wechsler deviation IQ distribution.</td>
</tr>
</tbody>
</table>

CA, chronological age; DA, developmental age; DMQ, developmental motor quotient; IQ, intelligence quotient; MA, motor age.
people who are at risk for loss of function over time. For these people, it is important to minimize the likelihood of a type II error (i.e., believing no change occurred when actually there was a change). In such cases, if a person’s score falls outside a true score range that is based on a lower level of confidence, it may indicate the need to conduct further diagnostic tests or to monitor the person more closely over time.

If normative data are available for a given test, a person’s score can be compared directly to the normative group’s performance by converting the score into a derived or relative score. Normative scores provide relative rather than absolute information (101). Normative data should not be considered as performance standards but rather as a reflection of how the normative group performed. Derived scores are expressed either as a developmental level or as a relative position within a specified group. Derived scores are calculated by transforming the raw score to another unit of measurement that enables comparison with normative values. Most norm-referenced tests provide conversion tables of derived scores that have been calculated for the raw scores so that hand calculations are not required. However, it is important for examiners to understand the derivation, interrelationship, and interpretation of derived scores. Specific calculation of these scores is beyond the scope of this chapter. For computational details and the practical application of these statistical techniques, the reader is referred to textbooks on psychological or educational statistics and measurement theory (92,94,101).

Selection of the particular type of score to report depends on the purpose of testing, the sophistication of the people reading the reports, and the types of interpretations to be made from the results (101). Table 8-2 summarizes various descriptive and standard scores that are commonly used. Figure 8-2 shows the relationship of these scores to the normal distribution and the interrelationship of these scores. Calculation of standard scores (e.g., $z$ scores, $T$ scores, stanines, developmental motor quotients, deviation IQ) is appropriate only with interval or ratio data. They express where a person’s performance is with regard to the mean of the normative group, in terms of the variability of the distribution. These standard scores are advantageous because they have uniform meaning from test to test. Consequently, a person’s performance can be compared between different tests.

**Written Evaluation**

Thorough documentation of testing procedures and results is essential in both clinical and research settings to permit comparison of test results between and within individuals. The tests administered should be identified clearly. Any deviations from the standardized procedures, such as altered test positions or modified instructions, should be documented (14). If multiple procedural options are available for a given test item (e.g., measuring for a flexion contracture at the hip), the specific method used should be specified in the report. The patient’s behavior, level of cooperation, alertness, attention, and
motivation during the evaluation should be documented. Any potential effect of these factors on test performance should be stated. Other factors that might have influenced the validity of test results also should be noted (e.g., environmental factors, illness, length of test session, activity level before the test session). It should be indicated whether optimal performance was elicited. If a person’s performance is compared with normative data, the degree of similarity of the person’s characteristics to those of the normative group should be stated. It is imperative to distinguish between facts and inferences in the written report.

The use of a standard written evaluation format facilitates communication between and within disciplines. In addition, computerized databases provide standardized formats useful for both clinical and research purposes. Serial examinations of a given person can be reviewed easily, and a patient’s status can be compared directly with that of other people with similar characteristics. Clinical and research applications of computer databases for documentation in rehabilitation medicine are discussed by Shurtleff (102) and Lehmann et al. (103).

**CONTEMPORARY MEASUREMENT TECHNIQUES**

There are several potential deficiencies in current measurement methods that can negatively impact the usefulness of medical rehabilitation tests and measures across the continuum of care and research. Jette and Haley (104) include “(a) narrowly defined scope of outcome measures; (b) the inability of different outcome instruments to talk to each other; (c) the classic trade-off between feasibility of existing outcome measures versus their limitations in detecting clinically relevant changes.” Two contemporary measurement techniques used for many years in educational testing, item response theory (IRT) and computer adaptive testing (CAT), can overcome these limitations.

**IRT Techniques**

Jette and Haley (104) state “IRT methods examine the associations between individuals’ response to a series of items designed to measure a specific outcome domain (e.g., physical functioning). Data collected from samples of rehabilitation patients are fit statistically to an underlying IRT model that best explains the covariance among item responses. IRT measurement models are a class of statistical procedures used to develop measurement scales. The measurement scales are comprised of items with a known relationship between item responses and positions on an underlying functional domain, called an item characteristic curve.” The form of the relationships is typically nonlinear. Using this approach, probabilities of patients scoring a particular response on an item at various functional abilities can be modeled. Persons with more functional ability have higher probabilities of responding positively to functional items than persons with lower functional abilities. These probability estimates are used to determine the individual’s most likely position along the functional dimension. When assumptions of a particular IRT model are met, estimates of a person’s functional ability do not strictly depend on a particular fixed set of items. This scaling feature allows one to compare persons along a functional outcome dimension even if they have not completed the identical set of functional outcomes. “Within rehabilitation, researchers have linked functional outcome items from an item pool to create a practical yet comprehensive set of short forms that can be applied in different rehabilitation settings. IRT methods open the door to understanding the linkages among items used to assess a common functional outcome domain, and in this way serve as the psychometric foundation for CAT.”

**CAT Techniques**

Jette and Haley (104) describe CAT programs as “a simple form of artificial intelligence that selects questions tailored to the test-taker, and thereby shortens or lengthens the test to achieve the level of precision desired by the user. Functional outcome CAT applications rely on extensive item pools constructed for each outcome area. They contain items that consistently scale along each functional outcome dimension from low to high proficiency, and include rules guiding starting, stopping and scoring procedures. CAT methodology uses a computer interface for the patient/clinician report that is tailored to the patient’s unique ability level. The basic notion of CAT test is to mimic what an experienced clinician does. A clinician learns most when he or she directs questions at the patient’s approximate proficiency level. Administering outcomes items that represent tasks that are too easy or too hard for the patient provides little information.” “A CAT is programmed to first present an item from the midrange of an IRT-defined item pool, and then direct subsequent functional items to the level based on the patient’s (or clinician’s) previous responses, without asking unnecessary questions. The selection of an item in the mid-range is arbitrary and the CAT can be set to select an initial item based on other information entered about the patient such as age, diagnosis, or severity of condition.”

**OBJECTIVE MEASUREMENT WHEN A STANDARDIZED TEST IS NOT AVAILABLE**

**Rationale for Systematically Observing and Recording Behavior**

Standardized tests and objective instrumentation are not always available to measure the parameters of clinical and research interests. Consequently, rehabilitation professionals often resort to documentation of subjective impressions (e.g., “head control is improved,” “wheelchair transfers are more independent and efficient”). However, functional status and behaviors can be documented objectively by observing behavior using standardized techniques that have been demonstrated to be reliable. Systematically observing and
recording behavior provides objective documentation of behavior frequency and duration, identifies the timing and conditions for occurrence of a particular behavior, and identifies small changes in behavior. Several of the procedures for objective documentation described below are based on the principles of single-case research designs. These research designs have been suggested to be the most appropriate method of documentation of treatment-induced clinical change in rehabilitation populations, owing to the wide variability in clinical presentation, even within a given diagnostic category (95,105,106). In addition, such designs have been recommended to evaluate and compare the effects of two different treatments on individual patients (104). Selected single-case research concepts that specifically pertain to objective documentation for either clinical or research purposes are presented in this chapter. The reader is referred to Hayes et al. (107), Barlow and Hersen (60), Bloom et al. (47), Kazdin (61), and Ottenbacher (95) for more thorough discussions of documentation using single-case research standardized testing techniques.

Procedures for Objective Observation and Recording of Behavior

Step 1: Identify the Target Behavior to be Monitored
The target behavior must be identified by specifying the parameters of interest and their associated conditions. The prerequisite conditions required must be defined, such as verbal directions, visual or verbal cues, or physical assistance provided. In addition, environmental conditions must be described because different responses may be observed in the therapy, inpatient ward, or home setting. The duration, frequency, and timing of the observation period also must be specified. Ideally, these conditions should be constant from one observation period to the next for comparison purposes.

Step 2: Operationally Define the Target Behavior
An operational definition is stated in terms of the observable characteristics of the behavior that is being monitored. The definition must describe an observable or measurable action, activity, or movement that reflects the behavior of interest. The beginning and ending of the behavior must be clearly identified. Objective, distinct, and clearly stated terminology should be used (60,95). The definition should be elaborated to point out how the response differs from other responses. Examples of borderline or difficult responses, along with a rationale for inclusion and exclusion, should be provided. An example of an operational definition used to determine success or failure in drawing a circle is provided in Figure 8-3.

Step 3: Identify the Measurement Strategy
There are five methods of sampling behavior: event recording, rate recording, time sampling, duration recording, and discrete categorization (47,60,95). Each of these methods will be described below, along with indications and contraindications for their use.

Event Recording
The number of occurrences of the behavior is tallied in a given period of time or per given velocity in the case of mobility activities. Indications for event recording include when the target response is discrete, with a definite beginning and end, or when the target response duration is constant. The target behavior frequency should be low to moderate, and the behavior duration should be short to moderate. It is best to augment the number of occurrences with real-time information to permit sequential, temporal, and reliability and agreement analyses. Contraindications for using event-recording techniques include behaviors that have a high incidence of occurrence because of the increased probability of error in counting the high-frequency behavior and behaviors that have an extended duration or that occur infrequently (60,95) (e.g., wheelchair transfers). Duration recording should be used in the latter case. The following is an example of event recording:

A man with hemiplegia successfully fastened 5 of 10 shirt buttons during a 10-minute period of time using his involved hand to hold his shirt and his uninvolved hand to manipulate the buttons. The number of successes, number of trials, and duration of the observation period were recorded.

Rate Recording
The number of occurrences of the behavior is divided by the duration of the observation period (e.g., the number of occurrences per minute). This method is indicated when the observation period varies from session to session. Rate recording is advantageous because it reflects changes in either the duration or the frequency of response and is sensitive for detecting changes or trends because there is no theoretical upper limit. The following is an example of rate recording.

FIGURE 8-3. The operational definition of a circle (dashed lines: circle path template; solid line: patient’s drawing of a circle). The patient is instructed to draw a circle inside the two dashed lines. An adequate circle is one in which the two ends meet, and the line of the circle stays within the circle path template. It can touch the edges of the template but cannot extend beyond the edges.
A child with Down syndrome exhibits five occurrences of undesirable tongue thrusting during a 10-minute observation period the first day and eight times during a 20-minute observation period the second day. The observations were made from videotapes recorded immediately after the child’s oral motor therapy program. An independent observer, who was blind to the child’s intervention program, performed the frequency counts. The rate of responding was 0.5 behaviors per minute (5 per 10 minutes) for the first day and 0.4 behaviors per minute (8 per 20 minutes) for the second day.

Time Sampling
This method involves recording the state of a behavior at specific moments or intervals in time. It also has been described in the literature as scan sampling, instantaneous time sampling, discontinuous probe time sampling, and interval sampling. Time sampling is analogous to taking a snapshot and then examining it to see if a particular behavior is occurring. This method often is used in industrial settings to determine exposure to risk factors or compliance with injury prevention techniques.

To monitor behavior using this method, the behavior of interest is observed for a short block of time (e.g., a 5-second observation period) at specified recording intervals (e.g., 5-minute intervals) during a particular activity (e.g., a 30-minute meal period). The recording interval is signaled to the observer by means of a timer, audiotape cue, or a tone generator. The target behavior is scored as either occurring or not occurring during the observation period of each recording period. Fixed (i.e., preset) or random intervals can be used, but it is important to avoid a situation in which the signal coincides with any regular cycle of behavior. The sampling should occur at various times throughout the day and in different settings to obtain a representative picture of the behavior frequency. The recording interval length depends on the behavior duration and frequency, as well as on the observer’s ability to record and attend to the person. The more frequent the behavior, the shorter the interval. For low to medium response rates, 10-second intervals are recommended. For high response rates, shorter intervals should be used (108). An advantage of this type of recording is that several patients can be observed simultaneously by one rater in a group setting (e.g., during meal times or recreational events) by staggering the recording intervals for each patient.

Variations of time sampling include observing the behavior during a single block of time that is divided into short intervals (i.e., interval recording) or during brief intervals that are spread out over an entire day (i.e., time sampling); combining time sampling and event recording, when the number of responses occurring during a given interval are recorded; and combining time sampling and duration recording, when the duration of the response during a given interval is recorded. The following are examples of time sampling:

To document a patient’s ability to maintain his head in an upright position, the nursing staff observed him for 15 seconds at 5-minute intervals during one 30-minute meal period, during one 30-minute self-care/dressing period, and during one 30-minute recreation period.

To estimate compliance of 12 industrial workers with suggestions provided in a back school program, the time individual workers spent in appropriate versus inappropriate postures was recorded for 5 minutes each hour during an 8-hour shift.

Duration Recording
Either the duration of the response or the length of the latency period is recorded. The duration is reported as the total time if the observation period is constant or as the percentage of time that a behavior occurred during observation periods of varying length. Indications for this method include continuous target responses, behaviors with high or even response rates, and behaviors with varying durations, such as a wheelchair transfer, for which a frequency count would be less meaningful. The behavior duration is timed with a stopwatch, electro-mechanical event recorder, or electronic keyboard. Variations of duration recording include timing the response latency (i.e., the time that elapses between a cue and a response), measuring the time required to complete a particular task, or monitoring the time spent performing a particular activity. The following are examples of duration recording:

- The amount of time that it takes an adult with a spinal cord injury to dress in the morning
- The length of time that a child is able to stand independently with and without orthotics before losing his or her balance.

Discrete Categorization
With this method of behavior measurement, several different behaviors of interest are listed and checked off as being performed or not performed. This method is useful in determining whether certain behaviors have occurred. It is indicated when behavioral responses can be classified into discrete categories (e.g., correct/incorrect, performed/not performed). An example of this method is a checklist of the different steps for performing a wheelchair transfer, such as positioning the wheelchair, locking the brakes, removing feet from footrests, and so forth. The observer checks off whether each of these steps was performed during a given transfer.

Step 4: Establish Interrater Reliability
There are four reasons for assessing interrater (i.e., interobserver) reliability and agreement.

1. To establish how consistently two observers can measure a given behavior
2. To minimize individual observer bias by establishing interrater reliability and then retraining observers if the level of reliability is unacceptable
3. To reduce the chances of an examiner altering or “drifting” from the standard method of rating by implementing periodic interrater reliability or agreement checks to ensure that observers are consistent over time
4. To examine the adequacy of operational definitions, rating criteria, and scoring procedures. Items that have poor agreement should be revised

Before the onset of data collection, two people should independently observe and score pilot subjects who have characteristics that are similar to those of the clinical or study population. Behaviors of interest are rated according to predetermined operational definitions. Interrater reliability and agreement are then calculated using an appropriate statistic (see section on Reliability and Agreement). The minimum acceptable level of agreement depends on the type of statistic calculated (see Table 8-1).

If interrater reliability or agreement is below the target level, improvement may occur by discussing operational definitions of the behaviors. If problems with reliability or agreement continue, it may be necessary to redefine behaviors, improve observation and recording conditions, reduce the number of behaviors being recorded, provide additional training, and, if necessary, further standardize the data collection environment (65,85). Interrater reliability or agreement should be reestablished once remedial steps have been taken. As stated previously, periodic checks of interrater reliability or agreement should be conducted in the clinic and at least once during each phase of a research study (64,65). Reliability and agreement data should be plotted along with clinical or research data to show the level of consistency in measurements.

Step 5: Report Scores and Graph Data
Baseline, intervention, and follow-up data should be plotted on a graph or chart to provide a pictorial presentation of the results. Graphing strategies include using standard graph paper or a standard behavior chart (i.e., six-cycle graph paper). Advantages of the latter are that it permits systematic, standardized recording using a semilog scale that allows estimation of linear trends. Extremely high and low rates can be recorded on the chart. Behavior rates that range from once per 24 hours to 1,000 per minute can be accommodated; therefore, data are not lost as a result of floor or ceiling effects. In addition, continuous recording of data for up to 20 weeks is permitted. For further information on graphing strategies, the reader is referred to White and Haring (109) and Carr and Williams (110) for use of the standard behavior chart in clinical settings.

The time period of data collection is plotted on the horizontal axis (e.g., hours, days, weeks) and changes in the target behavior on the vertical axis. Appropriate scaling should be used to accommodate the highest expected response frequency and the longest anticipated documentation period duration. The measurement interval on both axes should be large enough to permit visual detection of any changes in behavior. Interrater reliability data from each phase should be plotted on the same graph, along with the study results, as discussed previously.

Considerations When Reporting Scores
The percentage of correct scores often is reported because of the ease of calculation and interpretation. However, usefulness of this summary statistic is limited because it does not provide information on the number of times a patient has performed correctly (95). Consequently, it can be misleading if the total number of opportunities varies from day to day. For example, three successes of six trials on day 1 versus three successes out of four trials on day 2 would yield percentages of 50% and 75%, respectively. Based on percentage scores, it would appear that the patient’s performance was improved, and yet the absolute number of successes has not changed. Additionally, if an odd number of trials are administered on some days and an even number of trials on other days, performance changes may occur based on percentage scores simply because it is not possible to receive half credit for a trial on days when an odd number of trials are given (e.g., five successes out of ten trials vs. three successes out of five trials).

SUMMARY

Rehabilitation practitioners and researchers in rehabilitation medicine increasingly are using objective tests and measurements as a scientific basis for communication, to establish credibility with other professionals, and to document treatment effectiveness. The increased use of such measures has resulted in greater responsibility of the user for appropriate implementation and interpretation of tests and measures. Rehabilitation professionals must be familiar with the principles of objective measurement to use the tools properly.

The section “Psychometric Parameters Used to Evaluate Tests and Measures” of this chapter described the psychometric parameters used to evaluate the state of development and quality of available objective measures. The four basic levels of measurement—nominal, ordinal, interval, and ratio scales—were defined. The purposes for testing were discussed, including screening tests, in-depth assessment tests, and criterion-referenced tests. Several issues of practicality for selection and use of tests also were identified. The various forms of reliability, agreement, and validity described are of great importance for using the various measurements effectively. A test that does not provide reproducible results, or does not measure what it is purported to measure, is of no value and is potentially harmful by giving a false implication of meaningfulness. Consequently, caution in the use and interpretation of test results must be exercised when information on reliability or validity of a measure is not available or if their values are below accepted levels.

The section “Principles of Evaluation, Testing, and Interpretation” of this chapter discussed the principles of evaluation, testing, and interpretation that help to ensure that adequate reliability and validity are obtained from test administration. The issues of standardization, interrater reliability, and procedural reliability are of particular importance. Care must be taken during test administration to control for the potential rater errors of central tendency, standards, halo effect, logical errors, and examiner drift.

For many applications in rehabilitation medicine practice and research, standardized measures have not yet been
developed. Methods derived from single-subject research paradigms provide guidelines for objective measurement when a standardized test is not available. These guidelines, which are discussed in the section "Procedures for Objective Observation and Recording of Behavior" of this chapter, include identifying the behavior to be monitored, operationally defining the behavior, identifying the measurement strategy (e.g., event recording, rate recording, time sampling), establishing inter-rater reliability, and properly reporting scores and graphing the data.

Specific tests and objective measurement instruments are not discussed because of the number and broad spectrum of measures used by rehabilitation professionals.

The principles discussed in this chapter provide the framework for the readers critically to assess the measures available for their specific application needs. Such critical analysis will further emphasize the need for ongoing development and improvement of objective measures at the disposal of rehabilitation professionals.

REFERENCES

11. Hinderer KA, Hinderer SR, Deitz JL. Reliability of manual muscle testing using the hand-held dynamometer and the myometer: a compari-
27. Merbitz C, Morris J, Grip JC. Ordinal scales and foundations of misin-
30. Deitz JC, Tovar VS, Beeman C, et al. The test of orientation for rehabilita-
31. Deitz JC, Tovar VS, Thorn DW, et al. The test of orientation for rehabili-
36. Gans BM, Haley SM, Hallenborg SC, et al. Description and inter-
37. Gans BM, Haley SM, Hallenborg SC, et al. Description and inter-
38. Gans BM, Haley SM, Hallenborg SC, et al. Description and inter-
40. Gans BM, Haley SM, Hallenborg SC, et al. Description and inter-


46. Baumgartner TA, Jackson AS.

47. Bloom M, Fischer J, Orme JG.


Functional Evaluation and Management of Self-care and Other Activities of Daily Living

INTRODUCTION

In this chapter, we discuss self-care and activities of daily living (ADL) from the standpoint of rehabilitation across a wide range of ages and conditions. We first describe the importance of this domain of human activity to illustrate why self-care and ADL performance are important rehabilitation goals. We then describe approaches to assessing an individual’s ADL skills, sampling the range and scope of assessments in this area that are supported by published research. We attempt to identify best practices and areas of philosophical and procedural controversies. Finally, we turn to the intervention strategies. A collaborative goal-setting process to guide intervention is emphasized. We summarize key intervention strategies across a sample of conditions, citing evidence in support of intervention strategies, where available.

Our objectives in this chapter include

- To familiarize the reader with the central importance of self-care (ADL and IADL) in the life of the patient viewed from the perspective of discharge setting
- To describe a range of approaches toward measuring self-care performance
- To identify strengths and weaknesses of various ADL and IADL assessment instruments
- To describe collaborative goal-setting and development of a plan of care
- To review various strategies and principles for intervention within the categories of adaptation, remediation, and compensatory strategies, health promotion, disability prevention, environmental modifications, use of assistive technology devices (ATDs), and use of personal care attendants (PCAs)
- To summarize and research ADL-related interventions and plans of care

Enabling individuals to manage daily self-care is among the most important tasks undertaken by the rehabilitation team. This is because such tasks relate directly to the business of living and their performance signifies a return to participation in the routines of daily life. Self-care tasks include dressing, eating, bathing, grooming, use of the toilet, and mobility within the home. These are basic tasks included within the general category of ADL. Although able-bodied persons perform most self-care tasks routinely, such tasks can represent difficult challenges for persons with sensory, motor, and/or cognitive deficits.

Other important activities for living in the community are related to managing the requirements of daily life. These extended ADLs (EADLs) go beyond basic self-care and have been labeled by Lawton as instrumental ADL, or IADLs (1). These include food preparation, laundry, housekeeping, shopping, use of the telephone, use of transportation, use of medication, and financial management. Childcare is also an important responsibility in the daily routine of many people.

Unfortunately, there is no consensus on the classification of human activity. As a result, many terms with similar definitions used for ADL categories are used in the medical, health, and rehabilitation literature. Table 9-1 lists some of these. In this chapter, basic ADLs (BADL), such as eating, dressing, grooming, hygiene, and mobility, are described as personal or self-care tasks. Essential tasks for maintaining the living environment and residing in the community are described as IADLs or EADLs (see Table 9-1).

Self-care tasks may assume a symbolic meaning for the individual in a rehabilitation program, because attending to eating, dressing, and toileting tasks is basic part of the routine necessary for establishing a sense of identity (2) as well as for gaining acceptance in a social world (3). This is because appropriate dress, personal appearance, hygiene, and other expectations influence perceptions of the self and others (4).

In the developed nations, about 30% of a typical person's waking hours is spent performing self-maintenance activities, including basic self-care and household maintenance (5,6). For able-bodied persons, an average of more than 1 hour per day is spent in basic self-care activities (7). Research has shown that more than 70% of the variation in discharge decisions following stroke rehabilitation is determined by the ability to function independently self-care tasks necessary for bathing, toileting, social interaction, dressing, and eating (8,9).
In the United States, recognition of the importance of functional independence is reflected in population survey data collected by governmental agencies, such as the National Center for Health Statistics (10,11). An individual’s functional ability is an important predictor of nursing home placement, with research showing a high correlation between the number of dependencies in ADL and the risk of institutional placement (10,12).

### Self-care and ADL in Context

Current international models of disability consider the multiple factors that influence daily life and the ability to perform necessary life tasks (13). These models recognize the importance of the physical and social settings in which an individual lives, and how these factors come together to support or limit task performance and participation as a member of society. Others with whom an individual regularly spends time constitute that individual’s social nucleus, providing important support and social interaction and influencing activity choices and role requirements as well as the level of independence (14–16). This nucleus typically includes friends, acquaintances, and members of the individual’s immediate and extended family.

It is within this social situation that the importance of self-care is most apparent, because meeting self-care needs is vital to success in meeting expectations for social interaction. Self-esteem, or the value accorded oneself, is determined by how well self-evaluation matches the values perceived as important in the social environment (17). Self-esteem is influenced by social acceptance and by one’s success in achieving a desired social identity (18–21). Because the ability to perform self-care tasks contributes to both acceptance and identity, it can have a direct effect on self-esteem (19,22). Importantly, social factors, including social support, are important predictors of rehabilitation outcome as shown by studies of amputation (23), stroke (24,25), serious burn injury (26), spinal cord injury (20), hip fracture (27).

Typically, self-care activities are taken for granted by the person and society unless difficulties are encountered. Limitations in self-care tasks and dependency on others for their completion serve to diminish an individual’s self-concept and can lead to decreased morale and depression (28). A study of elderly patients found that a relationship existed between self-concept and functional independence, and that people who were dependent in ADL scored lower in measures of self-concept (19).

### TABLE 9.1 Terminology of Functional Performance

<table>
<thead>
<tr>
<th>Frequently Used Categories</th>
<th>Activities Typically Included in the Category</th>
<th>Correspondence with ICF Descriptions of Categories and Activities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-care</td>
<td>Bathing Communication</td>
<td>Self-care: caring for oneself, washing and drying oneself, caring for one's body and body parts, dressing, eating and drinking, and looking after one's health</td>
</tr>
<tr>
<td>Personal care</td>
<td>Dressing Eating</td>
<td>Mobility: moving by changing body position or location, or by transferring from one place to another, by carrying, moving or manipulating objects, by walking, running or climbing, and by using various forms of transportation</td>
</tr>
<tr>
<td>Basic ADL (BADL)</td>
<td>Grooming Mobility Toileting (sphincter control) Transfers</td>
<td>Communication: general and specific features of communicating by language, signs and symbols, including receiving and producing messages, carrying on conversations, and using communication devices and techniques</td>
</tr>
<tr>
<td>Instrumental activities of daily living (IADL)</td>
<td>Child care Financial management Food preparation</td>
<td>Domestic life: acquiring a place to live, food, clothing and other necessities, household cleaning and repairing, caring for personal and other household objects, and assisting others</td>
</tr>
<tr>
<td>Extended ADL (EADL)</td>
<td>Housekeeping Laundry Medication use Shopping Telephone use Transportation use</td>
<td>DOMESTIC LIFE: acquiring a place to live, food, clothing and other necessities, household cleaning and repairing, caring for personal and other household objects, and assisting others</td>
</tr>
<tr>
<td>Reintegration</td>
<td>Paid Work Volunteerism</td>
<td>Major life areas: carrying out the tasks and actions required to engage in education, work and employment and to conduct economic transactions, including education, work and employment, and economic life</td>
</tr>
<tr>
<td>Community participation</td>
<td>Leisure Recreation</td>
<td>Community, social and civic life: the actions and tasks required to engage in organized social life outside the family, in community, social, and civic areas of life, such as participation in spirituality and religion, political life, and citizenship, volunteerism, recreation, and leisure</td>
</tr>
</tbody>
</table>
Research has shown a clear relationship among self-concept, morale, and level of functional independence. For example, Chang and MacKenzie found that self-esteem was a consistent and significant predictor of functional ability at various intervals following stroke (20). Chemerinski et al. found that improvements in ADL performance were associated with remission of poststroke depression (21). These studies and others (29) indicate that an important goal of rehabilitation should be to help patients learn to take control of decisions about daily living, since this may contribute positively to their sense of efficacy, morale, and overall sense of well-being. More important, it may also increase life expectancy, since loss of hope and feelings of helplessness during early rehabilitation phases have been shown to be associated with shorter survival rates following stroke (30).

Within living settings, the presence of an individual with needs for care-giving affects the entire family or social group (31,32). When a member of the family can no longer perform expected activities, the daily routine may be upset, creating stress, diminished psychological well-being, and conflict (33). Family members must adjust their expectations of the individual who is disabled as well as adjust to changes in family routines and activities (34).

Family caregivers are important to the well-being of persons with disability and chronic illness (35). Over time families experience stages and time periods, each with characteristic issues. Important concerns related to self-care and care-giving needs must be considered in light of these stages, with recognition that needs change over time. The most significant change affecting caregiving is the number of family members who are available to provide support as a family life cycle matures.

Necessary adjustments made by families or caretakers confronted with rehabilitation challenges often include a readjustment of homemaking tasks or changes in priorities and may impose additional financial or social burdens due to the need to hire outside assistance or rely on volunteers (36). Studies have shown that levels of depression and anxiety, as well as somatic complaints, are more prevalent among caretakers and family members of disabled people living in the home environment than those typically found among members of the general population (33,34). Yeung et al. found that self-confidence in Chinese family caretakers was an important factor in psychosocial well-being (32). Caregiver burden, a term given to the general strain, isolation, disappointment, and emotional demands of caring for a member of the household with a disability, seems to increase in proportion with ADL needs (37–39). A study of families involved in caring for survivors of stroke found that family adaptation after 1 year was related to family stresses and demands, family resources, and family perceptions. In particular, family functioning was poorer when the patient developed psychological morbidity, when the patient was less satisfied with the recovery, and when the health burden of the stroke was greater (40). In recognition of the increasing importance of the role of household caregivers, interventional strategies, including counseling, education and training, and social support, have been reported. Recent meta-analyses have indicated that such interventions are effective in improving well-being and mental health and reducing the “burden” of care among caregivers (41,42).

During the rehabilitation process, the family can have a considerable influence on functional outcome (43). A stable and supportive family unit can be of great assistance, whereas families that are functioning poorly can impede rehabilitation. In some cases, poor outcomes can be traced to a lack of family involvement in the rehabilitation process (44). In other cases, too much support can encourage dependency (41). This indicates that the family should be involved in all aspects of rehabilitation, including evaluation and the setting of rehabilitation goals and treatment strategies before and after discharge (45).

A primary source of adjustment difficulties for people with physical disabilities comes from societal treatment of them as socially inferior (46). The common belief that strength, independence, and appearance are important aspects of self-worth is very damaging to people with disabilities. Interaction within a social group often depends on the ability to perform at the group’s expected level; otherwise, the person will not be included as a significant group member (18).

Self-care tasks are not publicly valued in the same manner as gainful employment (47,48). Ironically, they assume importance principally when one’s inability to perform them leads to perceived disadvantage or social stigma (49). Self-reliance in ADL helps to refute the idea that a person with a disability may be a financial or social burden to society. It is important to realize that social participation and quality of life are often the ultimate goals of patients, and this endpoint should influence shared goal setting. Physical health is an enabler of well-being, and the capacity to accomplish self-care represents the beginning set of tasks necessary for participation. As noted by Hogan and Orme, ambulation and self-care mastery alone are insufficient for attaining desired goals related to social participation (50). A research synthesis reported by Bays concluded that independence with ADL, and social support were key variables in the quality of life experienced by survivors of stroke (51). Lund et al. determined that social participation in various activities of life, including self-care, contributed to perceived quality of life in survivors of spinal cord injury (52). Cardol et al. (53) asserted that ethical approaches to planning and implementing care in rehabilitation should place greater emphasis on the autonomy of the individual. This is exemplified by an attentive attitude, opportunities for informed choices by the patient, and consideration for each patient’s preferences, needs, and social contexts (53). In some cases, active participation in goal setting by persons receiving rehabilitation may require special efforts to overcome lack of familiarity, perceived indifference, and other barriers to involvement (54).

**Self-care and Functional Performance**

Traditionally, intervention for people who have difficulty performing self-care tasks has begun with training in the acute care or rehabilitation environment. Typically, such intervention includes instruction in procedures to regain dressing,
grooming, hygiene, and food preparation and eating skills. In pursuit of these goals, rehabilitation sessions have been conducted within the patient’s hospital room or in simulated ADL settings within the facility. Intervention strategies involve teaching the individual functional skills or the use of assistive technologies so that compensatory strategies can be performed in the postdischarge environment.

Unfortunately, as suggested earlier, ADL training in a rehabilitation facility does not guarantee skill generalization to the discharge location. Patients may perform well in a rehabilitation facility, but skills may not always transfer to the individual’s pre-admission or discharge living setting. Environmental and psychosocial factors that directly influence task performance may be too varied between settings for the person receiving rehabilitation to generalize the learned skills (55). In addition, the individual may become dependent on the staff for self-care performance (56) or lack the opportunity to practice new skills on a regular basis. Consequently, performance following discharge may reflect a lack of confidence or motivation.

**Setting Rehabilitation Goals for the Discharge Environment**

The growing cost of specialized rehabilitation care has led to a reduction in hospital lengths of stay and has resulted in more rapid discharge from special care facilities. More care is now provided in outpatient settings and in the home environment.

This has advantages, since intervention in the home can be beneficial in achieving certain self-care skills and community reintegration (57–59). Community-based intervention has the benefit that the environment can be evaluated in terms of architectural, transportation, and communication barriers and how these support or limit the individual’s daily living skills.

**Collaboratively Planning Self-care and ADL Goals**

Current standards in rehabilitation require the involvement of the person receiving care as well as family members or caregivers, as appropriate in planning intervention (60). Controlled studies have shown that active collaboration in rehabilitation goal setting increases client satisfaction with care (61).

When goals are set in collaboration with the individual receiving care, the motivation to learn and maintain a skill is better than if rehabilitation professionals or caregivers determine the goals. It also appears that agreement on goals may influence functional outcomes by establishing clearer and more realistic goals (62). Each self-care behavior should be evaluated to see if the individual is motivated to learn and maintain it. Some studies have shown differences in the extent to which professionals and persons receiving care have congruent views regarding rehabilitation goals (63) but generally support the value of client participation in decision-making about care (61,64–67). This underscores the need for close collaboration between providers and recipients of care when planning intervention.

One of the first options the professional and person receiving rehabilitation should explore concerning the performance of any self-care task is whether the task is necessary or desired. The individual may choose not to perform some self-care tasks that were done before his or her illness. For example, a woman with hemiplegia who formerly rolled her hair on rollers on a daily basis may decide to have it cut in an easier-to-manage style rather than learn to use rollers with one hand. This type of decision should be based on individual preferences. Similarly, changes in societal styles and norms may also influence self-care goals, since greater diversity in clothing, hairstyle, and general appearance make it less likely that deviations from the norm will stand out.

In some instances, training procedures can be used to regain a desired skill. Following a cerebrovascular accident (CVA), for example, the therapist may be able to retrain the person to perform the task as it was performed before the CVA if there is sufficient return of voluntary movement. In some instances, the individual may no longer have the perceptual or physical capability to perform a task as before. However, he or she may be able to learn to accomplish the task using different movement patterns or with different body parts.

Environmental changes represent an additional array of intervention options that can be explored by the individual and his or her rehabilitation team as a means of gaining independence in self-care. In some instances, simply rearranging the physical environment may allow the disabled person to perform tasks independently. For example, moving dishes to lower shelves so that the patient can reach them from a wheelchair would represent a modification of the environment requiring only simple rearrangement. Structural changes in the physical environment also may be necessary. These can include major changes such as the architectural modification of rooms to accommodate wheelchair movement or less extensive improvements such as replacing round doorknobs with lever handles for a person who has weak grasp or installing bathroom rails and grab bars for persons with unsteady gait or balance difficulties.

The idea of universal design, which describes key principles for making environments, facilities, and objects useable for people regardless of their physical attributes or limitations, should have a positive impact on reducing barriers to activity and participation in the years ahead (68). This emerging environmental movement is broader than previous concepts of environmental accessibility (as described in the Americans with Disabilities Act and other legislation), yet highly relevant to rehabilitation and disability. Universal design emphasizes creating environments and objects that are simple and intuitive and that enable equitable and flexible use, have perceptible forms of information, require low physical effort, have tolerance for error, and have sizes and shapes appropriate for approach and use.

Assistive technology devices (ATDs) can be used to aid in the satisfactory performance of a desired task. These devices can range from simple, inexpensive articles, such as bathtub seats, to the use of expensive equipment such as computers for environmental control and communication (69). Many labor-saving devices are now widely available in catalogs and retail outlets catering to the general population. A line of fashionable apparel
designed for easy dressing and maintenance is now available for persons with limitations in range of motion. The rehabilitation professional’s role is to inform the patient of the existence and cost of these devices and to train the individual and caregiver in their use and maintenance.

Assistance from other people for the partial or total completion of a desired task is another option available to the individual receiving care (70). Assistance may come from spouses, friends, or paid PCAs. The role of the professional in this case must be to instruct the individual and/or the care attendant on optimal approaches to working together for the completion of identified self-care tasks (71). Figure 9-1 provides a decision chart that describes the process of goal setting summarized in this section.

Collectively, the personal and environmental intervention options described in this section form the basis for collaborative decision-making and treatment planning. It should be borne in mind that neither diagnosis alone nor the extent of impairment can serve as an adequate basis for planning self-care intervention. Together, the rehabilitation team and the individual receiving care must determine those approaches that represent the most realistic and achievable goals based on the abilities, values, and personal social circumstances of the recipient of care (61). Only in this way will optimal results be achieved after discharge.

**ISSUES IN ADL EVALUATION**

Granger (72) defines functional assessment as “a method for describing abilities and limitations and to measure an individual’s use of the variety of skills included in performing tasks necessary to daily living, leisure activities, vocational pursuits, social interactions, and other required behaviors.” Assessment must take place within a conceptual framework. An early model for this purpose was the World Health Organization’s International Classification for Impairment, Disability, and Handicap (ICIDH) (73). The ICIDH was revised in 2001 after an international revision effort, and is now called the International Classification of Functioning, Disability, and Health, or ICF (13). The ICF and its revised classification broadens the overall scope of the structure to include categories that view limitations to activities and participation as functional
consequences that are influenced by environment, personal factors, and problems with body function or structure. The ICF also addresses concerns about terminology (such as use of the word handicap) and adopts terms and definitions that are global in application and reflect differences between developed and underdeveloped nations. The revised model for the ICF now has two parts, each with two components with defined domains and constructs as reflected in Table 9-2.

As a conceptual framework for considering self-care, the ICF provides a means for recognizing that self-care tasks take place within particular living environments and may be performed differently based on an individual’s habits, culture, and social situation. This chapter attempts to emphasize that these contexts should be reflected in the functional assessment process, the determination of rehabilitation goals, and the type and location of interventions chosen.

Assessment has, as its ultimate purpose, the ability to make informed decisions. Scales and instruments designed to assess the ability of the individual to perform self-care tasks may assist in intervention or discharge planning by describing or documenting current abilities or monitoring changes in functional status. More global scales, which may include self-care components, are used to provide information on the effectiveness of rehabilitation programs, thus playing an important role in program evaluation.

**Historical Development of Self-care Evaluation Tools**

Assessment of the individual’s ability to function independently has been conducted in medical rehabilitation for nearly 50 years. In an early review of the problems of measurement and evaluation in rehabilitation published in 1962, Kelman and Willner found that poorly conceptualized outcome criteria, lack of standardization, disagreement about methods, multidimensional scales, and the influence of the setting on performance were barriers to effective management (74).

In the ensuing two decades, several published reviews of functional assessment from the standpoint of relevance, clinical utility, scale construction, standardization, reliability, and validity of scales concluded that the development of new scales should be curtailed, with greater effort devoted to the refinement and validation of existing scales (75–79). The challenges of functional assessment are summarized in the following sections.

**Determining Capability Versus Characteristic Behavior**

Current models of rehabilitation emphasize that function cannot be considered in isolation from its environmental context. This is made especially clear by the distinction between capability and actual behavior. Measures of capability represent what the person can do, whereas measures of actual behavior indicate what the person does.

Many self-care assessments in use during the past 30 years have been designed to measure what a patient is capable of doing within the care facility. An assessment of actual behavior, however, should ideally take place in the daily living environment in which the person will be performing the tasks. The setting influences actual performance that may differ from the individual’s capability. Awareness of the distinction between capability and actual behavior and their relationship to the patient’s environment has had an important impact on the development of new approaches to the assessment of self-care abilities. The ICF is based on a premise that disability is the result of an interaction between the person’s physical capabilities and the characteristics of a setting or environment. Environments have characteristics that can support or interfere with performance. It seems appropriate then that self-care performance assessment, if it is to be accurate and complete, should consider both the person’s physical capabilities and the characteristics of the environment.

As clients return home from hospitals, units, or residential facilities at an earlier stage, ease of administration becomes an
issue for determination of client progress. Although one way of determining ability is through observation of performance, another approach involves asking individuals themselves (or people living with them) to provide verbal reports. The literature is mixed about the validity of such reports. A study of stroke survivors by Knapp and Hewison found that caregivers generally perceived stroke survivors as more functionally impaired than did the survivors themselves (80). In contrast, a study by Wilz found that stroke survivors were able to accurately assess their level of impairment (81). Confusing the issue further, a study by Kwakkel et al. demonstrated that physical and occupational therapists consistently underestimated the degree of independence in ADL and level of recovery achieved by patients undergoing rehabilitation following stroke (82).

The point here is that the most accurate way to measure self-care performance is through actual observation of the individual within the performance environment. Even then, however, the maintenance or consistency of the individual’s performance over time will be unknown.

Contextual Factors in Self-Care Evaluation

A complete assessment of self-care function should consider factors that include the patient’s ability to manage devices that extend independence through environmental control; the family resources available to the patient in the environment to which he or she is to be discharged; the amount of time or energy required to perform tasks independently; and the degree of safety with which patients are able to perform tasks.

Developments in high technology for independent living have made it possible to assess people with severe disability in terms of their available movements and physical resources for controlling switches to activate environmental control units. Paradoxically, these devices are more likely to extend the patient’s ability to perform IADL more proficiently than self-care tasks.

The inability to perform self-care tasks independently, of course, does not require discharge to institutional care if human resources are available in an alternative environment. Frequently, the client can, and does, rely extensively on the assistance of a partner, other family members, or friends, or PCA to assist with self-care tasks. It can be argued that the presence of these resources, although commonly determined by the social worker in planning discharge, should be given early consideration in planning rehabilitation intervention.

Additional contextual considerations include the amount of time and energy required to perform the task independently versus the value of the task as perceived by the patient. It cannot be assumed, given competing requirements for time and energy, that all patients assign the independent performance of all self-care tasks the same degree of importance. Thus, the motivation to complete the task independently after discharge is likely to be a function of the alternatives available for task completion, the importance of the task to the patient, and the amount of time and energy required to perform the task independently in the face of competing demands.

Weingarden and Martin reported on a study of ten postdischarge spinal cord–injured patients to determine if time was a factor in the decision to retain, modify, or completely delegate dressing activities (83). Although all ten patients were capable of dressing independently at home, none did so routinely. The authors concluded that the person’s concepts of appropriate time and energy expenditure are important considerations in postdischarge decisions on the use of functional skills. Another study by Pentland, Harvey, and Walker explored time use by 312 men with spinal cord injury living in the community. Their study showed that the level of the lesion, the level of functional independence, and the setting did not predict the amount of time spent in personal care, productivity, leisure, and sleep. Nor did the time spent in these activities predict outcomes such as life satisfaction, perceived health, or adjustment to disability (84).

The degree of safety with which a task is performed may be of obvious importance to the practitioner but may not be apparent to the patient or those caregivers in the environment who may be providing assistance with self-care tasks. It is, therefore, important that training in self-care assistance be provided as part of the rehabilitation effort and that the ability of helpers to render this assistance in a safe and effective manner be assessed before the patient is discharged (85).

Defining Functional Outcomes

Many observers have noted that increased emphasis on accountability and the need for determining the benefit-cost ratios of rehabilitation have revealed ambiguities and a lack of consensus regarding definitions of rehabilitation success. For example, gains in self-care ability, although important to the patient, may not be perceived as beneficial within a system that perceives employability as the sole criterion of success. This has created additional pressure for the development of evaluation tools that consider post-discharge function in the living setting.

Fortunately, increased attention to these issues has resulted in increased research, which has encouraged the refinement or development and validation of several scales that assess self-care performance. Some of these scales possess characteristics suitable for program evaluation and research as well as clinical decision-making. The need remains, however, for greater awareness of the problems associated with functional assessment, greater consensus on appropriate measurement items, and consistency in definitions and terminology. Although awareness of the importance of using instruments that possess necessary measurement characteristics has improved, a tendency to create ad hoc instrument modifications continues, making comparisons between studies more difficult.

Approaches to Obtaining Performance Data

Functional evaluation can be accomplished by direct and indirect means. Direct evaluation involves first-hand observation of skill performance. Judgment based on subjective observation, however, is not always an accurate predictor of outcome (82). Indirect evaluation can be done by using client report
(gathered through mailed surveys, face-to-face or telephone interviews) or reported by proxies, such as members of the family or household. Evidence of the validity of such indirect approaches is equivocal (80).

In addition to creating threats to reliability, indirect reporting of functional performance can create threats to validity. A number of factors make it difficult to establish the validity of functional ability indices. These can include the impact of assistive technologies, environmental differences, and caregivers. Psychological factors, such as the patient’s level of motivation, the professional perspective of the rater, and the role expectations of the patient can compromise efforts at establishing the validity of these scales (86). There are also sources of error in observation, including the likelihood that the client’s performance is influenced by the evaluation process and therefore may not represent a typical or consistent performance.

Desirable Attributes of Self-care Scales

Reviews of functional status measures have identified several criteria for evaluating the quality or suitability of scales. These criteria derive both from psychometric standards and from expert opinion, and include psychometric properties of validity, reliability, sensitivity to change, and methods for assessing scaling properties, such as Rasch analysis (87, 88). Psychometric standards in the United States have been greatly influenced by the jointly developed standards of the American Psychological Association, the American Educational Research Association and the National Council on Measurement in Education (89). Within rehabilitation, the American Physical Therapy Association (APTA) (90, 91) published measurement standards in 1991, and the American Congress of Rehabilitation Medicine developed Measurement Standards for Interdisciplinary Medical Rehabilitation in 1992 (92). These standards provide important guidelines for the appropriate use and interpretation of measures, including self-care scales. They also provide definitions of important terms relevant to test development, and refer to technical problems that should be avoided by developers and users. Table 9-3 summarizes important characteristics that can be used in evaluating, comparing, and selecting rehabilitation measures.

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standardization</td>
<td>Scale has explicitly stated procedures for administration and scoring, performance data from a normal population, preferably of varying ages; information on the measurement properties of the scale (e.g., its computed reliability); and a statement of the necessary qualifications of the examiner. It is difficult to standardize scales that are useful both as measures of capability and as measures of actual functioning, because of varying contextual factors.</td>
</tr>
<tr>
<td>Scalability</td>
<td>Scaling procedures serve to quantify a person’s responses to a defined set of tasks so that they are distributed along a Continuum of performance. For an assessment device to be considered a true scale, it must be established that the tasks performed will cumulatively yield a score or descriptor that represents increasing capability or independence.</td>
</tr>
<tr>
<td>Reliability</td>
<td>An acceptable scale should provide a reliable measure of the client’s level of performance. Reliability refers to a scale’s accuracy and consistency in providing information, regardless of the time, setting, or person performing the assessment. Scales that have carefully defined methods and scoring criteria are likely to be more sensitive and consistent, and present a more accurate picture of the client than those that do not.</td>
</tr>
<tr>
<td>Validity</td>
<td>Validity is related to theoretical and methodological issues and depends on a number of factors. These include the extent to which the scores on the assessment are related to some external criterion, the degree to which the instrument contains items or tasks that represent the domain of interest, and the relationship of the instrument to other measures that collectively support various theoretical assumptions. A scale cannot be valid if it is not reliable.</td>
</tr>
<tr>
<td>Comprehensive</td>
<td>Self-care assessments are more useful if they determine performance levels for all BADL skills and are applicable to everyday diagnosis. Because clients have the universal need to perform or have performed for them basic self-care tasks, broad applicability is appropriate and facilitates comparison of research findings among differing client groups. To the extent that extended ADLs can be assessed, a scale becomes more useful as a measure of independence in everyday situations.</td>
</tr>
<tr>
<td>Performance-based</td>
<td>This characteristic eliminates measurement error due to selected observations or inaccurate memory. Although self-care abilities are increasingly being evaluated through interviews and surveys, this is most advisable only after discharge and if validation of such approaches has been demonstrated. Although performance-based scales should not emphasize speed, time can become an issue if the extent of time required makes it impractical.</td>
</tr>
<tr>
<td>Practical</td>
<td>The extent to which a scale is designed to facilitate decision-making and research is an important consideration in instrument selection. The number of items should be sufficient to permit reliability while requiring a reasonable length of time (i.e., 30–45 min) for administration. To the extent possible, items should not require equipment that would unduly limit locations where the scale could be administered. Scoring sheets should contain clear explanations of criteria and should be designed to permit accurate recording as well as coding for data processing. Terminology should be readily understandable, with abilities expressed using everyday language. The meaning of scores should clearly convey a client’s level of functional independence to professionals and caregivers who will be living with the client.</td>
</tr>
</tbody>
</table>
Resolving Measurement Problems in Functional Evaluation

A topic of considerable importance in the area of functional assessment is that of how numbers are assigned to performance observations and later interpreted. Most self-care scales use ordinal rankings to assign numbers to patient performance on selected items. Merbitz et al. (93) identified the limitations of such rankings, based on the characteristics of ordinal measurement scales. They pointed out that the properties of ordinal scales do not permit valid conclusions to be drawn on the summed or averaged scores obtained from them and that their misuse in this way is misleading and subject to invalid inferences. Further, because they are given equal weight in scoring, the implied assumption of most scales of functional ability is that the items being measured are equally important, either to the professional or to the individual being measured. This is clearly an invalid assumption.

Using techniques developed by Rasch (94) and others, however, observations of functional ability (from items and persons) can be translated into linear measures. The logit unit of measurement produced by the Rasch model is the natural log of the odds of a correct response. This transformation allows the researcher to interpret the person and item information using the same units of measure. Once the scores for the persons and items have been transformed using Rasch scaling, other information about the persons and items can be obtained using a variety of statistical methods. These transformed scales possess the properties necessary for valid inferences (95). In effect, these statistical transformations lead to scores, which are corrected for differences in items or raters. Such techniques are increasingly being applied to existing scales (96).

Patient Collaboration in Evaluation and Goal Setting

No discussion of functional evaluation is complete without a consideration of the values that may be reflected in the process. Although functional assessment is a judgmental process, it should not be unnecessarily value laden. In fact, the rehabilitation goal of independent functioning itself reflects a societal value not shared to the same extent by all cultures. There are profound differences in the cultural heritages and life experiences of individuals, and these differences bring differing sets of values about independence and self-care to the rehabilitation setting. Effective management of the client requires an appreciation of these value differences, and an appreciation that interdependence rather than independence is a condition typical of societal groups. It is, therefore, important for the overall functional evaluation to include methods for determining premorbid activity patterns and leisure interests as well as values and attitudes toward assistance. Characteristic methods for performing BADL and EADLs, as well as the characteristic aspects of the environment in which these have been performed, are important information in planning successful intervention strategies for management of personal care and other ADL necessary for life satisfaction and quality.

Selected Measures of Basic ADL (Self-care)

Although dozens of instruments designed to assess personal or self-care performance have been reported in the literature, consensus has not been achieved for use of a single scale. Scales range from those that focus on personal or self-care ADL skills (BADLs) to those that focus on IADL or EADL skills. Still others are more global and include physical functioning and/or EADL skills (Table 9-4). These have been called global scales.

One such global scale is the Functional Independence Measure (FIM). The past decade has shown a marked growth in the use of the FIM developed in the United States by Granger et al. (72) in the 1980s and now translated into several languages. The FIM and its associated measurement systems have been adopted as a key part of decisions influencing public funding of rehabilitation services in the United States. The FIM has not been embraced universally, however, and the Barthel Index (BI), developed by Mahoney and Barthel (97), with its modifications, continues to be used extensively throughout the world.

Besides the BI and the FIM, a handful of scales have demonstrated sufficient validity and frequency in the literature to warrant their review here. This review is not intended to be exhaustive, and it is acknowledged that there are other measures with self-care components that have been developed or are in use for specialized purposes that are not described here. To gain a complete picture of the client’s level of function and community reintegration within the new WHO (ICF) framework requires that items pertaining to household maintenance and community functioning also be considered. For this reason, additional items relating to household and community functioning often accompany the assessment of basic self-care items. In studying the association between the three widely used measures of functional change (FIM, BI, and Functional Assessment Measure [FAM]) and core sets of statements representing the domains of the ICF, Grill et al. concluded that many aspects of human functioning are not measured, making it necessary to add supplementary measures (98).

The following paragraphs report the composition, measurement properties, and scoring of several scales. Brief summaries of selected studies describing their use with varying client populations are also provided. Where available, evidence reporting reliability and validity is summarized. The scales to be reviewed in this section are the BI, the Katz Index of Independence in ADL (99), and the FIM (100).

The Barthel Index

In 1965, Mahoney and Barthel published a weighted scale for measuring BADL in disabled patients (Table 9-5) (97). The initial BI included ten items, including feeding, transfers, personal grooming and hygiene, bathing, toileting, walking, negotiating stairs, and controlling bowel and bladder. Items are scored differentially according to a weighted scoring system that assigns points based on independent or assisted performance. For example, a person who needs assistance in eating
### TABLE 9.4 Measures of ADL/IADL

<table>
<thead>
<tr>
<th>Assessment Tool</th>
<th>Authors/Year of Publication</th>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arnadottir (A-1) OT-ADL Neurobehavioral Evaluation</td>
<td>Arnadottir (1990)</td>
<td>Observation (late adolescent-adult)</td>
<td>This instrument is an observational tool used to look at neurobehavioral deficits through observation of self-care activities such as feeding, grooming, dressing, transfers, and functional communication.</td>
</tr>
<tr>
<td>Assessment of Motor and Process Skills (6th ed.)</td>
<td>Fisher (2006)</td>
<td>Observation (children and adult)</td>
<td>The original version of this instrument was designed to assess motor process skills through observation of 16 ADL motor and 20 ADL process skills. There is also a school version of the AMPS.</td>
</tr>
<tr>
<td>BI of ADL</td>
<td>Mahoney and Barthel (1985)</td>
<td>Verbal report/observation (adult)</td>
<td>This assessment addresses ten ADL areas including feeding, bathing, grooming, dressing, bowel and bladder control, toilet use, transfers, mobility, and stair climbing.</td>
</tr>
<tr>
<td>Behavioral Assessment Scale of Oral Functions in Feeding</td>
<td>Stratton (1981)</td>
<td>Observation (children)</td>
<td>This tool examines a child's feeding skills including items such as lip closure, tongue control, sipping, etc. Children are rated on a scale ranging from passive response to active participation.</td>
</tr>
<tr>
<td>Canadian Occupational Performance Measure (4th ed.)</td>
<td>Law et al. (2005)</td>
<td>Interview (children and adults)</td>
<td>Individualized client-centered interview evaluating self perception, may be used with an older child, or in conjunction with a caregiver.</td>
</tr>
<tr>
<td>Child Occupational Self-Assessment (version 2.1)</td>
<td>Keller et al. (2005)</td>
<td>Self-rating scale (children)</td>
<td>The child version of the Occupational Self-Assessment uses happy faces and stars to facilitate child self-evaluation of perceived occupational competence in a variety of occupational areas.</td>
</tr>
<tr>
<td>Frenchay Activities Index</td>
<td>Holbrook and Skilbeck (1983)</td>
<td>Self-report with recording form (adults)</td>
<td>Although originally developed for clients with stroke, this tool has been expanded to other users. Guidelines for scoring include descriptions of examples in each area, and amount of client involvement in doing the activity.</td>
</tr>
<tr>
<td>Functional Assessment Measure</td>
<td>Hall (1997)</td>
<td>Used in conjunction with the FIM</td>
<td>This tool was designed by clinicians in order to be used along with the FIM. The tool addresses cognitive, behavioral, communication, and community functioning measures.</td>
</tr>
<tr>
<td>Functional Evaluation for Assistive Technology</td>
<td>Raskind and Bryant (2002)</td>
<td>Checklist/observation and Interview (children and adults)</td>
<td>This evaluation takes an ecological approach using five scales to determine appropriate assistive technology for the client.</td>
</tr>
<tr>
<td>FIM (The Center for Functional Assessment Research at SUNY, Buffalo, 1999)</td>
<td>The Center for Functional Assessment Research at SUNY, Buffalo, 1999</td>
<td>Observation/history taking (adults; child's version available as well)</td>
<td>This tool is an 18-item ordinal scale encompassing seven levels designed to identify areas of dysfunction and function in self-care.</td>
</tr>
<tr>
<td>FIM for Children WeeFIM II</td>
<td>Hamilton and Granger (2005); Center for Functional Assessment Research, 2005</td>
<td>Observation/history taking (children)</td>
<td>The WeeFIM was recently updated to the WeeFIM II and now includes a module for 0–3 in addition to the tool for older children. The original WeeFIM consists of 18 items assessing self-care, motor and cognition—this is designed to measure the amount of caregiver assistance needed. The 0–3 module measures motor, cognitive, and behavioral performance.</td>
</tr>
<tr>
<td>Independent Living Scales</td>
<td>Loeb (1996)</td>
<td>Observational (adults)</td>
<td>The ILS tool is designed to assess cognitive skills necessary for independent living. Five scales include memory/orientation, managing money, managing home and transportation, health and safety, and social adjustment.</td>
</tr>
<tr>
<td>Assessment Tool</td>
<td>Authors/Year of Publication</td>
<td>Type</td>
<td>Description</td>
</tr>
<tr>
<td>----------------</td>
<td>----------------------------</td>
<td>------</td>
<td>-------------</td>
</tr>
<tr>
<td>Instrumental Activities of Daily Living Scale (IADLS)</td>
<td>Lawton and Brody (1969)</td>
<td>Client/practitioner interview/survey (adults)</td>
<td>This interview-based tool covers eight daily living tasks including use of phone, shopping, food preparation, housekeeping, laundry, transportation, medications, and finances. As with many self reports or interviews, the utility of the tool depends on the reliability/insight of the client, or familiarity of the practitioner with the client's functioning level.</td>
</tr>
<tr>
<td>Katz Index of ADL</td>
<td>Katz et al. (1963, 1970)</td>
<td>Interview/observation (adults)</td>
<td>Short assessment/index of ADL function including bathing, dressing, toileting, continence, transferring, and feeding.</td>
</tr>
<tr>
<td>Klein Bell Activities of Daily Living Scale</td>
<td>Klein and Bell (1982)</td>
<td>Observational rating scale (children and adults)</td>
<td>This scale considers ADL items in six different functional areas including dressing, elimination, mobility, bathing, hygiene, eating, and emergency/phone.</td>
</tr>
<tr>
<td>Kohlman Evaluation of Living Skills (3rd ed.)</td>
<td>Kohlman Thomson (1992)</td>
<td>Observation/interview (adults)</td>
<td>Typically designed for the psychiatric population, some studies have expanded beyond this area. The tool rates clients in ADL areas including self-care, safety and health, money management, transportation and telephone, and community skills.</td>
</tr>
<tr>
<td>Melbourne Low Vision ADL</td>
<td>Haymes et al. (2001)</td>
<td>Observation (adults)</td>
<td>This test is designed for persons with low vision and assesses current function and rehabilitative outcomes in ADL. The assessment uses both client interview and direct observation.</td>
</tr>
<tr>
<td>Melville-Nelson Self-Care Assessment</td>
<td>Nelson and Melville (2002)</td>
<td>Observational rating scale (adults)</td>
<td>This rating scale looks at client function in personal self-care ADL of bed mobility, transfers, dressing, eating, toileting, personal hygiene, and bathing.</td>
</tr>
<tr>
<td>Milwaukee Evaluation of Living Skills</td>
<td>Leonardelli (1988)</td>
<td>Observation (adults)</td>
<td>This evaluation was originally designed for the psychiatric population but is also used in psychiatric medical units. The evaluation consists of 20 subtests that may be used in their entirety or in any combination deemed appropriate to the client needs. Skill areas include areas such as communication, bathing, dressing, eating, hair care, safety in the home and community, use of money, etc.</td>
</tr>
<tr>
<td>Nottingham Extended Activities of Daily Living Index</td>
<td>Nouri and Lincoln (1987)</td>
<td>Self-report questionnaire (adults)</td>
<td>This evaluation is another self report instrument that measures 21 tasks in four areas including mobility, kitchen tasks, domestic tasks, and leisure.</td>
</tr>
<tr>
<td>Performance Assessment of Self-Care Skills (version 3.1)</td>
<td>Rogers and Holm (1994)</td>
<td>Observational rating scale (adults)</td>
<td>This instrument measures ADL and IADL in either the clinic or home (two versions). Scales address functional mobility, personal care, physical IADL, and cognitive IADL.</td>
</tr>
<tr>
<td>Pediatric Evaluation of Disability (PEDI)</td>
<td>Haley et al. (1992)</td>
<td>Observation/inventory and interview (children)</td>
<td>The PEDI is a widely used standardized tool to measure function in children with disabilities age 6 mo to 7.5 y but has been applied to older children who have conditions placing them in a developmental functioning level lower than their age. Domains assessed include self-care, mobility, and social function.</td>
</tr>
<tr>
<td>Refined ADL Assessment Scale</td>
<td>Tappen (1994)</td>
<td>Observational rating scale (adults)</td>
<td>This observational scale assesses five ADLs including toileting, washing, grooming, dressing, and eating. The scale is designed to be used in the natural morning routine and outcomes indicate the level of assistance needed.</td>
</tr>
<tr>
<td>Safety Assessment of Function and the Environment for Rehabilitation (SAFER)</td>
<td>Oliver et al. (1993) and Chu et al. (2001, 2002)</td>
<td>Interview/observational checklist (adults)</td>
<td>This scale consists of questions divided into areas including ADL, addictions, and personal behaviors. Areas assessed are designed to look at current function and provide a baseline for future assessment.</td>
</tr>
</tbody>
</table>
would receive five points, whereas independence in eating would be awarded ten points. A client with a maximum score of 100 points is defined as continent, able to eat and dress independently, walk at least a block, and climb and descend stairs. The authors were careful to note that a maximum score did not necessarily signify independence, since IADLs such as cooking, housekeeping, and socialization are not assessed.

The BI may be the most widely studied and used self-care scale globally. Several studies have shown that the scale has acceptable psychometric properties, including that it is sensitive to change over time, that it is a significant predictor of rehabilitation outcome, and that it relates significantly with other measures of client status. Granger et al. reported the stability (i.e., test-retest reliability) of the BI as 0.89, whereas interrater reliability coefficients were above 0.95 (100). One study of interrater reliability between nursing staff and non-clinical research assistant showed that although an overall agreement was within acceptable boundaries, two items out of ten showed weakness for agreement (transfer and dressing) (101). Schlote et al. found that the BI had excellent reliability, including the

### TABLE 9.4 Measures of ADL/IADL (Continued)

<table>
<thead>
<tr>
<th>Assessment Tool</th>
<th>Authors/Year of Publication</th>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Structured Assessment of Independent Living Skills (SAILS)</td>
<td>Mahurin et al. (1991)</td>
<td>Observational rating scale (adults)</td>
<td>This tool provides an assessment of function through direct observation of 50 items in four domain areas of motor, cognitive IADL, and social interaction skills. It is a widely used tool by several professionals in five domain areas.</td>
</tr>
<tr>
<td>Vineland Adaptive Behavior Scales-II (Sparrow et al., 2005)</td>
<td>Observation/ Interview/Rating Scale (Sparrow et al., 1984)</td>
<td>BADI/IADL, communication, and behavior</td>
<td>This tool provides an assessment of function through direct observation of 50 items in four domain areas of motor, cognitive IADL, and social interaction skills. It is a widely used tool by several professionals in five domain areas.</td>
</tr>
</tbody>
</table>


correspondence of scores between physician and nonphysician raters of the status of stroke survivors (102), while Green et al. found that the BI had excellent test-retest reliability (103) using a similar population. Shortened versions of the BI (BI-3 and BI-5) have also been developed, and these have shown evidence of satisfactory psychometric characteristics and predictive validity acceptable for outcome use (104,105).

The initial BI score was found to be the most reliable predictor of final rehabilitation outcome in the study of stroke patients conducted by Hertanu et al. (106). This study concluded that the BI was a more reliable predictor of rehabilitation outcome than estimates based on computed tomography showing the extent of the lesion after CVA. The BI also has been found to correlate significantly with the type of discharge and shorter length of stay for patients with CVA (107,108) and independent living outcome for patients with spinal cord injury (109), as well as participation of young adults with disabilities (110). Scores on the BI and the PULSES Profile have been shown to correlate substantially (111), as have scores on the BI and the physical subscales of general health status measures such as the SF-36, the Sickness Impact Profile, and the Dartmouth Functional Health Assessment Charts (112). Other studies have similarly provided evidence of concurrent validity with other measures of ADL (113,114). In a clinical trial of stroke rehabilitation, van Hartingsveld et al. used Rasch analysis to weigh item scores of the BI and found improvement in determining results (114,115). Bowel and bladder items emerged as having the poorest fit in analyzing outcomes. Some studies have reported both floor and ceiling effects when the BI is used (113,116–118) or have questioned its utility in trials of outcomes with particular disease groups (119) or have recommended its use in the subacute phases of recovery (120).

A modified version of the BI and the extended BI (EBI) (122), was developed to address the perceived limitations of the FIM and existing BI by adding items for comprehension, expression, social interaction, problem solving, memory/learning/orientation, and vision/neglect. Separate studies of large groups of neurologic patients have shown that the EBI is a reliable, valid, and practical instrument that is sensitive to changes over time (122–124). Because of the manner in which the new items are administered, rater training is necessary. Time required to administer the EBI was described as significantly shorter than the time needed to administer the FIM.

The Katz Index of Independence in ADL
The Katz Index of Independence in ADL was developed to study results of treatment and prognosis in the elderly and in those with chronic illness (125). Development of the index was originally based on observations of a large number of activities performed by a group of patients with fracture of the hip.

The index is based on an evaluation of the functional independence of patients in bathing, dressing, going to the toilet, transfers, continence, and feeding. Using three descriptors for rating independence in each of the six subscales, the rater is able to derive an overall grade of independence with the aid of specific rating criteria. Depending on the determined level of independence, a client is graded as A, B, C, D, E, F, G, or Other. According to the scale, a client graded as A would be functioning independently in all six functions, whereas a client graded as G would be dependent in all rated functions. Patients graded as “Other” are dependent in at least two functions but not classifiable as C, D, E, or F. Through observations over a defined period of time, the observer determines whether the client is assisted or whether the client functions on his or her own when performing the six activities. Assistance is classified as active personal assistance, directive assistance, or supervision.
In studies of the Katz Index involving more than 1,000 patients, the scale was found to result in an ordered pattern, so that a person able to perform a given activity independently also would be able to perform all activities done by those graded at lower levels (99). This hierarchical structure correctly classifies the functional ability of patients 86% of the time and reflects a desirable property of scalability (126).

The Katz Index of Independence in ADL has been used as a tool to accumulate information about recovery after CVA (127–129) and amputation from peripheral vascular disease (130), as a means of providing information about the need for care among patients in late poliomyelitis (131,132), following hip fracture (133,134), in persons with dementia (135–137) and as an instrument to study information about the dynamics of disability in the aging process (138–140). Bronnson and Asberg used the scale in a study of internal medicine patients in a general hospital in Sweden (141). Their study found a high degree of interrater reliability and high coefficients of scalability. Reijneveld et al. have also reported successful use of the Katz scale with Dutch, Turkish, and Moroccan populations (142).

**Selected Measures of IADL**

Increasingly, as rehabilitation professionals give greater attention to the functional context of the client, more consideration is being given to IADLs, or the individual’s ability to perform those tasks beyond basic self-care necessary to live in the community. Wade (143) has suggested that these types of ADLs be described as extended activities of daily living (EADL). Other terms in use for similar scale items are social ADL and advanced ADL (144).

In this section, several scales that focus specifically on IADL (EADL) performance are reviewed, including the Assessment of Motor and Process Skills (AMPS), the Nottingham EADL Scale, and the Frenchay Activities Index (FAI). These scales are chosen because of their global use and the availability of published studies of their reliability and validity.

**Assessment of Motor and Process Skills**

The AMPS is an observational evaluation that is used to simultaneously examine both the ability to perform IADL and the underlying motor and process capacities necessary for successful performance (145). The AMPS is an assessment system that requires a clinician to observe a person performing IADL as he or she would normally perform them. The individual to be measured selects two or three familiar tasks from among more than 50 possibilities described in the AMPS manual. After the observation, the clinician rates the person’s performance in two skill areas: IADL motor and IADL process. Table 9-6 lists the motor and process areas. Motor skills are defined as observable actions that are supported by underlying abilities, including postural control, mobility, coordination, and strength. The AMPS motor items represent an observable taxonomy of actions used to move the body and objects during actual performance. Process skills reflect the organization and execution of a series of actions over time in order to complete a specified task. Thus, process skills may be related to a person’s underlying attentional, conceptual, organizational, and adaptive capabilities. Like the AMPS motor skill items, the AMPS process skill items represent a universal taxonomy of actions that can be observed during any task performance.

During each IADL task performed for the assessment, and for each of the 16 motor and 20 process skills (see Table 9-6), the person is rated on a four-point scale: 1 = deficit, 2 = ineffective, 3 = questionable, and 4 = competent. The raw ordinal scores are analyzed using the Rasch approach referred to as many-faceted Rasch analysis. This approach rests on a mathematical model of likelihood that the person will receive a given score on each of the motor and process skill items. The observed counts of the raw scores of IADL motor and process skill items constitute ordinal (ranked) data. These counts are converted by logistic transformation into additive, linear measures. Once the raw scores are computer analyzed, the derived person ability measures (motor and process) become estimates of the person’s position on the two AMPS scales. That is, the AMPS motor and process scales represent continua of increasing IADL motor or process skill ability, and the person’s estimated position on the AMPS motor and process scales, expressed in logits, represents his or her IADL motor and process skill ability (145).

The many-faceted Rasch analysis used in the AMPS allows simultaneous calibration of three aspects of performance: item

<table>
<thead>
<tr>
<th>MOTOR SKILL GROUPS AND ITEMS</th>
<th>PROCESS SKILL GROUPS AND ITEMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Posture</td>
<td>Using knowledge</td>
</tr>
<tr>
<td>Stabilizes</td>
<td>Chooses</td>
</tr>
<tr>
<td>Aligns</td>
<td>Uses</td>
</tr>
<tr>
<td>Positions</td>
<td>Handles</td>
</tr>
<tr>
<td>Mobility</td>
<td>Heeds</td>
</tr>
<tr>
<td>Walks</td>
<td>Inquires</td>
</tr>
<tr>
<td>Reaches</td>
<td>Notices</td>
</tr>
<tr>
<td>Bends</td>
<td>Temporal organization</td>
</tr>
<tr>
<td>Coordination</td>
<td>Initiates</td>
</tr>
<tr>
<td>Coordinates</td>
<td>Continues</td>
</tr>
<tr>
<td>Manipulates</td>
<td>Sequences</td>
</tr>
<tr>
<td>Flows</td>
<td>Terminates</td>
</tr>
<tr>
<td>Strength and effort</td>
<td>Space and objects</td>
</tr>
<tr>
<td>Moves</td>
<td>Searches</td>
</tr>
<tr>
<td>Transports</td>
<td>Gathers</td>
</tr>
<tr>
<td>Lifts</td>
<td>Organizes</td>
</tr>
<tr>
<td>Calibrates</td>
<td>Restores</td>
</tr>
<tr>
<td>Grips</td>
<td>Adaptation</td>
</tr>
<tr>
<td>Calibrates</td>
<td>Accommodates</td>
</tr>
<tr>
<td>Endures</td>
<td>Adjusts</td>
</tr>
<tr>
<td>Energy</td>
<td>Navigates</td>
</tr>
<tr>
<td>Pacers</td>
<td>Benefits</td>
</tr>
<tr>
<td>Paces</td>
<td>Energy</td>
</tr>
<tr>
<td></td>
<td>Pacers</td>
</tr>
<tr>
<td></td>
<td>Attends</td>
</tr>
</tbody>
</table>
easiness, task simplicity, and rater leniency. Each of these item characteristics is determined by using a probabilistic model. The ability measure produced by the Rasch analysis is the estimated person ability plotted on a linear scale and is defined by the skill item easiness and task simplicity but adjusted for the rater who scored the task performance (151).

Because the person ability measures on the AMPS are adjusted for task simplicity, a clinician can use the ability measure to predict whether a person possesses the motor and process skills necessary to perform tasks that are more difficult than those the person was observed performing. Also, since the AMPS includes 50 possible IADL tasks and each person is observed performing only two or three, the number of possible alternative task combinations is very large. Regardless of how many different tasks the individual performs, however, the ability measure will always be adjusted to account for the ease and simplicity of those particular tasks, so direct comparisons can be made among persons even though they performed completely different tasks.

Several investigations using the AMPS with persons who have psychiatric (146) orthopedic, neurologic, cognitive (147,148), and developmental disabilities (149,150) have been reported. The ability of the AMPS to enable analysis of the separate contributions of motor and process variables has provided support for theoretical assumptions about the specific variables contributing to task limitations in various conditions. Validity studies have suggested that the AMPS can help predict home safety (151) and measure improvements following intervention in psychiatry (152–155), developmental disabilities (156); multiple sclerosis (157,158) and stroke (159–162); and acquired brain injury (163,164), cerebral palsy (165,166), dementia (167–172), and in students with difficulties interfering with function in school (173–175). The AMPS has demonstrated consistency across gender (176), raters (177), settings (149), and different cultural groups (178,179). These investigations have established the reliability (180) and validity of the AMPS. The AMPS approaches the assessment of functional performance using many-faceted Rasch analysis, with clients selecting the item to be performed from an inventory of possible IADL items.

**The Nottingham EADL Index**

The Nottingham EADL Index (NEADL) was developed by Nouri and Lincoln (181) in 1987 and is widely used throughout Europe and in other countries. This self-report scale has 21 items organized into four sections, which can serve as subscales. The sections include mobility, kitchen tasks, domestic tasks, and leisure activities. Scoring is along a four-item range of discrete categories, ranging from not done at all to done alone easily. Unfortunately, there are no guidelines for assigning scores. Because of the range of EADL tasks reported, the scale has intuitive appeal as an outcome measure of rehabilitation and social participation. Table 9-7 lists specific items included in the Nottingham EADL.

<table>
<thead>
<tr>
<th>TABLE 9.7</th>
<th>Nottingham Extended ADL Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mobility Questions—Do you:</td>
<td>Not at All</td>
</tr>
<tr>
<td>Walk around outside?</td>
<td></td>
</tr>
<tr>
<td>Climb stairs?</td>
<td></td>
</tr>
<tr>
<td>Get in and out of the car?</td>
<td></td>
</tr>
<tr>
<td>Walk over uneven ground?</td>
<td></td>
</tr>
<tr>
<td>Cross roads?</td>
<td></td>
</tr>
<tr>
<td>Travel on public transport?</td>
<td></td>
</tr>
<tr>
<td>In the Kitchen—Do you</td>
<td></td>
</tr>
<tr>
<td>Manage to feed yourself?</td>
<td></td>
</tr>
<tr>
<td>Manage to make yourself a hot drink?</td>
<td></td>
</tr>
<tr>
<td>Take hot drinks from one room to another?</td>
<td></td>
</tr>
<tr>
<td>Do the washing up?</td>
<td></td>
</tr>
<tr>
<td>Make yourself a hot snack?</td>
<td></td>
</tr>
<tr>
<td>Domestic Tasks—Do you:</td>
<td></td>
</tr>
<tr>
<td>Manage your own money when you are out?</td>
<td></td>
</tr>
<tr>
<td>Wash small items of clothing?</td>
<td></td>
</tr>
<tr>
<td>Do your own shopping?</td>
<td></td>
</tr>
<tr>
<td>Do a full clothes wash?</td>
<td></td>
</tr>
<tr>
<td>Leisure Activities—Do you:</td>
<td></td>
</tr>
<tr>
<td>Read newspapers or books?</td>
<td></td>
</tr>
<tr>
<td>Use the telephone?</td>
<td></td>
</tr>
<tr>
<td>Go out socially?</td>
<td></td>
</tr>
<tr>
<td>Manage your own garden?</td>
<td></td>
</tr>
<tr>
<td>Drive a car?</td>
<td></td>
</tr>
</tbody>
</table>
A growing literature on use of the Nottingham scale is beginning to demonstrate that it is a suitable instrument for evaluating EADL function in the community. Although most reported studies involve subjects who have received rehabilitation following stroke (182–184), the measure has been used with other diagnostic groups such as those with pulmonary problems (185–187), elderly persons in the community (188,189), and patients with hip replacement (190). For the latter group, Harwood and Ebrahim compared the responsiveness of the Nottingham EADL with two other measures (191).

In this study, the NEADL was not viewed as sensitive to change in function as the SF-36 or the London Handicap Scale as a measure of activity and social participation for patients with hip replacement. Despite this the scales have shown evidence of acceptable scalability, concurrent validity, and construct validity. For example, Hsueh et al. (192) used the Nottingham EADL in a study of stroke patients in Taiwan. They found that with minor modification to two items, the Nottingham EADL had satisfactory scalability and reproducibility and correlated with age and scores on the BI (192). Schloet et al. (102) studied the reliability of several scales in measuring stroke outcomes. They found that 25% of the items did not attain satisfactory reliability levels, but overall the scale showed high interrater correspondence. The NEADL scale has also been useful as a measure of the effectiveness of rehabilitation strategies, such as ADL training (193) and as an outcome measure in several studies of stroke rehabilitation (194–204) and in studies of the effectiveness of day hospitals, intermediate care facilities, and community-based care programs in facilitating independence (199,203,205).

**Frenchay Activities Index**

The FAI was developed initially for use in clinical social work for stroke patients and has emerged as a frequently used measure of EADL (206). The index is designed as a mailed questionnaire to be completed by self-report. The FAI consists of 15 items divided into two sections or subscales. The first section pertains to activities performed within the 3 months preceding completion of the scale and includes standard mobility, household maintenance, and meal preparation items. The second section pertains to items performed in the 6 months before scale completion and reports on work, leisure, travel, and household/car maintenance. Items are scored on a four-point scale from 0 to 3 according to very well defined guidelines.

Turnbull et al. studied 1,280 people over 16 to construct preliminary norms and to determine evidence of reliability and validity (207). They concluded that the FAI is reliable and shows good evidence of validity with an elderly population but would benefit from adding items relating to sport, physical exercise, and caring for children. This would make it a tool more useful for a broader segment of the population (207). Green et al. studied the test-retest reliability of the FAI and other scales of stroke outcome. They found that the FAI had only moderate reliability using kappa coefficients (mean differences for items 0.6 ± 7.1) and had higher random error when stroke survivors were measured twice within a 1-week interval (103).

Although most of the studies using the FAI have been related to outcomes following stroke (112,208–226), the index has also been used for other populations, including those with complex disabilities (227), lower limb amputation (228,229), acquired brain injury (230), multiple sclerosis (231), and caregivers (162). The scale has also been translated for and used to study rehabilitation outcomes in Japan (232), China (233), Denmark (224), and Spain (234,235).

Carter et al. compared the postal version of the FAI with an interviewer-administered version, using a population of stroke survivors discharged from the hospital (236). Kappa statistics for individual item agreement for the subjects tested ranged from 0.35 to 1.00. Overall, the correlation between total scores was 0.94. Piercy et al. estimated the intrarater reliability of the FAI as 0.93 in a study of interview administered versions of the scale by two raters (210). A study by Schuling and others suggested that the reliability of the FAI could be improved by deleting two items and by creating two subscale scores, one for domestic and the other for outdoor activities (237). Studies using the FAI and measures of BADL (particularly, the BI) have demonstrated that the scales measure different factors and may be useful in combination (233). Chen et al. found wide differences between reports by proxies and patients using the self-administered version of the FAI following stroke and against the use of this version for clinical purposes (238).

**Combined ADL/IADL Measures**

**The Functional Independence Measure**

The FIM evolved from a Task Force of the American Congress of Rehabilitation Medicine and the American Academy of Physical Medicine and Rehabilitation, which met to develop a reliable and valid instrument that could be used to document the severity of disability as well as the outcomes of rehabilitation treatment as part of a uniform data system (239,240).

The FIM consists of 18 items organized under six categories, including self-care (e.g., eating, grooming, bathing, upper-body dressing, lower-body dressing, and toileting); sphincter control (i.e., bowel and bladder management); mobility (e.g., transfers for toilet, tub or shower, and bed, chair, wheelchair); locomotion (e.g., walking, wheelchair, and stairs); communication, including comprehension and expression; and social cognition (e.g., social interaction, problem solving, and memory). Using the FIM, patients are assessed on each item with a seven-point scale, ranging from complete independence (value = 7) to complete dependence (total assistance required = 1 (Fig. 9-2) (241).

The original components of the FIM were developed under a federal grant to the State University of New York at Buffalo. The FIM is now part of a system of data collection and outcome under proprietary license held through association with that university. The Uniform Data System for Medical Rehabilitation (UDSMR) holds the rights to the FIM.
FUNCTIONAL INDEPENDENCE MEASURE™ AND FUNCTIONAL ASSESSMENT MEASURE

Scale:
7. Complete Independence (timely, safely)  (Patient Stamps)
6. Modified Independence (extra time, devices)
5. Supervision (caring, coaching, prompting)
4. Minimal Assist (performs 75% or more of task)
3. Moderate Assist (performs 50%-74% of task)
2. Maximal Assist (performs 25%-49% of task)
1. Total Assist (performs less than 25% of task)

SELF CARE ITEMS
1. Feeding
2. Grooming
3. Bathing
4. Dressing Upper Body
5. Dressing Lower Body
6. Toileting
7. Swallowing*

SPHINCTER CONTROL
8. Bladder Management
9. Bowel Management

MOBILITY ITEMS (Type of Transfer)
10. Bed, Chair, Wheelchair
11. Toilet
12. Tub or Shower
13. Car Transfer*

LOCOMOTION
14. Walking/Wheelchair (circle)
15. Stairs
16. Community Access*

COMMUNICATION ITEMS
17. Comprehension-Audio/Visual (circle)
18. Expression-Verbal, Non-Verbal (circle)
19. Reading*
20. Writing*
21. Speech Intelligibility*

PSYCHOSOCIAL ADJUSTMENT
22. Social Interaction
23. Emotional Status*
24. Adjustment to Limitations*
25. Employability*

COGNITIVE FUNCTION
26. Problem Solving
27. Memory
28. Orientation*
29. Attention*
30. Safety Judgement*

*FAM items

<table>
<thead>
<tr>
<th>Adm</th>
<th>Date</th>
<th>D/C</th>
<th>Date</th>
<th>Adm</th>
<th>Date</th>
<th>D/C</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Granger et al. (243) studied the FIM scores for stroke patients before and after implementation of the IRF-PAI. They found evidence of differences in the ratings in four areas, which they attributed to coding changes designed to influence case mix and reimbursement. Accordingly, they urge caution in data comparisons involving years before and after implementation of IRF prospective payment guidelines.

Several studies have shown that the FIM is a reliable instrument. Ottenbacher et al. reported a meta-analysis of 11 reliability studies in 1996 (244). These 11 investigations...
included a total of 1,568 patients and produced 221 reliability coefficients. The analysis demonstrated a median interrater reliability for the total FIM of 0.95 and median test-retest and equivalence reliability values of 0.95 and 0.92, respectively. The median reliability values for the six FIM subscales ranged from 0.95 for self-care to 0.78 for social cognition. For the individual FIM items, median reliability values varied from 0.90 for toilet transfer to 0.61 for comprehension. The reliability of motor subscale items was generally higher than that for items in other subscales. Additional reliability studies of the FIM have also shown similar findings.

A large number of validity studies have been reported since the inception of the FIM that have demonstrated that the scale has concurrent (245–252), predictive (253–256), and construct validity (239,248,257–261). These results have also been demonstrated with foreign (translated) versions of the scale (200,262–268). Studies have also been conducted comparing various modes of test administration, including interviews, telephone reports, and direct observations (269–271). These suggest that the FIM retains acceptable reliability under different conditions of administration.

Some limitations regarding scalability and sensitivity in the FIM have been noted. Hall et al. reported ceiling effects of the FIM at rehabilitation discharge, and particularly at 1 year after injury in a moderate to severely injured TBI population (272). Muecke et al. (256) studied the FIM as a predictor of rehabilitation outcome in lower-limb amputees. The FIM did not predict outcome of patients who were functioning lower (in bottom quartile) at admission, but it did predict rehabilitation success well in patients functioning at a higher level at admission.

To address the issue of sensitivity for use with brain injury survivors, an adjunct to the FIM, the FAM was developed by clinicians representing each of the disciplines in an inpatient rehabilitation program. Twelve items were developed in the areas of cognitive, behavioral, communication, and community functioning (273). These items are not designed to stand alone but are added to the 18 items of the FIM. The total 30-item scale combination is referred to as the FIM+FAM and requires approximately 35 minutes to administer. In 1995 a users group in the United Kingdom developed a special version of the FAM, keeping the seven-level structure, but attempting to improve the objectivity of scoring for ten items viewed as more subjective in scoring (274,275). According to Hall and Johnston, the FAM appears to add sensitivity beyond the FIM only for post-acute rehabilitation functional assessment (276,277). Although the scale items have shown acceptable reliability (230,278), a 2-year study by Gurka et al. showed that the addition of the FAM items only modestly increased the ability to predict employability and community integration in survivors of brain injury (279).

Another adaptation of the FIM, the self-reported functional measure, has been developed to quantify the ability of patients to care for themselves when they enter rehabilitation treatment and to chart their progress until they are discharged into the community or to another facility. Recent studies of the self-reported functional measure show that the instrument predicts inpatient hospitalization but not outpatient health care use (280) and that it can also predict caregiver hours (281).

**Canadian Occupational Performance Measure**

Although the Canadian Occupational Performance Measure (COPM) is not exclusively a measure of self-care, it measures self-care performance goals as identified by clients and is included as an example of a criterion measure that focuses on intervention goals viewed as important to the client (282–287). The COPM incorporates roles and role expectations from within the client’s living environment using a semi-structured, individualized interview.

The COPM uses a client-centered approach to measure perception of performance over time (285,288). The therapist and client collaborate within the administration of the measure to identify goal areas, plan intervention, and provide a baseline for re-assessment. Three subsections of the interview encompass self-care, productivity, and leisure. Scores yield measures of perceived performance and satisfaction in each of the three areas. The COPM may also be used to measure specific ability limitations in order to gain an understanding of why the client may be having difficulty in a particular functional area.

The COPM measures the client’s identified problem areas in daily functioning. In those instances where a client cannot identify problem areas (e.g., a young child, an individual with dementia), a caregiver may respond to the measure. The instrument takes into account client roles and role expectations and, in focusing on the client’s own environments and priorities, ensures the relevance of identified areas in the assessment process. The COPM can be used to measure client outcomes with different objectives for treatment, whether those aims involve restoration of function, maintenance, or prevention of future disability. The COPM may be used across cultures (289,290), diagnostic categories (289,291–293), and age groups (146–148). Its use has even been expanded to include systems, communities, and external stakeholders (288).

The COPM is administered in a five-step process using a semi-structured interview conducted by the therapist together with the client and/or caregiver. The five steps in the process include problem identification/definition, initial assessment, occupational therapy intervention, reassessment, and calculation of change scores. The original version included a procedure whereby rated importance was used as a weighting factor in calculating the performance and satisfaction scores. However, this has been eliminated based on findings from pilot studies that indicated the equivalence of scores whether or not importance weights are included.

In administration of the measure, problems are identified and defined jointly with the client and appropriate caregivers. Once the problem areas are defined, the client is asked to rate the importance of each activity on a scale of 1 to 10. The client (or caregiver) is also required to rate his or her ability to perform the specified activities and his or her satisfaction with performance on the same scale of 1 to 10. Following the initial portion of the interview, the client is asked to identify the
five top problems on a ten-point rating scale and indicate both perception of performance and perceived satisfaction. This information facilitates goal prioritization in preparation for treatment. These scores are then compared across time. There are two scores: one for performance and one for satisfaction. Administration time takes 30 to 40 minutes on average.

During initial development, the authors reported findings on an extensive pilot study of the COPM that involved administration in several countries, including New Zealand, Greece, and Great Britain (294). The scale has since been translated and used in several other countries. Early findings indicated that the average change scores for performance and satisfaction were approximately 1.5 times the standard deviation of the scores, indicating sensitivity of the instrument to perceived changes in occupational performance by clients. The COPM is viewed as a flexible instrument and appeals to clinicians who value the client-centered philosophy underlying its development (295). Some reports have indicated that patients occasionally experience difficulty with the process of self-rating of performance (296), and the suitability of the measure for use with patients demonstrating cognitive or affective difficulties has been questioned, although some studies have shown these concerns to be overstated (297). Additional concerns have been raised regarding the length of time it takes to administer the measure (288).

Reliability studies have reported intraclass correlation coefficients of \( r = 0.90 \) and above for both performance and satisfaction scores (298). Trombly et al. studied the achievement of goals by adults with traumatic brain injury and found that improved scores, as measured by the COPM, accompanied client perceptions of progress on scales of independent living and social participation (299). A comparative study of rehabilitation settings for survivors of stroke using the COPM showed that participant satisfaction with goal achievement was independent of setting and consistent with the results of performance measured by IADL and health outcome scales (300). A study by Simmons et al. found that the use of the COPM, in combination with the FIM, enhances accuracy in prediction of outcomes for rehabilitative services for persons in adult physical disabilities settings (146). More recently, the reliability and validity of the COPM with parents and pediatric populations has been studied (147,148). Cusick et al. (147) studied the internal reliability and validity of an adapted version of the COPM with young children. The study concluded the adapted version was psychometrically robust and appropriate to pediatric populations.

Additional recent studies continue to demonstrate positive results in the psychometric properties of the COPM (289,290, 293,301,302). Dedding et al. (302) studied the convergent and divergent validity of the COPM in two Amsterdam hospitals. The study demonstrated evidence of the instrument’s validity and the authors concluded the unique nature of the COPM in affording clients the opportunity to self-rate any self-care, productivity, or leisure activity. These and several other studies have demonstrated that the COPM correlates well with other measures of ADL outcome, motivates active participation and adherence to rehabilitation regimens, and improves satisfaction with services for a variety of diagnostic groups and ages (61,301,303–307). The COPM appears to provide useful information regarding self-care performance from the standpoint of the recipient of care.

Self-care Evaluation of Children

Self-care assessment of the pediatric client requires a number of special considerations. These pertain to incorporating developmental milestones into the structure of the assessment, interacting with the child during the assessment process, and when appropriate, obtaining a parental view of the child’s performance.

Within pediatric self-care assessment, practitioners consider the child’s physical, cognitive, psychological, and social emotional skills along with the environmental impact on daily function. Personal, social, and parental factors influence a child’s self-care performance. A child brought up in a home where value is placed on grooming, may develop those skills in advance of a child whose home environment places emphasis on academics or sport. Consideration of the interaction of the child, environment, and self-care activity is integral; holistic evaluation should include both direct evaluation of the child and assessment of the environmental supports, and barriers to occupational performance. Two widely used instruments that have attained sufficient use to establish their value in the clinical setting as useful measures of functional performance in children are the WeeFIM and the Pediatric Evaluation of Disability Inventory.

**FIM for Children (WeeFIM)**

In 1987, the FIM was adapted to meet the need for a reliable and valid functional assessment tool that would be useful in measuring the severity of disability in children. The resulting Functional Independence Measure for Children (WeeFIM) (308) was designed to measure the functional ability in a developmental context. In 2005, the WeeFIM II System Clinical Guide was developed and added a 0 to 3 module in addition to an Internet-based software program to the existing WeeFIM system (309). The WeeFIM II is designed for children ages 3 to 7 (or those falling below a developmental age of seven) and includes a minimal data set of 18 items that measure performance in the areas of self-care, mobility, and cognition. The scale uses the same seven-point ordinal scale to assess level of function (graded from dependence to independence) as the parent tool, the FIM. The WeeFIM 0 to 3 module is designed for children ages 0 to 3 and measures developmental performance in motor, cognitive, and behavioral performance utilizing a three-level rating system (3-usually, 2-sometimes, 1-never). The 0 to 3 module is designed to measure early functional performance and serve as a measurement for functional outcomes over time.

In conjunction with the development of the WeeFIM II, clinicians may now subscribe to the WeeFIM II system.
through the UDSMR. Subscription to the system includes specific modules including inpatient and outpatient modules and the 0 to 3 assessment tool. In addition, subscribers also receive quarterly benchmark reports, clinical and technical support, software for reporting and submitting online data, and a variety of other services to facilitate use of the FIM. The software system may also be used to interface with data reporting for the Joint Commission on Accreditation of Health Care Organizations.

Although continued research is needed on the WeeFIM II and the use of the 0 to 3 module, there is sufficient data to warrant the use of the WeeFIM in a variety of pediatric populations. Studies of the WeeFIM have shown a strong correlation between the scale scores and age, with the subscale scores involving gross and fine motor skill demonstrating the highest correlations (310). Data showed that tasks on the WeeFIM demonstrate a developmental sequence, with an observed positive relationship between the complexity of tasks and the age at which children achieve independence in their performance (311). Repeated evaluations of the scale and comparisons of personal and telephone interview ratings have demonstrated that the scale has good stability and equivalence reliability (312). In different studies and under varying conditions, the intraclass correlation coefficients for the six subscales have ranged from 0.73 to 0.99, and test-retest reliability has been estimated at 0.98 for children with disabilities and 0.99 for able-bodied children (312). Total score intraclass correlation coefficient values have consistently been greater than 0.95 (236). Comparisons of personal assessment and telephone interview ratings have found the scale to be consistent for items, subscales, and total test scores under those varying conditions of administration (312).

Comparisons of WeeFIM scores with other developmental tests, including the Vineland Adaptive Behavior Scales and the Bartelle Developmental Inventory Screening Test, found subscale correlations of 0.42 to 0.92 and total score correlations of 0.72 to 0.94 (313). Additional research comparing the WeeFIM to two standardized language tests, the Symbolic Play Test and the Reynell Language Development Scale, demonstrated high correlations in a Hong Kong pediatric population, the authors concluded from their pilot study that the WeeFIM appears to be an easy to use functional assessment of language that children with developmental delays (314). Comparisons of the WeeFIM with the Pediatric Quality of Life Inventory appear to be fair to moderately correlated in areas of physical health, but significant differences were not found in areas of psychosocial function, suggesting the tools may measure different constructs within the psychosocial area (315).

Several studies have been conducted on the validity of the WeeFIM and have used the instrument to measure the developmental and functional status of children with and without disability and across cultures (316–318). The scale has also been used to assess the functional status and rehabilitative progress in those with genetic impairments (240), heart defects (241), cerebral palsy (319–322), spina bifida (323), Down’s syndrome (324), and Rett syndrome (325). It is also useful in documenting progress and development following extreme preterm (246) and very-low-birth-weight delivery (247,248), primary brain tumors (249), and dorsal rhizotomy (250), in addition to rehabilitation following traumatic brain injury (251,252,326) and pediatric spinal injury (253). The WeeFIM has also been used to measure rehabilitation outcomes in larger populations across diagnostic, age, and rehabilitation cohorts (327). In summary, the WeeFIM is a useful instrument for measuring self-care performance in children. Additional research is warranted to continue to validate psychometric properties in the newer version, including the 0 to 3 module and the WeeFIM’s application to other cultures.

**Pediatric Evaluation of Development Inventory**

The Pediatric Evaluation of Development Inventory (PEDI) is described as a comprehensive assessment that samples key functional capabilities and performance in children from the ages of 6 months to 7.5 years. The scale may also be used for older children if their functional abilities fall below that expected of 7.5 year old children without disabilities (328). The scale is designed to be used with children having a variety of disabling conditions and can be administered by professionals or by structured interview, or parental report. The PEDI addresses both capability and performance in the areas of self-care, mobility, and social function. These three domains are further divided into subunits that comprise each task. Capability is determined by identifying the functional skills for which a child has demonstrated mastery, with scores reflected on the Functional Skills Scales. Two other subscales are provided. One is the Caregiver Assistance Scale, which measures the extent of help provided to the child during typical daily situations, and the other is the Modifications Scale, a measure of environmental modifications and equipment used routinely in daily activities. The PEDI has 197 functional skill items and 20 items that measure caregiver assistance and environmental modifications. The PEDI was designed to determine functional capabilities and deficits, to monitor progress, and to evaluate therapeutic outcomes. Table 9-8 lists item domains and complex activities included in the PEDI.

During the development of the PEDI, content validity was determined through use of a multidisciplinary panel of experienced experts (329). Items were derived from a wide array of functional performance and development scales. Normative data were collected from 412 children and families from the northeastern United States with a sample stratified to represent national population demographics while retaining equal representation across the target age groups using 6-month intervals. A detailed manual with scale development data, administration instructions, and scoring has been developed. The scale also has published scoring forms and software.

Six domain scores are provided that enable a profile of relative strengths and weaknesses in both functional skills and
caregiver assistance across the domains tested. No composite summary score is provided; with the rationale that this would obscure meaningful differences in functional performance within specific domains. Scaled scores can be computed to provide an indication of where a child’s performance falls relative to the possible maximum. Item difficulty for the PEDI was determined through Rasch analysis, which was also used to estimate goodness of fit between individual subject profiles and the overall hierarchy, intended for each subscale. Since each scale is self-contained, it can be used individually or in combination with other scales. The average time for administration is 45 to 60 minutes.

The psychometric properties of the scale are reported in the administration manual (328). Reliability data (internal consistency) for the six scale scores were computed using Cronbach’s alpha, with coefficients ranging from 0.95 to 0.99. Using the clinical samples, ICC values for inter-interviewer reliability for the scales was estimated at 0.84 to 1.00. Values for independent respondents ranged from 0.74 to 0.96. Selected modifications in the PEDI were made based on these data. Initial scale data reflected an expected progression of functional skills according to age. Initial concurrent validity was established through comparison of scores on the PEDI with scores on the Battelle Developmental Inventory Screening Test and the WeeFIM. These correlations were generally high for self-care and mobility but lower for social function. In early studies of the scale’s ability to detect change, results were mixed, with one clinical sample of children with mild to moderate traumatic injuries demonstrating positive changes on the PEDI in all domains. Another clinical sample involving children with multiple significant disabilities showed positive change after 8 months only on the mobility scale. Some scores for this group decreased, indicating that the children were falling behind their peers in age-expected functional levels (328). Ludlow and Haley studied the influence of setting (context) on rating of mobility activities and found that parents in the home setting tend to use stricter criteria in their ratings than rehabilitation professionals in the school setting, although both can be trained to attain a satisfactory level of consistency (330).

Since its initial development and normative studies, several clinical studies using the PEDI have been reported. These have related to measuring the status of very-low-birth-weight children at age 5 and for various rehabilitation or surgical interventions with children in various diagnostic categories, including traumatic brain injury (331–334), spinal bifida (335), cerebral palsy (336–338), and osteogenesis imperfecta (339,340). In addition, several reports have been published where the PEDI has been used to measure outcomes following targeted medical and surgical interventions for cerebral palsy (341,342), including studies of posterior rhizotomy (343–345), the use of botulinum-A toxin injections (346,347),

### Table 9.8: Content of the Pediatric Evaluation of Disability Inventory

<table>
<thead>
<tr>
<th>Self-Care Domain</th>
<th>Mobility Domain</th>
<th>Social Function Domain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Function skills scale</td>
<td>Types of food textures</td>
<td>Toilet transfers</td>
</tr>
<tr>
<td>Use of utensils</td>
<td>Chair/wheelchair transfers</td>
<td>Chair/toilet transfers</td>
</tr>
<tr>
<td>Use of drinking containers</td>
<td>Car transfers</td>
<td>Car transfers</td>
</tr>
<tr>
<td>Toothbrushing</td>
<td>Bed mobility/transfers</td>
<td>Bed mobility/transfers</td>
</tr>
<tr>
<td>Hair brushing</td>
<td>Tub transfers</td>
<td>Tub transfers</td>
</tr>
<tr>
<td>Nose care</td>
<td>Method of indoor locomotion</td>
<td>Method of indoor locomotion</td>
</tr>
<tr>
<td>Handwashing</td>
<td>Distance/speed indoors</td>
<td>Distance/speed indoors</td>
</tr>
<tr>
<td>Washing body and face</td>
<td>Pulls/carries objects</td>
<td>Pulls/carries objects</td>
</tr>
<tr>
<td>Pullover/front-opening garments</td>
<td>Method of outdoor locomotion</td>
<td>Method of outdoor locomotion</td>
</tr>
<tr>
<td>Fasteners</td>
<td>Distance/speed outdoors</td>
<td>Distance/speed outdoors</td>
</tr>
<tr>
<td>Pants</td>
<td>Outdoor surfaces</td>
<td>Outdoor surfaces</td>
</tr>
<tr>
<td>Shoes/socks</td>
<td>Upstairs</td>
<td>Upstairs</td>
</tr>
<tr>
<td>Toileting tasks</td>
<td>Downstairs</td>
<td>Downstairs</td>
</tr>
</tbody>
</table>

Complex activities assessed with caregiver assistance and modifications scale

<table>
<thead>
<tr>
<th>Eating</th>
<th>Grooming</th>
<th>Bathing</th>
<th>Dressing upper body</th>
<th>Dressing lower body</th>
<th>Toileting</th>
<th>Bladder management</th>
<th>Bowel management</th>
<th>Chair/toilet transfers</th>
<th>Car transfers</th>
<th>Bed mobility/transfers</th>
<th>Tub transfers</th>
<th>Indoor locomotion</th>
<th>Outdoor locomotion</th>
<th>Stairs</th>
<th>Functional comprehension</th>
<th>Functional expression</th>
<th>Joint problem solving</th>
<th>Peer play</th>
<th>Safety</th>
</tr>
</thead>
</table>

Management of bladder: Stairs

Management of bowel: Safety
and surgical release (348). Ketelaar et al. studied the properties of 17 scales assessing the functional motor abilities of children with cerebral palsy and concluded that the PEDI was one of only two measures that demonstrated acceptable psychometric properties while having the capability to document changes in function over time (349).

Studies of children from outside the United States have been reported to ascertain the suitability of using the PEDI with other cultural groups. Custers et al. compared profiles of non-disabled Dutch children with normative profiles and found enough differences to recommend the cross-cultural validation of the PEDI before use in the Netherlands (350). More recently, the reliability of the Dutch version was found to have positive psychometric properties, with the inter-interview reliability exceeding that of test-retest and inter-respondent reliability (351). The authors concluded that the psychometric properties of the Dutch version were “good” but that minor adaptations needed to be made (351). Similar recommendations were made after a study of children in Puerto Rico (352). These studies led to recommended item modifications for use in those countries (353). A study of the applicability of the PEDI in Slovenia found statistically significant differences in functional skills and caregiver assistance scores between Slovene children and the American normative group (354). The authors concluded that the American normative data are not fully appropriate for reference with the Slovene population (354). In Norway, a study of the applicability of the PEDI utilized a previous translation of the PEDI utilizing international guidelines (355) and found the Norwegian sample scored significantly lower for functional skills in mobility tasks than the American children. Other areas were similar between the two groups yet the authors recommended further research from which Norwegian reference values can be derived (355). In summary, the PEDI has wide range applicability across diagnostic groups. Although it is widely used internationally, further research is needed to explore its reliability and validity in various cultures. (Readers should note that the PEDI is now under revision and should consult the developers for latest version.)

**MANAGEMENT OF BASIC SELF-CARE SKILLS**

The functional evaluation process characterizes the strengths and weaknesses of each client in relation to specific daily living skills. The practitioner or rehabilitation team uses this information to develop options for a plan of care that will assist or enable the client to become more independent in selected activities. Research has shown that patients who collaborate in identifying goals and selecting treatment options become active and responsible participants in their own rehabilitation. Most intervention approaches fall into one of four categories: remediation, compensation, disability prevention, or health promotion (356). These categories are consistent with what people do when they cope with functional limitations without professional advice.

**FIGURE 9-3.** Self-care goals and intervention strategies.

<table>
<thead>
<tr>
<th>Evaluation Data</th>
<th>Problem List</th>
<th>ADL Goals</th>
<th>Intervention Strategies</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>HISTORY/NARRATIVE</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Client is male, 67 y.o., retired school administrator. Active premorbidly in volunteer work and outdoor hobbies.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>MEDICAL DIAGNOSES</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Right CVA Depression</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SENSORIMOTOR</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proprioception intact Grade 3(Fair) muscle strength in Left wrist, elbow, and shoulder flexors and extensors</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No visual field cut 2 point discrimination absent on R digits</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>PERFORMANCE LIMITATIONS</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unable to initiate or complete dressing or bathing tasks without verbal cues and physical assistance.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>COGNITIVE</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Follows simple directions</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>FUNCTIONAL LIMITATIONS</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left Hemiparesis Apathy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>AFFECTIVE</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No interest in activities</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**REMEDIATION**

Techniques to encourage use of affected extremity Improve strength, coordination

**COMPENSATION**

Train in use of dressing skills Introduce dressing aids (stocking aid, zipper pull, velcro closures) Train family members

**DISABILITY PREVENTION**

Grab bars in tub/shower Non-slip surfaces
AN OVERVIEW OF INTERVENTION APPROACHES

The major categories of ADL intervention described earlier are briefly described in the following sections. Determination of the most appropriate treatment approach is, ideally, collaboration among the client, caregiver, and health-care team. The therapeutic approach that is the best match intervention considers the clients’ needs, desires, and expectations as well as capacity for learning, the prognosis for the impairments, the time available for intervention, and the anticipated discharge environment.

Remediation
Remediation is the re-acquisition of functional or structural deficits, or the acquisition of new skills in the area of skilled movement, cognition, or social function. In this approach, one expects that lost skill or ability will be regained. Rehabilitation approaches are used to establish or reestablish the client's skills, habits, and routines. Consequently, the use of this approach implies an active learning process whereby patients must adapt to functional and environmental limitations as they affect the demands of everyday life. The initial acquisition of skills by those who are developmentally or congenitally disabled at birth is a markedly different learning process from the reacquisition of daily living skills by those who have been independent at such tasks before becoming disabled. When the goals of remediation are to develop skills in a person with a congenital condition, the training process is described as habilitative. When the goal is to achieve previous functional levels for a person with an acquired disability, the training process is described as restorative.

The best candidates for remediation or restoration of skills often involve a client who may make considerable contributions to the therapy process based on his or her previous knowledge and understanding of how the task was performed. Such patients can monitor their own errors and often use appropriate strategies to minimize deficits. In this case the treatment session is used to develop a practice strategy and the client practices self-care at each opportunity whether or not the therapist is present. Problems occurring between treatments are discussed, and possible solutions can be practiced during the next treatment session. Implied in the process of remediation is the ability to learn. Learning strategies and implementing teaching strategies to optimize learning are important in this process.

Learning Strategies
The first key to establishing a positive learning environment for a client is to ensure that the set goals are relevant and meaningful to the client. Meaning refers to the sense that is made or the personal implications that are drawn from, engaging in self-care tasks. Nelson et al. (357) define meaning as an interpretive process in which increasing the purpose of a task can improve the meaningfulness of a given activity. For example, embedding an exercise within an activity of daily living (or occupation) added purpose to the exercise and thereby increased the likelihood of accomplishment of the task. Involving the client and family in the goal-setting stage can also help engineer a match between learning and capabilities by improving meaning and therefore motivation to accomplish the task.

As research has unveiled the importance of specific learning strategies in both acquisition and re-acquisition of functional skills, rehabilitation professionals have incorporated these strategies into training. There are a variety of classification schemes for describing these learning strategies. In practice, four stages of learning described by Bertoti are helpful in differentiating between the initial stages of learning (e.g., acquisition) and the long-term stages of learning that may include maintenance, fluency or proficiency, and generalization. This incremental learning process, moving from initial instruction to mastery with graded assistance, is sometimes termed scaffolding (358).

Optimal learning in the initial stages requires different strategies from those for retention or long-term learning, but most of the teaching strategies are used at both the initial and long-term learning strategies, albeit in different ways. Some of these strategies for learning are highlighted in Table 9-9. Research also indicates that learning is enhanced when the learner engages in a naturally occurring task versus use of progressive resistive exercise and by using real objects in a natural context (357,359–363).

The strategies employed by the therapist in each stage includes anticipation of the needed instruction, application of physical or verbal guidance, timing of extrinsic feedback, emphasis of intrinsic feedback, and the orchestration of practice (358,364).

Instruction
Instruction prior to initiation of any task is crucial for motivation and for clarification about the task and can take the forms of verbal, demonstration, and modeling. Verbal instructions help to focus the client’s attention on important aspects of the task. The therapist may choose to break down verbal instruction into one or two essential elements. Verbal instructions that are both brief and clear are the most helpful to clients. In addition, a therapist may use demonstration or modeling of the task to either replace or enhance the verbal instructions. Demonstrating how the client is to achieve the task has clearly been demonstrated to promote learning. Effective teaching strategies require a thorough task analysis to identify barriers to learning and to determine prior knowledge and current capabilities of the learner. In addition, effective training includes the use of active learning strategies, engagement in motivational and collaborative activities, adaptation to the learners’ abilities, and use of real objects (362,363). The teaching and implementation processes are highlighted in Table 9-10.
Guidance
Guidance, also referred to as facilitation, assists the client with execution of an action. Guidance has been shown to be effective in the acquisition phase of learning novel tasks, particularly when it is graded and is then extinguished to match the client’s skill development (358). The implementation of guidance is most effectively used when the following circumstances exist: when the client is in the initial learning stage, is performing tasks that are naturally completed at a slower rate, or is unfamiliar with the task. It is also used for prevention of injury or reduction of fearfulfulness during movement. The overall goal is to reduce the use of guidance while promoting independent practice.

Feedback
Feedback is critical to the rehabilitation process and can greatly enhance the success in performing ADLs most efficiently. However, care must be taken to select the appropriate type and quantity of feedback based on the needs of the client. There are several types of feedback that incorporate both intrinsic (i.e., internal sensory cues) and extrinsic (e.g., verbal, tactile, knowledge of results) information. These are highlighted in Table 9-11.

<table>
<thead>
<tr>
<th>TABLE 9.9</th>
<th>Strategies Used During Different Stages of Learning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strategy</td>
<td>Acquisition Stage</td>
</tr>
<tr>
<td>Extrinsic Feedback</td>
<td>Use of multiple types of cues (verbal, tactile, videography, and biofeedback), given both frequently and concurrently with the task performance</td>
</tr>
<tr>
<td>Practice</td>
<td>Encourage both physical and mental/imagery practice</td>
</tr>
<tr>
<td></td>
<td>Provide opportunities for the task repetition</td>
</tr>
<tr>
<td></td>
<td>Consistency</td>
</tr>
<tr>
<td></td>
<td>Use part practice to master components</td>
</tr>
<tr>
<td></td>
<td>Blocked practice</td>
</tr>
<tr>
<td></td>
<td>Correct only significant errors</td>
</tr>
<tr>
<td></td>
<td>Give opportunities to fixate on closed skills</td>
</tr>
<tr>
<td></td>
<td>Require cognitive activity and problem solving</td>
</tr>
<tr>
<td>Guidance</td>
<td>Manual contact in conjunction with verbal cues</td>
</tr>
<tr>
<td></td>
<td>Trial and error may promote frustration</td>
</tr>
<tr>
<td></td>
<td>so limit the variation of practice to promote a balance toward success of task completion</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TABLE 9.10</th>
<th>General Procedures for Implementing Teaching and Learning Strategies during Skill Acquisition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identify and address barriers to learning the activity or occupation.</td>
<td></td>
</tr>
<tr>
<td>Determine the learner’s prerequisite skill set and learning needs.</td>
<td></td>
</tr>
<tr>
<td>Engage the client and caregivers in joint goal setting then employ active learning strategies, observation, and demonstration.</td>
<td></td>
</tr>
<tr>
<td>Provide a nurturing, safe, and supportive environment.</td>
<td></td>
</tr>
<tr>
<td>Provide activities that pose an adequate challenge to the client, using grading and adaptation to create this appropriate challenge.</td>
<td></td>
</tr>
<tr>
<td>Provide opportunities for practice in context, using the appropriate practice schedule for skill level.</td>
<td></td>
</tr>
<tr>
<td>Facilitate use of feedback (both internal and external) and physical guidance during skill acquisition.</td>
<td></td>
</tr>
<tr>
<td>Transfer learning by engaging client in active problem solving, use errors as a positive learning opportunity.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TABLE 9.11</th>
<th>Types of Feedback Used in Intervention</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type of Feedback</td>
<td>Intrinsic</td>
</tr>
<tr>
<td>Proprioceptive</td>
<td>Knowledge of results</td>
</tr>
<tr>
<td>Tactile</td>
<td>Common clinical mechanisms</td>
</tr>
<tr>
<td>Vestibular</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Visual</td>
<td></td>
</tr>
</tbody>
</table>
Practice

Practice is an important aspect of all motor training. Mere repetition of activity is not therapeutic training. The therapist or other self-care trainer must provide task structuring, strategic prompts, and suggestions for improvement of performance. The client or learner must learn to monitor and correct performance errors. Over time, desirable behaviors must be systematically rewarded and undesirable behaviors ignored or extinguished.

Practice usually occurs only at the time or times each day that a task is appropriate. During the initial or acquisition stages of instruction, therapists can promote generalization and mastery by teaching under natural conditions, using real equipment rather than simulated tools, involving multiple teaching conditions (locations, instructors, materials), and selecting the instructional examples carefully, with attention to those that best sample the range of variation likely to be encountered in task performance. Learners are encouraged to do what they can to help—sometimes starting a task that the therapist will need to finish and sometimes completing a task that only the therapist can initiate.

For the client with an acquired disability, relearning self-care independence is a distinctly different process than for the habilitation client. First, there is a loss of self-esteem and sense of failure and frustration when one is unable to perform those tasks that often are taken for granted by non-disabled persons. Initial learning usually is motivated by intrinsic rewards of increased competency at self-care tasks and by the positive social reinforcement of parents and other caregivers. In learning a task, negative reinforcement (i.e., avoiding unpleasant experiences or consequences) such as avoiding embarrassment over having to ask for assistance for feeding or toileting may be far more effective than positive social reinforcement. The therapist who tries to use social praise to reinforce practice of toileting skills will find that it is not effective. In fact, it probably will be viewed as demeaning to praise an adult in a situation reminiscent of a childhood experience.

Although many physically disabled adults with acquired deficits may not fit the above description, those who also have significant cognitive or perceptual deficits may not be appropriate. Conditions such as apraxia may significantly influence the effects of self-care training and may need to be remediated before ADL training. Recent studies, while inconclusive, suggest that such concurrent treatment may be useful (365). For such clients and those for whom other approaches have failed to produce results, compensatory training and modification should be considered.

Other remediation strategies focus on the restoration of biological, physiologic, or neurologic processes. For example, the practitioner may incorporate motor or sensory techniques to fully or partially develop or restore sufficient voluntary control or movement to enable task accomplishment. Techniques derived from theories of neuroscience, biomechanics, and motor control are included among remediative approaches. The objective is to recover sufficient perception, cognition, and voluntary movement to enable task performance in a safe and effective manner. It is noted, however, that studies have shown that body structure and function predict less than half of the variation in task performance. In particular, the literature on stroke rehabilitation shows that gains in organ/physiologic (impairment) skills are small and do not automatically result in improved functional performance. Trombly analyzed the results of several studies in which correlations between motor impairment and ADL were reported. Her findings indicate that the amount of variance in ADL accounted for by motor impairment was 31%. Most (approximately 69%) of the variance associated with ADL performance was derived from other factors unrelated to the physical impairment. Motivational and environmental factors explain most task performance variations (359–361).

In the restorative approach, training is frequently combined with compensatory strategies, such as the use of ATDs, prostheses, or orthoses. For example, the individual with only one functional arm can continue to be independent in dressing by learning to select certain clothing and/or by substituting motions and using devices for dressing. When remediative approaches are inadequate or are too costly in terms of time, energy, or expense, compensatory or adaptive approaches are necessary.

When able-bodied children learn self-care tasks initially, they do so over extended periods of time. Anthropologists note that there is a regular sequence of self-care independence that is supported by child development studies across cultures. Feeding, grooming, continence, transfers, undressing, dressing, and bathing usually occur in this order, with normal acquisition influenced by the appearance of readiness skills.

Even such limited information as this may be useful in habilitation training. Using a distributed practice schedule (i.e., teaching self-care activities only during those times they would normally be performed) is critical for the person unfamiliar with the concept of the task. Effectiveness in self-care training involves learning the appropriate times and natural sequences of daily activities. The client who is relearning a task often retains an appreciation of when it is to be performed, but the client being habilitated needs to learn not only the skills but also the context appropriate to each task. Thus, in habilitation training, the acquisition of self-care skills may occur over extended periods and involves the careful structuring of tasks, frequent monitoring to correct performance errors, and feedback.

**MAJOR SELF-CARE CONSIDERATIONS BY FUNCTIONAL LIMITATION**

This section concerns special problems and issues related to ADL performance limitations associated with specific functional deficits. Perhaps the most important concept in habilitation aimed at daily living skills is that neither the type nor the severity of the disability can be used exclusively to predict how independent a given client will be. For example, Putzke et al. determined the predictors of ADL function...
among cardiac transplant candidates and found no correlation between certain expected physical variables and IADL capacity (366). Limited capacity for exertion, which is used by many occupational and physical therapists to recommend appropriate activities, was not a determinant of the activities those cardiac clients performed as part of their daily routine. The researchers concluded that client perceptions and not cardiac symptoms accounted for the patterns of independence/dependence in household and social activities. A study of self-care and aging by Norburn et al. (367) identified three types of self-care coping strategies. These were related to the use of equipment or devices, specific changes in behavior, and modifications of the environment. The study also found that receiving assistance sometimes was used to supplement self-care coping strategies.

Is it useful, then, to discuss self-care considerations by functional limitations? If the goal is to become aware of special problems or issues unique to a particular disability and not to predict outcomes, then the answer clearly is affirmative.

Some studies of physical function have provided useful theoretical associations between diseases and reports of difficulty with clusters of tasks within IADL routines. In a useful statistical analysis, Fried et al. studied 5,201 men and women age 65 and older in four U.S. communities (368). They analyzed difficulties associated with 17 tasks of daily life and found four clusters of activities where difficulty with one task was associated with reported difficulty with other tasks in that group. The groups included activities primarily dependent on mobility and exercise tolerance, complex activities heavily dependent on cognition and sensory input, selected basic self-care activities, and upper-extremity activities. These groups were then examined to see if they were associated with underlying impairments based on regression analysis using 15 chronic diseases and conditions. They found that physiologic and disease characteristics significantly correlated with difficulty in each of these four groups. The implications of this study pertain to the appropriateness of outcome measures, which the authors suggest should be chosen with defined physiologic rationale. However, the study also provides a general indication of patterns of ADL/IADL consequences associated with the 15 conditions (verbal, tactile, knowledge of results). These are highlighted in Table 9-12.

Several useful conclusions related to the performance of life tasks can be gleaned from this study. First, it is evident that problems with body structure and function seldom affect isolated tasks or task categories. Thus, some health functional difficulties influence performance across all task categories. These could include depression, joint pain and stiffness, visual impairments, shortness of breath, and both generalized

---

**TABLE 9.12** ADL/IADL Task Groups Found to Be Statistically Associated with Underlying Medical Conditions

<table>
<thead>
<tr>
<th>Group</th>
<th>Mobility</th>
<th>Complex (Cognitive/Sensory)</th>
<th>Self-care</th>
<th>Upper Extremity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mobility</td>
<td>Walk 1/2 mile</td>
<td>Pay bills</td>
<td>Using toilet</td>
<td>Gripping</td>
</tr>
<tr>
<td></td>
<td>Climb ten steps</td>
<td>Meal preparation</td>
<td>Dressing</td>
<td>Reaching</td>
</tr>
<tr>
<td></td>
<td>Transfer from bed to chair</td>
<td>Shopping</td>
<td>Bathing</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Walk in own home</td>
<td>Light housework</td>
<td>Eating</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Heavy housework</td>
<td>Telephone use</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Lifting or carrying ten pounds of groceries</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complex (Cognitive/Sensory)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Balance disorder</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Claudication</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Joint pain/stiffness</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Shortness of breath</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>General weakness</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hand grip weakness</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Depression</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cognition/memory impairment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Visual impairment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Impairment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Congestive heart failure</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cancer</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Arthritis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Stroke</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Diabetes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Emphysema</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hypertension</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disease association</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(difficulty with one or more tasks in group)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

and hand grip weakness. Although these associations can help practitioners to anticipate certain activity and participation limitations associated with given diseases and their accompanying functional pathology, each case must be considered in the context of a particular client's circumstances.

From the standpoint of rehabilitation intervention, these data suggest that some generalized strategies for improving endurance and strength can have the potential for broadly benefiting ADL/IADL performance on several tasks. In those cases where the underlying etiology of task difficulty cannot be addressed, such as with weakness and shortness of breath accompanying congestive heart disease, general compensatory strategies to address performance of self-care and IADL tasks will be necessary.

The following sections summarize recent studies related to interventions and outcomes for selected areas of task-related functional pathology. Practitioners are encouraged to consult the many intervention textbooks now available in the rehabilitation sciences for theoretical and applied (procedural) guidelines and details associated with implementing intervention strategies (369–371). Although many interventions have not yet been subjected to controlled study; it is important for practitioners to be aware of the literature in order to determine the available evidence underlying a particular intervention technique.

**Cognitive and Perceptual Deficits**

Diminished cognitive and perceptual function can be a consequence of senile dementia, traumatic brain injury, stroke, or developmental disability. The ability to attend; perceive and select relevant information in the environment; make logical decisions; plan and execute safe, goal-directed action; and remember events involves aspects of executive function, which may be compromised in cognitive dysfunction. Individuals with cognitive and/or perceptual loss must be approached and managed with greater care and attention than other clients.

The association between cognition (particularly executive function) (372) and ADL/IADL performance is well documented. Studies of community-dwelling seniors (373–375), persons following hip fractures (134,376,377), clients with cerebral palsy (378), and stroke clients (379–381) are among those documenting this association. However, Perneeczky et al. argue that the presence of cognitive impairment affecting complex ADL should not be used as a criterion for diagnosing dementia (382).

Sometimes cognitive loss is associated with agitation, particularly following traumatic brain injury or in senile dementia. For example, Bognar et al. found that the presence of agitation predicted longer length of stay and decreased functional independence following traumatic brain injury (383).

In stroke, generalized cognitive deficits are most common in clients with frontal lobe insults. Cognitive deficits involving visuospatial abilities are associated with right hemisphere strokes and apraxia. Disorders of learned movement or motor planning are found with left hemisphere strokes. Unilateral neglect, or inattention to the side opposite the affected hemisphere, is also a difficult consequence of stroke and can occur in up to half of clients with right brain damage and less frequently in left hemisphere strokes. Unawareness of the stroke or its results can be found in clients with strokes in the nondominant hemisphere. This appears to be related to the location and severity of brain damage, or the extent of cognitive impairment (384–386). Typically, these deficits interfere with regaining skills in ADLs. Clients with spatial constructional problems often have difficulty in eating, dressing, grooming, and walking. Those with unilateral neglect experience difficulty with reading, writing and numerical tasks, as well as in using a phone or driving.

Several studies have found that perceptual motor abilities are significant predictors of self-care and ADL in clients with CVA (387–390). In a systematic review, Rubio and Vandeven summarized the research on perceptual deficits and ADL-targeted intervention approaches following stroke (391). They acknowledged the convincing evidence, associating perceptual difficulties with poorer ADL outcomes and noted that little attention has been paid to agnosia and that controlled studies comparing various intervention options are needed.

The literature documents many attempts to use perceptual training in addition to task-oriented (functional) training as a way of improving ADL outcomes in clients with perceptual problems, such as unilateral neglect following CVA (392). Strategies for perceptual remediation include (a) visual scanning training with verbal cuing for visual anchoring, (b) increased client awareness of neglect to encourage use of compensatory mechanisms, (c) increased sensory stimulation to the affected side, and (d) activation of the affected extremity on the affected side of the body. Overall, although it has been shown that interventions targeted for perceptual skills can have short-term effects in improving perceptual skills (393), research has not demonstrated unequivocally that perceptual training in addition to traditional functional intervention improves outcomes in ADL over functional training alone. Clients with such limitations can and do make progress, albeit more slowly than similar clients without such deficits. At least one study has shown that training that includes specific compensation strategies along with traditional ADL strategies can improve overall performance in ADL skills for CVA clients with left hemisphere stroke (394). Unfortunately, there is a dearth of information on the long-term results of perceptual interventions.

Treatments for apraxia or motor planning, and execution deficits have included such strategies as manual guided movement, the use of objects and situations that elicit automatic motor responses, and systematic stepwise training (“backward chaining”). Controlled studies of the long-term efficacy of these approaches for improving functional ADL skills have not been reported.

**Depression**

Numerous rehabilitation studies have shown that the presence of depression following stroke and other health events (such as hip fracture, post-polio syndrome, traumatic brain injury) can
predict poorer functional outcomes and diminished well-being (395). Families and caregivers can be an important source of monitoring for depression, and mild depression often resolves or can be managed through behavioral interventions such as encouragement, attention, or active participation in suitable activities. More severe depression should be confirmed by accurate diagnosis. Symptoms may be effectively managed through medication (396). However, it appears for clients with CVA that the reduction in symptoms, particularly if started early, can result in improved levels of ADL function at discharge (397) but its presence reduces the proportion of excellent outcomes. A controlled study by Desrosiers et al. showed that a leisure education program could reduce depression and increase participation in leisure activities by stroke survivors (398).

Craig et al. reported on a program of cognitive behavior therapy designed to reduce depression and improve self-esteem during rehabilitation following spinal cord injury. Their study improved anxiety, mood, and self-esteem of those receiving the intervention but did not result in outcomes that were significantly different from those of clients receiving traditional rehabilitation. Interestingly, studies of suicide following spinal cord injury show that frequencies are higher among those with marginal functional disability than in those with complete lesions.

**Upper-Extremity Impairment**

Upper-extremity impairment can be bilateral, as in spinal cord injury and arthritis, or unilateral, as may be the case following stroke, traumatic injury, amputation, peripheral nerve injury, or complex regional pain syndrome (CPRS/reflex sympathetic dystrophy). Bilateral upper-extremity amputation, loss of motion due to hemiparesis of nerve injury, and weakness are serious threats to self-care independence and may be helped with prosthetic and orthotic devices. Unilateral upper-extremity impairment most often is compensated for by use of the uninvolved extremity. When disability affects the preferred hand, there will be a need to transfer the skill to the other hand. This may or may not require coordination and dexterity training.

Persons with unilateral upper-extremity amputations will often be able to accomplish most of their ADL with one arm. Thus, it is important to fit the unilateral upper-extremity amputation with a prosthesis within 3 months of losing the arm in order to facilitate training for bilateral patterns of use. The prosthesis always functions as the nonpreferred extremity in stabilizing objects (398).

Cosmesis may be a key factor for the client in the selection of a prosthesis, which can be body powered, myoelectric, or hybrid. Bilateral upper-extremity amputations, particularly if above the elbow, are more challenging but can result in successful independence. The prosthetic training for a bilateral above-elbow amputee usually requires a carefully designed series of training sessions. Initial training of BADL requires an extensive time, effort, and motivation on the part of the amputee. Often the basic tasks of writing, eating, brushing teeth, washing the face, and requiring assistance with dressing are the most realistic goals during rehabilitation following the prosthetic fitting. Later training can concentrate on additional activities in the daily routine (398).

**Lower-Extremity Impairment**

Mobility and transfer limitations are the most significant self-care problems for lower-extremity impairments, including those with hip fractures, amputations, arthritis, and paraplegia. Often wheelchairs or other ambulation aids are needed for independent mobility. Bathroom safety equipment, dressing aids such as extended handle shoehorns, and raised seats often are useful items that may be used either temporarily or permanently. Rearrangement of living space to permit wheelchair access or access using other mobility aids often is necessary. Ramps, chairlifts, and additional railings may be needed if stairs are present. For the person using a wheelchair, it also may be helpful to consider rearrangement of shelves, drawers, and closet space to permit frequently used items to be reached from the wheelchair. Studies have shown an association between the need for mobility devices following lower limb amputation and reduced ADL independence (229,399,400).

**Upper- and Lower-Extremity Impairment**

Persons with tetraparesis or tetraplegia from traumatic spinal cord injury, cerebral palsy, muscular dystrophy, multiple sclerosis, or amyotrophic lateral sclerosis must rely on a wide range of options for self-care independence. In most cases, there will be a need for attendant care, assistive devices, and modifications of living space. Such people often require high-technology devices, such as environmental controls, augmentative communications, and other microprocessor-based systems to be fully independent in their own homes (401). Promising developments are being shown in the technology associated with implanted neuroprostheses for restoring hand grasp and enabling ADL performance in tetraplegia following spinal cord injury (402,403).

**Pain**

The disabling consequences of low back pain or CRPS/reflex sympathetic dystrophy frequently include intractable pain, sensory changes, edema, and movement disorder, which can include restricted or involuntary movement and atrophy of the upper or lower extremity. Pain syndromes (404–407) can result in marked functional limitations affecting ADL performance, usually caused by pain rather than by restricted motion (408). Treatment often consists of medications or TENS to assist with pain control and compensatory training for alternative methods in accomplishing daily living tasks (408). In a recent systematic review, it was concluded that information based on a biopsychosocial model is a more efficient strategy to modify clients’ beliefs on pain and consequences, and increase the adherence to exercises. Thus, it seems necessary to focus as much attention on a client’s beliefs about chronic pain and gaining participation in regular exercise practice to prevent pain pathology (409).
Loss of Use of One Side of the Body
Perhaps the most important finding of studies related to ADL intervention for hemiparesis is that the side of the lesion may predict a pattern of problems but may not predict self-care outcomes. The cognitive and perceptual problems associated with different lesion sites and extents of damage were summarized earlier.

The differences between left- and right-side brain damage do not affect prognosis for self-care independence directly. Instead it appears that clients with left-sided brain damage are easier to train, receive more training, and progress more rapidly. Clients with right-sided damage tend to be more difficult to train, take longer to train, and tend to receive less training as a result. It remains possible that attitudes and skills of therapists and the economics of medical rehabilitation may be biased toward those with left-sided brain damage. Consequently, therapists and physicians need to be aware of the implications of their decisions about intensity and duration of training for clients with right- versus left-sided brain damage. It seems especially important to note that both groups show gains over time in self-care and that the hospital discharge prognosis for ADL independence is not necessarily the long-term self-care prognosis.

Historically, many therapeutic programs for stroke concentrated on the neurologic substrates of stroke, seeking to normalize muscle tone as a precursor to emphasis on functional training. More recent controlled studies have shown that motor learning strategies may represent a more effective approach in achieving improved outcomes in ADL (410,411).

An intervention approach known as constraint-induced movement therapy (CiMT; also forced-use therapy) has received increasing attention. This approach uses a regimen involving intensive training of the affected extremity during which the unaffected arm is constrained in a sling for much of each day. The intent is to foster cortical reorganization to recruit other areas of the brain to assume motor planning and control functions. Some encouraging results, including improvements in ADL function, have been demonstrated in clinical studies, but the approach is still viewed by some as experimental (412–416). A study used modified CiMT by adding a mental practice component and found encouraging results (417).

Limitations in Joint Range of Motion
Problems of limited range of motion that result in difficulty reaching common ADL items are best addressed by rearrangement of living space. An occupational therapist or other health care provider making a home visit can suggest safety equipment for the bathroom, removal of obstacles that could contribute to falls, and other adaptive devices that are needed. With respect to training, an emphasis on compensatory training regarding work simplification and energy conservation techniques is needed. Principles of joint protection and energy conservation are described in Tables 9-13 and 9-14.

<table>
<thead>
<tr>
<th>TABLE 9.13 Strategies for Joint Protection</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Strategies for Joint Protection</strong></td>
</tr>
<tr>
<td>Avoid positions that foster deformity</td>
</tr>
<tr>
<td>• Buildup handles on pens/keys</td>
</tr>
<tr>
<td>• Scissors/knife to open packages</td>
</tr>
<tr>
<td>• Electric appliances</td>
</tr>
<tr>
<td>• Faucet handles with levers</td>
</tr>
<tr>
<td>• Mugs, not cups, with small handles</td>
</tr>
<tr>
<td>Avoid tight gripping</td>
</tr>
<tr>
<td>• Use a sharp knife with a large handle in a sawing motion or adaptive scissors.</td>
</tr>
<tr>
<td>• Use buildup handles on cooking/mixing utensils, buckets, tools, shears, or briefcases.</td>
</tr>
<tr>
<td>• When sweeping/mopping keep handle at waist level in front of you with back straight to avoid poor posture.</td>
</tr>
<tr>
<td>• Use a flat hand to open jars or use assistive devices such as a Zim jar opener.</td>
</tr>
<tr>
<td>• Use lightweight utensils/cookware.</td>
</tr>
<tr>
<td>• Use cutting board with prongs.</td>
</tr>
<tr>
<td>• Use a flat hand to wash windows or countertops. Do not grip sponge tightly to wring water, flatten sponge against side of sink with palm to remove water.</td>
</tr>
<tr>
<td>Avoid placing excessive or constant pressure on joints</td>
</tr>
<tr>
<td>• When rising from a sitting position, use the heel of your hand to assist yourself.</td>
</tr>
<tr>
<td>• Avoid resting your chin on your knuckles.</td>
</tr>
<tr>
<td>• Avoid lifting heavy objects or repetitively gripping/lifting objects during work/ADL.</td>
</tr>
<tr>
<td>• Place heavy objects on wheels/slide them.</td>
</tr>
</tbody>
</table>

(Continued)
### TABLE 9.13 Strategies for Joint Protection (Continued)

<table>
<thead>
<tr>
<th>Avoid prolonged static positions</th>
</tr>
</thead>
<tbody>
<tr>
<td>• During activities such as typing, reading, desk work, and sewing, take frequent breaks to rest or move joints.</td>
</tr>
<tr>
<td>• When driving a car long distances, periodically release grip on steering wheel by resting/shaking hands one at a time.</td>
</tr>
<tr>
<td>• A foam steering wheel cover allows a less forceful grip.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Use proper body mechanics</th>
</tr>
</thead>
<tbody>
<tr>
<td>• When lifting objects from below waist level, bend knees and keep back straight.</td>
</tr>
<tr>
<td>• Do not lift heavy/bulky objects overhead. Use a stool to reach high places.</td>
</tr>
<tr>
<td>• When using a twisting motion, always turn hand toward the thumb (e.g., turning door knobs, keys in locks, jar lids, or wringing clothes).</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Use the largest, most stable joint for the task</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Lift an object with both hands using both forearms to carry close to body instead of gripping handles.</td>
</tr>
<tr>
<td>• Push objects with your weight instead of pulling them with your fingers.</td>
</tr>
<tr>
<td>• Carry purse on shoulder or forearm rather than in hand. Lighten purse/briefcase as much as possible.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Respect pain and fatigue</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Learn to distinguish the usual pain/fatigue that decreases after completing an activity from the pain that persists longer than 1 hr after the activity ends.</td>
</tr>
<tr>
<td>• Stop activity when signs of fatigue are noticeable.</td>
</tr>
<tr>
<td>• Eliminate activities resulting in persistent pain.</td>
</tr>
<tr>
<td>• Get enough rest: 10–12 hrs each night taking naps as necessary during the day. Rest has been shown to have a significant effect on how joints feel.</td>
</tr>
<tr>
<td>• Rest joints when they are flared-up or painful in deformity, preventing positions to prevent tightening and lost ROM.</td>
</tr>
<tr>
<td>• Learn placing techniques to accomplish what you want while keeping pain at a manageable level and to be able to do both of these things day after day.</td>
</tr>
<tr>
<td>• Use time, not pain, as guide to limit activity.</td>
</tr>
<tr>
<td>• Deformity-preventing joint resting positions in a firm bed are as follows:</td>
</tr>
<tr>
<td>• Chest expanded</td>
</tr>
<tr>
<td>• Small pillow used below neck/head</td>
</tr>
<tr>
<td>• Ankles neutral using a pillow</td>
</tr>
<tr>
<td>• Spine straight with hips extended</td>
</tr>
<tr>
<td>• Shoulders held back with arms straight</td>
</tr>
<tr>
<td>• Knees should be extended and straight.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Prioritize and organize activities</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Develop realistic expectations of what can be accomplished.</td>
</tr>
<tr>
<td>• Spend most of the time on enjoyable activities.</td>
</tr>
<tr>
<td>• Organize and plan for frequent rest.</td>
</tr>
<tr>
<td>• Eliminate activities that cannot be stopped if the joints get tired.</td>
</tr>
<tr>
<td>• Learn deep relaxation to release tension by meditation, music, relaxation tapes, imagery, or biofeedback.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Exercise</th>
</tr>
</thead>
<tbody>
<tr>
<td>• For muscle strength and range of motion</td>
</tr>
<tr>
<td>• Full general range of motion or aquatic therapy; however, isometric exercises in a pain-free ROM are best during a flare-up</td>
</tr>
</tbody>
</table>

---

### SUMMARY OF REMEDIATION STRATEGIES

Evaluation and management strategies for promoting independence in self-care represent, from the client’s standpoint, one of the most practical and important aspects of medical rehabilitation. As a rehabilitation goal, performance of BADL and EADL is now being viewed as an important outcome with profound implications for quality of life. In this chapter, emphasis has been given to the importance of client/professional collaboration in setting goals and selecting methods for managing life tasks after discharge. These options include determining if the task is feasible, if retraining or new training is desirable, if the environment needs to be altered, if assistance needs to be provided through other people, or if ATDs, including robotics and environmental control systems, may be useful.

In this collaborative process, the client’s preferences, experiences, and post-discharge living environment assume at least as much importance as the diagnosis or physical limitations. Moreover, such factors as costs of time and energy must enter into decisions about the value of various options for accomplishing daily living tasks. The active participation of the
individual in determining a plan of care represents the desired outcome of rehabilitation. This goal is ultimately to enable the individual to participate in those activities that bring life satisfaction. When remediation is not possible, compensatory techniques might be useful. This section highlights the standard techniques used by rehabilitation professionals, often in conjunction with remedial approaches.

Adaptation and Compensatory Strategies

When habilitation or relearning fails or is inappropriate, compensatory options remain available. In compensatory intervention, the task to be performed can be adapted or eliminated, performed by a caregiver, or modified through changes to the environment. Table 9-15 summarizes a number of specific applications to disability to adapt or compensate for skills. Prescription and training on assistive devices also enables performance, despite cognitive deficits, or diminished strength or sensation. Finally, an agent or caregiver can assist with task requirements or perform them entirely according to the requirements of the person receiving assistance. Often a skill the client would like to perform in a normal manner can be accomplished successfully some other way. The person with bilateral above-elbow amputations may not do well at feeding using prostheses but may develop superior toe prehension and use the feet rather than the hands to eat, write, and manipulate tools. Adaptive equipment may substitute for lost or impaired abilities that limit function. The use of such devices will be discussed in the next section.

Regardless of how compensatory training is approached, the philosophical principle that should guide intervention is that there are many approaches to accomplishing the

<table>
<thead>
<tr>
<th>TABLE 9.14</th>
<th>Application of Energy Conservation Principles to Specific ADL</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bathing/grooming</strong></td>
<td></td>
</tr>
<tr>
<td>Sit to shave or apply makeup.</td>
<td></td>
</tr>
<tr>
<td>Place a stool in shower/tub and chair outside tub to sit to dry/dress/undress.</td>
<td></td>
</tr>
<tr>
<td>Use a long-handled sponge to wash legs/feet/back to avoid bending.</td>
<td></td>
</tr>
<tr>
<td><strong>Dressing</strong></td>
<td></td>
</tr>
<tr>
<td>Wear loose clothing and slip on shoes.</td>
<td></td>
</tr>
<tr>
<td>Have one dressing area where all clothing can be reached.</td>
<td></td>
</tr>
<tr>
<td>Gather all clothes before dressing.</td>
<td></td>
</tr>
<tr>
<td>Use a long-handled shoe horn, sock aid, or reacher for donning shoes, socks, and pants.</td>
<td></td>
</tr>
<tr>
<td>Sit between UE/LE dressing and after dressing, if necessary.</td>
<td></td>
</tr>
<tr>
<td><strong>Cooking</strong></td>
<td></td>
</tr>
<tr>
<td>Prepare recipes in large quantities.</td>
<td></td>
</tr>
<tr>
<td>Place in separate containers and freeze extra.</td>
<td></td>
</tr>
<tr>
<td>Plan dinners in advance.</td>
<td></td>
</tr>
<tr>
<td>Utilize a big apron or cart on wheels to collect all items at one time to avoid numerous trips to the cabinet.</td>
<td></td>
</tr>
<tr>
<td>To prepare meal and wash dishes, slide heavy objects across countertops.</td>
<td></td>
</tr>
<tr>
<td>Place all frequently used items on countertop or shelves for easy access, to avoid bending or reaching.</td>
<td></td>
</tr>
<tr>
<td>Use paper plates or gather dishes on a cart.</td>
<td></td>
</tr>
<tr>
<td>Soak dirty pots/pans to make scrubbing easier.</td>
<td></td>
</tr>
<tr>
<td>Use a dishwasher when possible.</td>
<td></td>
</tr>
<tr>
<td><strong>Housekeeping</strong></td>
<td></td>
</tr>
<tr>
<td>Sit at the dryer to fold clothes.</td>
<td></td>
</tr>
<tr>
<td>Eliminate the need to iron by buying fabric that have wrinkle protection.</td>
<td></td>
</tr>
<tr>
<td>Make a shopping list or alternatively.</td>
<td></td>
</tr>
<tr>
<td>Have groceries delivered if possible or use a shopping cart out to car to avoid carrying packages.</td>
<td></td>
</tr>
<tr>
<td>Use labor saving devices when possible (e.g., dishwasher, washer, and dryer).</td>
<td></td>
</tr>
</tbody>
</table>
### TABLE 9.15 General Compensation Techniques by Impairment

#### Weakness
- Use lightweight objects, tools, utensils.
- Let gravity assist activities, e.g., use of wheeled carts for transportation.
- Provide external support, e.g., use of a splint or brace to support wrist, sitting down if legs are weak.
- Use an assistive device to compensate for lost function, e.g., use of a universal cuff to improve grasp, if grasp is weak.
- Use assistive techniques, e.g., slide objects along a surface to transfer from place to place.
- Use power-assisted devices and tools, e.g., electric screwdriver.
- Use biomechanical principles to aid in increasing strength, e.g., decrease resistance of opening a door by installing lever handles.
- Use largest joint possible for task.
- Use assistive devices commonly available in community, e.g., automatic door openers to open doors.
- Use both hands when possible.
- Use of energy conservation techniques.

#### Endurance
- Use energy conservation techniques (see Table 9-14).
- Use lightweight objects, tools and utensils.
- Engineer matches between capabilities and activity demands.
- Use principles of joint protection (see Table 9-13).
- Sit to perform tasks, e.g., use shower or bath chairs.

#### Limited range of motion
- Use of assistive devices, e.g., dressing sticks, researchers, or sock aids.
- Eliminate need for end range activities, e.g., use of slip on shoes, or elastic shoelaces to prevent need to bend.
- Use adaptive garments, e.g., larger size, stretch fabrics.
- Use of adaptive closures, e.g., larger buttons, Velcro fastenings, zippers.
- Adaptive utensils, e.g., built-up handles, elongated handles, long straws, universal cuffs, plate guards.
- Adaptive equipment, e.g., flexible shower hose, long-handled devices, electric toothbrush.
- Safety rails for beds, toilets, and bathtubs eliminate the need to reach long distances.
- Assistive technology, e.g., computers for writing.

#### Coordination
- Provide stabilization for object being worked on.
- Stabilization of the most proximal body part so control is improved distally.
- Use assistive devices where possible to prevent movement, e.g., Dyecem, non-skid surfaces.
- Use weighted or heavy utensils.
- Use adaptations to prevent lack of fine motor skills, e.g., Velcro closures for fastening.
- Use seated positioning whenever possible.
- Lever-type doorknobs and faucet handles.

#### Unilateral use of body
- Teach one-handed techniques for activities, e.g., use of adaptive one-sided dressing techniques.
- Use of adaptive devices, e.g., electric razors.
- Use shower or bath chairs to eliminate need to stand.
- Use wall mounted devices, e.g., hair dryers and jar openers.
- Assistive technology for communication devices, e.g., keyboarding for writing, use of headphone for use of phone, use of one-touch dialing.
- Utility cart or wheeled devices for transportation of objects.
- Use of electric devices, e.g., food processor for cooking.

#### Low vision or blindness
- Organize living space, to eliminate clutter, and identify a place for every item and stress need for each item to be replaced after use, e.g., utensils are kept in the same place, items in the refrigerator and in dresser drawers.
- Use organizational strategies for scanning, e.g., left to right, top to bottom.
- Hang matching clothing together.
- Use of Braille labels and/or optical scanners to distinguish labels on medication and clothing colors.
- Use of voice activated devices, e.g., Dragon Naturally Speaking for writing, books on tape, pay bills electronically.
- Use of assistive devices that provide adequate tactile, auditory or kinesthetic cues.
- Use high contrast colors.
- Ensure adequate lighting in workspaces.
- Use devices to magnify workspace or work activities, e.g., magnifiers for reading.
same task. Innovative alterations in task performance may allow people to do something for themselves that under other circumstances they depend on others to perform. It is characteristic of compensatory training that the end result of a client’s activity, whether it is clean teeth or tied shoelaces, is most important, rather than the method used to perform the task.

The adaptation process can typically be thought of as having five steps, highlighted in Table 9-16. These tasks include task analysis, problem identification, implementation of compensation principles and proposed solutions, training and reevaluation for fit. As in the rehabilitative or habilitative process, or in any learning or teaching situation, the task must first be analyzed to determine the demands of the activity. The demands would include both the environmental context in which the task takes place and the performance demands (e.g., reach, manipulation, mobility, communication). Problems presenting barriers to the performance can then be identified, and a match between task demands and performance limitations can be engineered.

The principles of compensation can then be applied. Solutions are proposed and problem-solving strategies modeled for the client. The solution can then be implemented, when equipment is being recommended the rehabilitation specialist needs to be a consumer advocate and understand safe application and use of the equipment. Training and practice, with variability in context are needed to ensure proper use and consistency in use (420–422). This training consists of introduction of device, choice of as natural context as possible for practice, determining the timing of the practice, providing instruction at the client’s level, using good teaching strategies, and reinforcing the use of the device or strategy (371). Finally, the device must be assessed for reliability, durability, and safety. As an example for this sequence, if a client with a hip fracture is not safe to stand in the shower, a shower seat and grab bars might be determined to be the solution for safe bathing at home. The grab bars must first be made of a heavy-duty
material to withstand repeated use at the anticipated force levels and be anchored securely in the studs. Then the therapist must determine teaching strategies that would be most helpful to his or her client, and the client might both practice in a clinical and home setting to ensure reliability in a variety of settings. The therapist might conduct an inspection of the installed device in the client’s home and ensure that the client is using the device properly (369–371).

Environmental Modifications

Both training and devices for aiding in independent self-care must be appropriate to the living space of the client. Modifications of living spaces and the architectural barriers they often impose may greatly enhance meeting ADL goals. Necessary modifications may range from minimal, in the case of rearrangement of furniture, to extensive, when apartments or homes must be specially designed for the disabled. Intermediate to these extremes are the cases in which modifications or additions to existing space and equipment may be used to enhance function.

Although it is ideal to design facilities with sufficient flexibility to accommodate access and support the activities of a population of diverse abilities, many homes and public facilities currently in use create barriers to performance and participation. One compensatory strategy to support self-care needs is to recommend modifications in existing space. During the course of a home visit and with the aim of making environments accessible, usable, safe, and negotiable, the rehabilitation practitioner can suggest many simple accommodations that can facilitate the attainment of self-care goals. Home modification assessments are often essential to determining whether a client is safe to return home.

It is useful here to provide some examples of self-care modifications that support function when specific impairments are present. When upper-extremity range of motion is limited, pots, pans, cosmetics, canned goods, and other essential daily living items can be placed out on counters rather than kept in their traditional places, which often are difficult or impossible to reach. For the client with a visual field deficit such as homonymous hemianopsia, moving the bed and furniture into the client’s intact visual field when viewed from the doorway may make it easier and safer to move around the room. Of course, the view and placement of objects relative to the bed also must be strategically considered for the same reasons.

The range of possibilities for modification of living space to meet the needs of disabled people is extensive. Examples of common modifications include widening doorways, adding ramps, converting dens or family rooms into a wheelchair-accessible bedroom, and modifying door handles and flooring to improve mobility.

Resources such as the U.S. Department of Health and Human Services, Administration on Aging, Rebuilding Together, give explicit guidelines for many common problems (423–426). Unfortunately, lack of information usually is not what prevents clients and caregivers from effecting changes in living space. Although cost is sometimes a barrier, one study showed that only 52% of recommended home modifications were made, with the primary reason for non-adherence being a lack of belief in the benefits of making changes. Some health care providers have attacked these problems directly by developing partnerships between governmental and community agencies to fund programs designed to eliminate architectural barriers within the home.

Use of Assistive Technology Devices (ATDs)

The use of ATDs has as its special mission the application of technology to increase a person’s performance capabilities by compensating for diminished function. Both low- and high-technology devices are available. High-technology systems are characterized by sophisticated electronic components. These include computers, robots, speech synthesizers, and environmental control systems. These electronic aids to daily living allow persons with very limited voluntary movement or degenerative neuromuscular conditions to operate a wide variety of household and workplace appliances without assistance. A study by Jutai et al. showed that electronic aids are perceived positively by clients using them and are expected to be useful by those who are not yet users but anticipate acquisition of the devices (427).

Low-technology items are simple mechanical aids, such as built-up handles for those with arthritis or shoelaces that can be tied with one hand. Such low-technology items are far more numerous than high-technology devices, yet they are sometimes overlooked. The technology of remote control has become increasingly available in modern life. For a person unable to reach a light switch, radio, television, thermostat, door lock, or curtain cord, an environmental control system can provide a new degree of independence. Ultrasonic or infrared signals, sent from a command center, may allow a person with a disability to use a variety of electronically activated appliances.

Perhaps reflecting the growth in the specialty of rehabilitation engineering, the past several decades have seen an unprecedented increase in the numbers and kinds of devices available to assist persons with functional limitations. The increase in the number of ATDs has been so significant that most rehabilitation practitioners cannot keep abreast of developments. Fortunately, in the United States, useful assistive technology databases are available to help identify suitable technologies and other information. In the United States, two federally supported web-based rehabilitation information sources currently exist, REHABDATA (428), provided by the National Rehabilitation Information Center, and ABLEDATA (429), operated under the auspices of the National Institute of Disability and Rehabilitation Research, U.S. Department of Education, ABLEDATA provides a current and searchable database describing commercially available ATDs. REHABDATA provides a broad range of information relevant to disability and rehabilitation, including devices, accessibility, programs, and research.

Research provides insight into the most appropriate strategies for deploying devices and aids to living. For example,
studies have found that the use of assistive technology is more effective than personal care by others for reducing disability in persons with disability (430,431). Fuhrer noted the need for maturation of outcomes research in assistive technology, calling for shared databases, the development of theory for assistive technology intervention, and the implementation of a multi-stakeholder approach for outcomes research (432). Lenker and Paquet (433) propose a model to guide assistive technology outcomes research that views the acquisition and use of devices as a dynamic process that changes over time based on perceived benefits by the user. Jutai et al. (434) propose that outcomes include effectiveness, social significance, and subjective well-being. A measure validated by Sherer et al. may be useful for predicting outcomes with ATD use (435).

To determine which systems or devices each person needs, a comprehensive team evaluation leads to a list of possible solutions for each identified problem, providing ample information for the client to make the ultimate decision in the selection of equipment. In every instance, the goal of assessment is to find the simplest, least expensive device that best meets the needs for the ADL goal. Figures 9-4 through 9-7 provide examples of low-tech devices that are useful in different basic self-care tasks.

Use of Personal Care Attendants (PCAs)
For many more people with severe disabilities, there is no combination of training, devices, or environmental modifications that will enable them to function independently in self-care. For such people to live independently outside an institution, they must depend on family members for support or PCAs to assist them. Part of the rehabilitation process for people who are going to require attendant care is that they learn to recruit, hire, supervise, and, if necessary, terminate PCAs.

Often the most difficult task for the client is that of defining the tasks that require assistance, the degree of assistance

FIGURE 9-4. A long-handled shoe device can assist in dressing for persons with restricted range of motion.

FIGURE 9-5. A reaching device with a claw end such as this stocking aid can be used to assist with dressing.

FIGURE 9-6. Specially designed cups can help compensate for weak palmar grasp.

FIGURE 9-7.
that is desired, and the hours of the day when these tasks need to be performed. Because hours, pay, and working schedules (i.e., often attendants are needed 7 days per week) often are not competitive with other types of employment, turnover among attendants is high. The occupational therapist or other appropriately trained health care professional often can assist the client in writing the job description for a PCA and in some cases may work with the client to be an effective supervisor. Perkinson et al. note that there are multiple options and stages for caregiving whether or not the role is assumed by family members or PCAs. They recommend the use of checklists with required tasks and procedures to assist professionals in understanding the complexities of involving family caregivers or attendants, and emphasize that ethical issues of care are not diminished simply because family members are involved in the caregiving (436).

An important and difficult aspect of working with PCAs is to be able to define all the self-care and daily living activities that will require assistance. Many tasks, such as brushing one’s teeth, dressing, and grooming, are performed on a daily basis. Other tasks, such as washing one’s hair or having a bowel movement, may be performed less frequently but still quite regularly. Still other tasks, such as doing the laundry or changing the bed, are even less frequent but still are performed regularly. Finally, some tasks, such as mending torn clothing or washing windows, are infrequent and not performed regularly. This list must then be categorized into three distinct classes of activities: those that the person can perform alone with or without ATDs, those that the client can perform with some assistance from others, and those that must be performed by the PCA. Research has shown that factors influencing success in the use of PCAs are mutual effort, skill, and commitment of the PCAs and employer with a disability (437).

**Disability Prevention**

Disability prevention pertains to those parts of a plan of care that promote safety or prevent health problems. These strategies seek to identify risk factors and implement preventive measures before an injury or adverse health event occurs. Disability prevention can take the form of training for the client and/or the caregiver, such as in teaching joint conservation techniques for clients with arthritis or in educating clients and caregivers with spinal cord injury on methods for preventing pressure sores or complications of bowel and bladder dysfunction. Monitoring for complications such as contractures, deep vein thrombosis (DVT), or recurrent stroke following CVA is also important.

Another approach to disability prevention involves exercise to promote fitness and counter the negative consequences of a sedentary lifestyle. Using a randomized controlled trial involving elders with osteoarthritis living in the community, Penninx et al. compared an aerobic exercise program, a resistance exercise program, and a control group (receiving attention only to control for socialization effects) for their effects on ADL disability assessed during an 18-month follow-up. The cumulative incidence of ADL-related disability was significantly reduced in the exercise groups in comparison to the control group (438).

Disability prevention also takes place through the use of environmental modifications oriented toward safety, such as attention to home hazards, lighting, and the installation of handrails and bathroom grab bars. Removal of throw rugs and other common obstacles for mobility or sensory limitations may make a nighttime trip to the bathroom considerably safer. Scales oriented toward the systematic assessment of living environments for safety and accessibility have been reported in the literature.

Another important area of safety and disability prevention concerns the prevention of falls. Fall prevention may include exercise programs to improve strength and postural control, improved medication management, careful attention to the selection and use of assistive technologies, and as indicated earlier, environmental modifications to improve safety. Research has shown that foot problems, assistive technology use, dementia or cognitive deficits, and the use of medications for depression, psychosis, or anxiety are factors that increase the risk of falling (439).

**Health Promotion**

A fourth intervention category involves strategies that promote health. Although some of the approaches in this category are similar to preventive strategies, because they may also serve to increase available resources of energy and time, they are considered to be health enhancing (440–443). One example of a strategy in this category is careful planning of the daily routines surrounding BADL and EADL. Conserving available energy and time in the face of competing demands makes it appropriate for the client and caregiver to determine what tasks may be better assigned to caregivers even though the client is capable of performing them with or without assistance.
SUMMARY
Evaluation and management strategies for promoting independence in self-care represent one of the most practical and important aspects of medical rehabilitation. As a rehabilitation goal, performance of BADL and EADL is now being viewed as an important outcome with profound implications for quality of life. In this chapter, emphasis has been given to the importance of client/professional collaboration in setting goals and selecting methods for managing life tasks after discharge. These options include determining if the task is feasible, if retraining or new training is desirable, if the environment needs to be altered, if assistance needs to be provided through other people, or if ATDs, including robotics and environmental control systems, may be useful.

In this collaborative process, the client’s preferences, experiences, and post-discharge living environment assume at least as much importance as the diagnosis or physical limitations. Moreover, such factors as costs of time and energy must enter into decisions about the value of various options for accomplishing daily living tasks. The active participation of the individual in determining a plan of care represents the desired outcome of rehabilitation. The goal is ultimately to enable the individual to participate in those activities that bring life satisfaction.

ACKNOWLEDGMENT
The work of Richard Schwartz, Karin Barnes, PhD, OTR, and Kenneth Ottenbacher, PhD, OTR was instrumental to the development of earlier versions of this chapter and is acknowledged with appreciation. We also thank Charles Hayden and Sarah Gibson for valuable assistance with the present chapter.

REFERENCES


CHAPTER 9 | FUNCTIONAL EVALUATION AND MANAGEMENT OF SELF-CARE AND OTHER ACTIVITIES OF DAILY LIVING | 279


CHAPTER 9 | FUNCTIONAL EVALUATION AND MANAGEMENT OF SELF-CARE AND OTHER ACTIVITIES OF DAILY LIVING | 281


PART I | PRINCIPLES OF ASSESSMENT AND EVALUATION


Disability Determination

The medical evaluation and treatment of physical impairments and associated disabilities embody the clinical practice of rehabilitation medicine. The physiatric emphasis toward evaluating and managing the disabling aspects of illness and injury sets us apart from more traditional medical fields, whose scope and focus rest primarily on the diagnosis and treatment of pathology and medical impairment. However, the past decade has witnessed an expanding interest among general medical providers in acquiring the skills of medical impairment rating (IR) and disability evaluation. A number of organizations now offer training courses on topics pertaining to disability evaluation (e.g., the American Academy of Disability Evaluating Physicians [AADEP, www.aadep.org]; American Board of Independent Medical Examiners [ABIME, www.abime.org]; American Academy of Physical Medicine and Rehabilitation [AAPM&R, www.aapmr.org]), and a certification examination in the emerging field of disability medicine is also available to interested health care providers through the ABIME.

This chapter is intended to underscore the importance of physiatry to disability medicine and of disability determinations to the physiatrist’s practice. Furthermore, it is intended to provide the physiatrist with a conceptual understanding of the current disablement model, a working vocabulary of terms commonly used in disability determinations, and a comparative understanding of the similarities and differences among the major U.S. disability systems within which such determinations take place. In addition, detailed discussions are provided concerning the evaluating and reporting requirements of the independent medical examiner, the IR process, return-to-work determinations, and some of the legal/ethical challenges facing the disability examiner.

MODELS OF DISABLEMENT

The consistency (or lack thereof) of definitions and criteria for disablement is a source of confusion immediately evident to the physiatrist who engages in the process of disability determinations. The taxonomy and associated nomenclature of disablement vary widely, both locally and internationally. Terminologies and definitions may differ significantly between disability systems or even among jurisdictions within a particular disability system. The core concepts and theoretical underpinnings of disability as a social construct are fully discussed in Chapter 54. Nevertheless, the current reference framework for communicating about disability is also briefly reviewed here, to provide the physiatrist with a fundamental and common perspective from which to view the various interpretations placed on the constructs of impairment and disability within each particular disability system.

ICF 2001

The World Health Organization (WHO) has recently adopted the International Classification of Functioning, Disability and Health (ICF) (1). A new terminology and conceptual model of disablement have been introduced that provide a common framework within which to view human functioning and disability from the perspective of the body, the individual, and society. Human functioning encompasses body functions and body structures, activities (personal sphere), and participations (societal sphere). Disability encompasses impairments, limitations (on activity), and restrictions (to participation) that may occur in the presence of a health condition (disorder or disease). Contextual factors, including environmental factors and personal factors, interact within the model as well (Fig. 10-1). Consequently, an individual’s functioning and disability can be considered as the dynamic interaction of one’s health condition and contextual factors in each case.

The disability examiner should become familiar with the constructs and applications of the ICF model to the field of medical disability determinations, where the validity and reliability of reporting requirements are increasingly being stressed.

In practice, however, the real-world applications of some of these terms of disablement are used differently. Workers’ compensation jurisdictions and many insurance companies use disabled to imply that the person can no longer perform the substantial and material duties of an occupation. So, for example, when an individual cannot lift “twenty ton per day” (e.g., throw 400 sacks weighing 100 lb each per day), they are considered disabled for performing that particular job.

To be disabled in the context of the Social Security Administration (SSA) means that a person must be disabled for “all substantial gainful activity” (SGA) in order to receive benefits from this agency. A ballerina who crushes her great toe has most likely concluded her career as a dancer and is disabled from this occupation in the eyes of her long-term disability carrier. SSA would certainly not see her as disabled and would...
point out that there are a wide array of occupations to which the dancer could still potentially apply herself.

Finally, there are important modifiers to the concept of disability. Under workers’ compensation, for example, if a registered nurse can no longer engage in heavy lifting, bending, and stooping on a frequent basis as required by his or her job description, he or she is disabled from his or her regular duty and may be restricted to light duty. In general, when a worker has an acute musculoskeletal (e.g., back) injury, a physician may place the patient on temporary disability, but it is unlikely there will be any permanent disability. When treatment for a particular condition has continued without improvement for more than 12 months, many would consider that the patient’s temporary disability has become a permanent disability. If so, a disability rating is then required by many workers’ compensation jurisdictions, and the physician examiner must determine if the patient’s disability is total (100% of the whole person) or partial (some fraction thereof).

The terms impairment and disability are often used somewhat interchangeably and incorrectly. For example, physicians may be asked in certain jurisdictions (e.g., Iowa) to use the AMA Guides (2) to provide a disability rating although the title and introduction to the book clearly state it is intended to rate impairment and not disability. The reason is simple—different jurisdictions rate and provide compensation for physical impairments differently. There is no way the AMA Guides could satisfy the rules of disability ratings in all of these different settings.

### MAJOR U.S. DISABILITY SYSTEMS COMPARED

The major U.S. disability systems include tort claims liabilities, workers’ compensation, Social Security, private disability insurance claims, and various federal disability programs. A discussion of each of these disability systems follows. Table 10-1 compares these systems graphically (3). The reader is also referred elsewhere for more extensive and in-depth discussions of this topic (4) than is possible here.

### Tort Claims

Before the 1880s, a worker who was injured on the job could only seek redress against his or her employer by bringing civil suit under common law, where cases were decided based on legal precedent. Tort liability was applicable and implied civil wrongdoing, whereby the burden of proof rested with the claimant. A successful tort claim required proof that (a) a legal duty existed, (b) a breach of legal duty occurred, and (c) harm or injury occurred as the direct result. The employer’s defense rested on grounds of employee contributory negligence, assumption of risk, and the “Fellow Servant Doctrine” implicating co-worker contributory negligence. Winning meant a large lump-sum payment for damages, whereas losing meant no reimbursement for medical expenses or time-loss compensation. The odds of winning were unacceptably low, with less than one in five claims settled in favor of the plaintiff. In their present form, tort claims often arise out of personal injury cases (i.e., medical malpractice, vehicular accidents) where criteria for recovery for damages vary by jurisdiction (5).

### Workers’ Compensation

Workers’ compensation arose out of inadequacies of recovery of claims under the tort system. The first workers’ compensation system was introduced by Germany in 1884 (6), and the first U.S. industrial compensation system was established in Wisconsin in 1911. Other states rapidly followed suit. As a result, workers’ compensation now exists in all 50 states as well as the District of Columbia (6).

In the United States, workers’ compensation law is determined on a state-by-state basis. Each state has its own compensation system, and there are significant differences from one system to another. In general, though, compensation systems share the following common features (3):

1. The system is no-fault, and the employer assumes liability for all claims arising as a result of injury or illness directly attributable to the workplace. In return, the worker gives up the right to bring suit against the employer for separate damages in most circumstances except where wanton negligence applies.
2. Benefits are provided only for work-related conditions that are medically determined to arise directly out of and in the course of employment. Causality is established by demonstrating that a compensable condition occurred while the employee was at work and engaged in employment activity.
3. An employee may forfeit eligibility if he or she was intoxicated or engaged in unsafe practices at the time of injury or if he or she was incarcerated or refused to return to work after being cleared medically to do so.
4. The condition must persist beyond a statutory waiting period (typically 7 days).
TABLE 10.1 Major U.S. Disability Systems Graphically Compared

<table>
<thead>
<tr>
<th>Eligible individuals</th>
<th>Work-related disability</th>
<th>Social Security</th>
<th>Disability Insurance</th>
<th>Federal Employees’ Compensation</th>
<th>Longshore and Harbor Workers’ Compensation Program</th>
<th>Railroad Workers and Seamen</th>
<th>Black Lung Benefits</th>
<th>Veterans Disability Programs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Workers’ Compensation</td>
<td>Nonfederal workers injured out of and in the course of employment</td>
<td>Workers &lt;65 or survivors who are or have contributed to the Social Security Trust Fund (SSDI), disabled children, aged, and blind</td>
<td>Those covered by a group or individual long-term disability plan after a period of short-term disability defined by the policy</td>
<td>Federal employees, including U.S. postal service</td>
<td>Maritime employees such as seamen, longshoremen, harbor workers, shipworkers (not seamen)</td>
<td>Railroad workers and seamen</td>
<td>Coalminers</td>
<td>Honorable or general discharge from the armed forces or a survivor of a veteran</td>
</tr>
<tr>
<td>Rating schedule</td>
<td>AMA Guides in many states; special schedules in FL, MN, CA</td>
<td>Disability Evaluation Under Social Security (Listing of Impairments)</td>
<td>None</td>
<td>AMA Guides</td>
<td>AMA Guides</td>
<td>None</td>
<td>CXR, PFTs, ABGs, and physical examination</td>
<td>Physician’s Guide for Disability Evaluation Examinations</td>
</tr>
<tr>
<td>Employability status</td>
<td>Unable to work in one’s own occupation or in lighter duty, if available</td>
<td>Unable to engage in substantial gainful employment that pays $500/month for &gt;12 months</td>
<td>Inability to engage in own occupation up to 2 years or in any occupation, thereafter, depending on the individual plan</td>
<td>Loss of earnings (no schedule loss) due to disability resulting from personal injury sustained while in the performance of duty</td>
<td>Wage loss and schedule loss benefits for injuries arising out of and in the course of employment</td>
<td>Sickness and unemployment benefits from the Railroad Retirement Board</td>
<td>Total disability due to pneumoconiosis</td>
<td>Wage loss and schedule loss for the average person unable to follow a substantially gainful occupation</td>
</tr>
<tr>
<td>Benefits</td>
<td>Survivor benefits, medical and rehabilitation expenses, wage loss benefits. Tort immunity for the employer</td>
<td>Monthly stipend</td>
<td>Wage compensation</td>
<td>66.6%–75% of wages, reasonable medical care. Lump sums not available</td>
<td>Full medical care, death benefits, lump-sum awards, 66.6% of weekly wages</td>
<td>Railroad workers and seamen must pursue action for damages under the Federal Employers’ Liability Act</td>
<td>Disability pension, death benefits, hospitalization, medical care, orthotics, prosthetics, durable medical goods, adaptive modifications $1993 as of 1999</td>
<td></td>
</tr>
<tr>
<td>Maximum monthly benefit</td>
<td>Determined by state statute</td>
<td>Generally 60%–70% of employment income</td>
<td>75% of wages if worker married or has dependents</td>
<td>200% of the current national average weekly wage</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

SSI, supplemental security income; SSDI, social security disability insurance; CXR, chest x-ray; PFTs, pulmonary function tests; ABGs, arterial blood gas. Reprinted from Rondinelli RD, Katz RT, eds. Phys Med Rehabil Clin N Am. 2001;12:500, with permission.
5. A claim must be filed within a specified time period. In general, the employer must be notified within 30 days of illness or injury, and a claim must be filed within 1 year of illness or injury or within 2 years if death is the result.
6. Disability may be temporary or permanent, partial or total.
7. Benefits under workers’ compensation include survivor benefits in case of death, medical and rehabilitation expenses, and wage-loss benefits (generally, two thirds of wages) during the period of temporary disability. Table 10-2 summarizes coverage/benefits under workers’ compensation.

Many states have adopted a second injury fund to encourage employment of individuals with preexisting disabilities, whereby the employer is protected from excess liability for compensation due to preexisting conditions. The second injury fund in Missouri, for example, began after World War II, when employers were encouraged to hire wounded veterans. The principal concept of a second injury fund is to compensate injured parties for the “synergistic” effects of prior and present injuries. For example, if a worker who had previously been blinded in one eye suffered loss of vision in the second eye, the impairment and disability would be much greater than twice the value of the first eye. This compensation for the “synergistic” or “combined” effect of two injuries is the principal intent of a second injury fund. Readers should investigate further by consulting the statutes in their particular state or jurisdiction.

Social Security Disability Insurance and Supplemental Security Income
The SSA is the largest U.S. disability program, assisting between one third and one half of those persons qualified as disabled (7). It includes the following two separate disability programs:

- Social Security Disability Insurance (SSDI) was established in 1956 to assist workers more than 50 years of age who became totally and permanently disabled. It is funded according to the Federal Insurance Contribution Act (FICA) payroll tax, combining deductions for old age and social disability insurance (OASDI). In general, workers are eligible if they are less than 65 years of age and if they have worked in a job covered by SSDI for at least 5 of the 10 years immediately preceding onset of disability. A requisite criterion of disability determination is proof that the individual is unable to engage in any SGA by reason of any medically determinable physical or mental impairment that can be expected to result in death or that has lasted or can be expected to last for a continuous period of at least 12 months (8). Beneficiaries of SSDI can exit the system by death, by reaching age 65, by undergoing continuing disability review (CDR) in which significant medical improvement is determined such that the individual is no longer eligible for disability, or by returning to work. Less than 5% of beneficiaries “fail” CDR and/or eventually return to work (8).

- Supplemental Security Income (SSI) was established in 1972 to provide support for indigent people who were blind, disabled, and less than 65 years of age. SSI operates as a federal-state partnership funded by general revenue (i.e., federal and state income tax). Individuals are eligible even if they have never worked, if they can be shown to have low income/assets (through “means testing”), and if their medical condition is severely incapacitating such that they cannot engage in SGA, according to the same criteria as SSDI (8).

Benefits to recipients of SSDI or SSI are primarily in the form of monthly stipends. However, recipients of SSDI become eligible for Medicare benefits after 24 months of entitlement. Similarly, recipients of SSI automatically become eligible for Medicaid. In some cases, recipients of SSI who also meet means test criteria for Medicaid are entitled to both types of coverage (8).

Private Disability Insurance
Approximately 40 million Americans have private long-term disability coverage, usually available through the workplace. Policies may be individual or group. Group policies are typically made available through the employee’s company and are more affordable. Definitions of disability and criteria for entitlement vary and are stipulated by contractual language rather than mandated by statutory provisions. Typically, employees who become disabled for their usual and customary occupation are covered by short-term disability for an initial period of 90 days, after which their long-term disability policy takes effect. In general, the duration of coverage is finite and specified (typically 2 years). Subsequently, the disabled person will continue to receive benefits only if unable to perform the functions of any occupation as provisionally defined (5).

An important feature of long-term disability policies is whether there is own occupation versus any occupation coverage. Own occupation coverage provides the insured with disability benefits (typically in the range of 60% of normal salary reimbursement to provide incentive to return to work) if they are not able to provide the essential elements of their particular job. Thus, a neurosurgeon would receive reimbursement if
no longer able to perform surgery, even if other physician responsibilities could be completed. Any occupation coverage means, within limits, the employee would be reimbursed only if he or she could no longer perform meaningful work in any related occupation. Again, the criteria vary according to the insurer. Less expensive group long-term disability plans tend to have own occupation coverage for approximately 2 years, and then the worker must be disabled from any occupation to receive further benefits. More expensive individual long-term disability plans tend to have more restrictive own occupation provisions.

**Federal Employees Compensation Act**
The Federal Employees Compensation Act (FECA) provides compensation benefits to civilian employees of the federal government for work-related disability. Coverage extends to federal employees of the U.S. Postal Service and Peace Corps, and certain nonfederal employees, including state and local law enforcement and the Civil Air Patrol. It is a no-fault system, so that federal employees cannot sue the federal government or recover damages under any other statute for work-related injuries (5).

**Longshore and Harbor Workers Compensation Act**
The Longshore and Harbor Workers’ Compensation Act (LHWCA) provides benefits to shoreside maritime employees such as shipbuilders and repairmen. It is a no-fault system administered by the U.S. Department of Labor (5).

**Federal Employers Liability Act and Jones Act**
The Federal Employers Liability Act (FELA) provides disability benefits to employees of the interstate railroad industry. It is a potentially adversarial system reminiscent of tort claims liability, insofar as the claimant can file suit in either federal or state civil court and must prove negligence on the part of the railroad. The employer’s defense is limited to comparative negligence, whereby recovery for damages may be proportionately reduced according to percentage of employer versus employee culpability.

The Jones Act is similar to FELA but covers civilian sailors for permanent disability suffered while in the service of a ship on navigable waters. The claimant must file suit against the ship’s owner, and cases are often settled out of court (5).

**Federal Black Lung Program**
The Federal Black Lung Program provides coverage for coal miners engaged in surface or underground activity, for total disability caused by pneumoconiosis (black lung) arising from employment. Diagnosis is ascertained through specific chest x-ray findings confirmed by certified “B-readers” according to National Institute of Occupational Health and Safety (NIOSH) standards and the International Labour Organization classification system. Disability claims are referenced according to U.S. Department of Labor standards for spirometric and arterial blood gas measurements (5).

**Department of Veterans Affairs**
Veterans who have received an honorable or general discharge from active military duty are eligible for compensation and pensioning within the Department of Veterans Affairs (VA). Entitlement is considered service connected if the disabling condition is due to illness or injury incurred or aggravated during the period of active duty. Non—service-connected entitlement refers to conditions determined to be unrelated to the period of active duty (9). Benefits include disability pensions for service-connected disability as monthly payments to the veteran or spouse/surviving children in the event of death, hospitalization and medical care, orthotic and prosthetic devices, durable medical equipment, and adaptive modifications to home and/or motor vehicle to accommodate disability (5).

**PREVALENCE AND COSTS OF DISABILITY**

The prevalence of disability varies widely in different western countries. For example, the United States, Canada, and Great Britain each have a prevalence of approximately 2% to 3%, West Germany and the Netherlands 4%, and Sweden has climbed to an astonishing 8% of the population. In Britain, the disability rates for chronic low back pain (LBP) are increasing exponentially (10). Our system of social justice dictates that those who are disabled due to severe or catastrophic impairments (e.g., head injury, spinal cord injury, severe multiple trauma) are deserving and in need of a “safety net” of financial and other assistance. However, there are many patients who seek disability support from this same social safety net, often without any clear pathology (e.g., with LBP). If 8% of the working-age population is on disability (many of these due to LBP), this places a huge burden on the remaining persons in the workforce. The prevalence of work disability in the United States by state has been summarized in a recent report. In 2001, the U.S. Census Bureau estimated that there was a resident population of 176,953,784 in the United States between ages 18 and 64; 3.4% of these were disabled. West Virginia, Mississippi, and Kentucky had the highest rates (6.8%, 6%, 6%, respectively), while Utah and Alaska had the lowest (2.0% and 2.1%, respectively). (Source: Disabled Beneficiaries and Dependents Master Beneficiary Record file U.S. Census Bureau, 2001 estimates.)

The costs to society of this “wealth of disability,” much of it related to LBP, are simply staggering (11). They include medical expenditures, lost wages, lost production, consumer cost increases, employee retraining, and litigation. The costs of workers’ compensation claims by case in 2002 varied from $21,087 in California to $3,959 in Wyoming. (Source: National Council on Compensation Insurance. Public Policy Institute of NYS, Inc., 152 Washington Avenue, Albany, NY 12210 [http://www.ppinys.org/reports/jtf2004/workerscomp.htm].) Medical costs previously made up roughly one third of workers’ compensation costs but now make up greater than 55% of total losses (National Council on Compensation Insurance data).
Clearly, the financial stakes and clinical implications of medical disability evaluation and reporting are enormous and require much diligence, preparation, and adequate training on the part of the community of involved medical practitioners. This will help to ensure that claims are handled as competently, straightforwardly, and compassionately as possible and with attention to the fair representation of the interests of all parties including claimants, employers, insurers, and adjudicators alike.

PHYSICIAN EVALUATING AND REPORTING REQUIREMENTS OF THE INDEPENDENT MEDICAL EXAMINATION

In cases of industrial injury where a legal claim dispute arises, the claimant or insurer or a neutral party (often an administrative law judge) may engage the services of an impartial physician to render an independent medical examination (IME) for purposes of determining disability. The disability examiner performing an IME must not have prior familiarity with, or direct involvement in, the case at the time of referral. The disability examiner is expected to represent the interests of all parties to the dispute fairly and impartially and therefore cannot be the treating physician of record.

The physiatrist acting as disability examiner for purposes of an IME can typically be expected to address the following issues in the evaluation and report (12).

Diagnosis and Severity
What is the diagnosis and extent of severity of the condition in question? The disability examiner must review available treatment records and perform a detailed history and physical examination pursuant to the presenting complaint(s), in order to identify and appropriately document objective pathology and supportive findings. Additional diagnostic tests or procedures may be requested and authorized to render a diagnostic impression and do not by themselves denote a treating relationship.

Prognosis
What is the probability of improvement or resolution of impairment, and what residual symptoms or loss of function is expected?

Causality
What is the cause of the condition? Causality refers to the association between a given cause (event capable of producing an effect) and an effect (a condition that can result from a specific cause) within the medical probability. Medical probability is the physician’s estimate that something is more likely than not (likelihood of causation exceeds 50%), as opposed to medical possibility (likelihood of causation is equal to or < 50%). Typically, a work-related injury must be shown to have a proximate cause whereby a direct and unbroken sequence of events produced the condition, without which cause the condition would not have occurred (13). Judgments regarding causality are typically not difficult when a previously healthy worker experiences a clear-cut injury with obvious clinical findings. However, they can be quite challenging when a worker has a history of problems in the currently affected area or when the alleged cause of symptoms is cumulative trauma rather than a single event (14). A careful history is of paramount importance to the determination of causality and should identify the circumstances and events during which injury occurred, documenting any previous injuries, treatments, and preexisting impairment or dysfunction and identifying important psychosocial contributing factors if applicable. In cases of occupational illness, as opposed to injury, such conditions (e.g., carpal tunnel syndrome, cumulative trauma disorders) may be regarded as arising proximately out of employment. Accordingly, the history should focus on occupational activities and risk factors, and outside recreational activities of a repetitive nature should be noted (13).

Necessity and Appropriateness of Diagnostic Testing and Treatments Rendered
Have the necessary diagnostic tests been performed? Has the treatment rendered been medically necessary, effective, and appropriate in terms of intensity, frequency, and duration? The disability examiner should ensure that a sufficient and thorough diagnostic workup and treatment period have been provided.

Additional Diagnostic Testing and Treatments Needed
What additional diagnostic and medical or therapeutic recommendations might improve outcome and hasten recovery? The disability examiner may recommend additional tests and treatments but should avoid contributing to the development of an “illness conviction” mindset through an endorsement of excessive diagnostic inquiries or inappropriate and perhaps futile treatment efforts. Only those treatments with reasonable expectation of further reducing the objective impairment and/or improving function should be endorsed.

Maximum Medical Improvement
Maximum medical improvement (MMI) is the point when a medical impairment becomes stable, so that additional diagnostic tests and/or therapeutic interventions are not reasonably expected to produce further improvement (2). MMI is felt to have occurred when a “sufficient healing period” has elapsed (the AMA Guides previously recognized 6 months as a sufficient healing period; the fifth and subsequent sixth edition no longer specify the duration of sufficient healing), when the medical condition has resolved, or when there is no reasonable ongoing or anticipated progress toward resolution of the condition. From a physiatric perspective, this end point is reached when the impairment is judged stable such that additional physical and functional improvements are no longer tenable (i.e., “maximum functional improvement”). Consequently,
an MMI determination should be based primarily on lack of demonstrable progress toward reducing impairment or achieving measurable functional gains, rather than symptom reduction per se.

Sometimes a patient meets criteria for MMI when the condition is no longer improving but is expected to deteriorate over time. An example would be a fracture into a joint. In this situation, the physiatrist can still declare the patient to have reached MMI. The issue of long-term deterioration is handled in either of two ways, depending on the patient's compensation system. One method requires the physiatrist to provide an estimate of the patient's future medical needs to maintain stability and/or to minimize progression of impairment regarding the condition. The other is to identify that the patient has reached MMI and is ready for claim closure but to note that a reopening of the claim might be necessary and appropriate at a future date.

Impairment Rating and Apportionment
What is the permanent IR? The permanent total or partial physical impairment associated with the condition should be determined according to methods and procedures outlined in the AMA Guides (2), where permitted by law (see below). Jurisdictional rules and reference systems may vary, and the disability examiner must comply with local requirements or risk judicial impeachment of his or her medical opinions and conclusions otherwise rendered.

Return-to-Work Restrictions
Can the injured worker return to the former job and, if so, how soon? Job description, job analysis, and functional capacity evaluation (FCE) can assist the disability examiner in determining fitness for duty following work injury (see later). When valid performance measurements are available, those data should guide the physiatrist's estimate of ability of an injured worker to perform the essential functions of the job safely and effectively. When performance is invalid because of inconsistencies and submaximal effort, the physiatrist's assessment becomes increasingly subjective and conjectural.

MEDICAL IMPAIRMENT RATING
Purpose and Derivation of the AMA Guides
The AMA Guides (2) has been developed as a standard reference to assist the physician in evaluating and reporting medical impairment of any human organ system. The estimates derived according to the AMA Guides are generally applicable to disability claims evaluations under workers' compensation and some private disability systems. In the United States, the AMA Guides is currently used by 44 states, 2 commonwealths, and federal employee compensation systems (in about 90% of U.S. jurisdictions), as required by statute, regulations, or administrative/legal practice in workers' compensation cases (2). Under workers' compensation law, physician input is mandated to the determination of causality and the coordination and direction of medical care, MMI determination, IR, and return-to-work release and/or restrictions. The AMA Guides has undergone periodic revisions since its inception in 1971. The sixth edition is the current and preferred edition; however, prior editions are still in use and mandated in some jurisdictions. Readers are encouraged to become familiar with the laws and requirements of the particular jurisdiction in which they practice and to comply fully with prescribed guidelines and reference materials locally available.

Limitations of the AMA Guides
The impairment percentages listed according to the AMA Guides are intended to represent an informed estimate of the degree to which an individual's capacity to carry out daily activities has been diminished (2). The AMA Guides provides a disclaimer that it is not to be used for direct estimates of work disability (2). Despite this, impairment percentages derived according to the AMA Guides are frequently used directly to determine disability percentages for work loss compensation purposes.

Subjectivity on the part of the claimant and/or examiner, as well as consistency of effort put forth by the claimant, may skew or otherwise affect the reliability of impairment measures (15,16). The physician examiner can frequently expect to encounter elements of symptom magnification, particularly in the presence of chronicity and pain. Exaggerated displays of pain behavior and related inconsistencies should be noted and documented; when properly accounted for, they should not result in inflated IRs or in inappropriate penalizing of the claimant who exhibits them.

The conceptual issues surrounding the definition of pain and measurement of pain behavior are well documented (17). Detailed discussions of the relationships between pain and suffering (18) and pain and disability (19) are presented elsewhere, and the reader is referred to Chapter 24 for a more detailed discussion of pain issues. For purposes of the IR process, the experience of pain is not directly and objectively measurable. The relationship between pain and suffering is essentially subjective. In general, pain behavior should not serve as the sole basis for an IR in the absence of objective corroborative findings.
of thumb/finger/hand, wrist, elbow, and shoulder. The lower extremity is divided into three subunits to include the foot and ankle, knee, and hip. The spine and pelvis are divided into cervical, thoracic, lumbar, and pelvis subunits. Within each regional unit, impairment is calculated for the smallest applicable subunit. Separate impairments within a subunit are added before impairments between subunits are combined within a region. Similarly, individual impairments within a region can be combined between regions. When two impairment values are combined within or between regional units or subunits, the smaller value ($B$) is combined with the larger value ($A$) by the following formula (2) whose combined value ($C$) is

$$C = A + B [1 - A]$$

This adjustment is necessary so that the cumulative impairment for a series of regional units or subunits does not exceed 100%. A reference table of “combined values” is provided to facilitate these conversions (2). If requested to do so, the disability examiner can convert impairments of the upper extremity to a “whole-person” estimate by multiplying by 0.6; similarly, impairments of the lower extremity can be converted by multiplying by 0.4. Impairment of the spine is always determined to the “whole person.”

The “Diagnosis-Based Impairment” Model

In an effort to enhance precision of IRs and increase interrater agreement and to improve upon the evidence base for IRs, the most recent AMA Guides (2) has become more diagnosis based, with the diagnoses themselves being more evidence based. This “diagnosis-based impairment” (DBI) method employs a common grid format with five columns for successive impairment classes 0 to 4 and rows corresponding to commonly accepted diagnoses (2) arranged hierarchically in aggregates of soft tissue, muscle-tendon, and bone-ligament-joint pathologies. For each anatomical region, the grid specifies the diagnostic criteria, which must be fulfilled to identify the proper impairment class for a given condition, and the range of impairment rating IR available for that class and diagnosis. Once the diagnosis and IC are determined, a “severity grade” and specific IR within class is further determined by an iterative process, using specific information from the claimant’s functional history, physical exam, and associated diagnostic test results. This new approach integrates information from the traditional physical exam such as range of motion (ROM) where applicable, without placing undue weight upon such measurements, which can be time-consuming to obtain and are often prone to influence of claimant and examiner subjectivity. Furthermore, the functionally based history, although also subjective, can be sufficiently validated by simple additions to the in-office exam and does provide some new level of sensitivity to the impact of impairment on activities of daily living (ADLs).

Deviating from the AMA Guides

The AMA Guides is intended to be just that—guides to aid the physician in analyzing and reporting data concerning medical impairment. However, in some cases (e.g., pain), objective and reliable data are lacking; in others, adequate normative data concerning function are unavailable. In such cases, the judgment and consensus represented by the AMA Guides are not intended to supersede the individual rating physician, who is expected to exercise independent judgment and go beyond the boundaries and limits specified by the AMA Guides when the situation warrants.

WORK DISABILITY DETERMINATION AND THE AMERICANS WITH DISABILITIES ACT

Medical impairment may affect employability—the capacity to meet the demands and conditions of employment set forth by an employer. Employability requires the physical capacity to travel to and from the job site, to be present at the job site for a sustained period of time, and to perform a predetermined array of tasks and duties in exchange for wages. Knowledge of medical impairment per se may not be a necessary or sufficient criterion for determining employability, and the physiatrist making such determinations must also be aware of and consider other criteria and perspectives offered including the Americans with Disabilities Act (ADA) (20).

Under Title I of the ADA, individuals with disabilities are afforded protection against discrimination in the workplace for private businesses that employ 15 or more persons (20). Accordingly, employability can be viewed in terms of a formal job description (typically available from an employer upon request) that lists the essential functions and associated physical task demands for that specific job. The essential functions are fundamental job duties that, if removed or altered, would change the nature of the job itself. Failure of the employee to perform the essential functions could result in termination for cause. Accommodation involves modification of a job description or workplace to enable an employee with an impairment or disability to otherwise meet the essential functions. The ADA mandates the availability to the disabled of reasonable accommodation (i.e., one that can be accomplished without undue hardship on the employer [in terms of added costs or logistic difficulties] or a direct threat to the health and safety of the disabled employee or any coworkers) (21).

For example, a patient with a herniated nucleus pulposus treated conservatively within the last 3 months now applies for a job that requires lifting seventy pounds, 50 times per day. Most physicians would feel comfortable that the worker would not be physically able to perform the job without a direct threat to his own health, and hopefully if the worker sued, the court would uphold the company decision. Similarly, if a worker had an uncontrolled seizure disorder, and the job required him to operate a cherry picker, which held other workers, most would agree the risk of seizure while operating the machine would be a direct threat to the workers in the compartment.
At present, the ADA concepts and mandates have not been fully integrated into workers' compensation law and instead represent a separate and parallel system. However, the physiatrist may render an ADA-compatible return-to-work determination for an impaired worker that specifically addresses essential functions and employer willingness or ability to accommodate in each case. In essence, work disability can be operationally dealt with in relation to medical impairment, essential job functions, and options for reasonable accommodation as the following scenarios that Table 10-3 illustrates.

Assume an injured worker suffers permanent impairment as a result of partial amputation of the dominant hand. The impact of this defined impairment on disability has a wide potential range, depending on whether the employee is, for example, a construction worker, police officer, concert musician, or surgeon. If the job requires the safe operation of heavy equipment at a construction site and the worker can perform the essential functions without accommodation, no work disability is present. If the job requires the ability to safely and reliably handle small firearms, accommodation may be required and may necessitate reassignment to a dispatching operation. With that accommodation, no work disability is present. If the job involves live concert performances, the individual may be unable to meet the essential functions but may still be able to compose and deliver studio performances through the use of an electronic synthesizer or other suitable equipment. In this case, some degree of work disability is present in spite of reasonable accommodation. If the job requires a high degree of manual dexterity, such as surgery, accommodation might involve use of a surgical assistant or physician extender to provide intraoperative manual assistance. However, medicolegal and financial constraints might pose sufficient logistic barriers to preclude such accommodation (even if available); consequently, work disability is present.

### Functional Capacity Evaluation

FCE is a systematic, comprehensive, and objective measurement of an individual's maximum work abilities (22). The components of the FCE include

1. Generic testing of workers' capabilities as defined by the Dictionary of Occupational Titles (23).
2. Job evaluation to assess the occupational and material handling requirements of a particular job according to the job description.
3. Work capacity evaluation of ability to safely and dependably perform and sustain job-related activities in response to specific job demands (13,24).

FCE generally includes tests of strength, flexibility, endurance, coordination and reaction speed, functional performance, and safety. Functional tests include material handling (e.g., lifting and carrying, pushing and pulling), sitting, standing, walking, reaching, stooping, crouching, balancing, and climbing. In some cases, hand activities to assess grip strength, manipulation, and dexterity are included. These parameters are quantified in terms of weight and frequency. Terms used to modify frequency include occasional, which generally refers to ≤33% of the time during the work day; frequent or 34% to 66% of the work day; and constant or frequent, which is 67% or more of the work day. Strength requirements for any occupation can be placed into categories varying from sedentary to very heavy work, as described in Table 10-4.

An FCE is typically administered by a trained physical or occupational therapist, and choice of equipment for FCE testing should be driven by considerations of safety, reliability, validity, practicality, and utility in each case. The validity and reliability of test results are limited in the presence of submaximal effort. The availability and feasibility of FCE testing are often limited by cost.
The conflicts inherent to the roles of treating versus rating physician have been previously documented (26) and are summarized briefly as follows: The treating physician acts primarily as a patient advocate and seeks to diagnose and treat in order to minimize suffering. In treating the injured worker, the physician may prioritize symptom alleviation and functional recovery ahead of return-to-work considerations. Case termination becomes an objective only after medical impairment is minimized and return to work, when feasible, has been achieved. By contrast, a rating physician may face an inverse set of priorities, whereby end points of IR and work disability determination are the primary objectives of the referring party. Case termination is of paramount importance in order to enable a rating to occur (see MMI determination), and return-to-work considerations are a priority objective to enable case closure. Satisfactory diagnostic and therapeutic results are of interest in promoting successful return to work. The physiatrist, as treating or rating physician, can bridge this conflict and perform both roles equally by maintaining a priority focus on functional improvement. Diagnostic evaluations and medical and rehabilitative treatments remain viable objectives as long as functional improvement during treatment can be demonstrated. When functional improvement is no longer tenable, case termination is warranted, and an IR and work disability determination can be made. If further care of a “non-restorative” nature appears indicated beyond case closure, the physiatrist may advocate for treatment outside the workers’ compensation system through alternative legal and administrative channels. It is thereby possible for the injured worker and physiatrist to maintain a functionally based therapeutic alliance throughout their course of interactions.

### Legal and Ethical Considerations

Medicolegal aspects of rehabilitation medicine are discussed in detail in Chapter 21, to which the reader is referred for discussion of expert witness testimony and related issues. It is important to note that medicolegal accountability by physicians continues to increase for their opinions rendered with respect to workers’ compensation disability and IMEs. The physiatrist can expect to be frequently deposed or offer courtroom testimony regarding findings and opinions pursuant to a disability determination and may be held accountable for details recorded months or even years previously. Consequently, data collection, organization, and reporting should be carried out in a thorough, systematic, and sufficiently detailed manner to facilitate retrieval of specific information at any future time.

The disability-evaluating physiatrist must embrace a sphere of concern beyond patient advocacy to also include the interests of the employer and insurance carrier alike. Pressure may be applied to terminate treatments that are primarily palliative in nature, to hasten the injured worker’s return to work or

### TABLE 10.4

<table>
<thead>
<tr>
<th>Degree of Strength</th>
<th>Amount of Lifting/Carrying</th>
<th>Posture; Other Activities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sedentary work</td>
<td>Occasional: ≤10#</td>
<td>Primarily sitting; walking and standing at most occasionally</td>
</tr>
<tr>
<td>Light work</td>
<td>≤20, ≤10# frequently</td>
<td>Significant walking/standing or primarily sitting but requiring pushing and pulling of arm and/or leg controls</td>
</tr>
<tr>
<td>Medium work</td>
<td>≤50#, ≤20# frequently</td>
<td>Unspecified</td>
</tr>
<tr>
<td>Heavy work</td>
<td>≤100#, ≤50# frequently</td>
<td>Unspecified</td>
</tr>
<tr>
<td>Very heavy work</td>
<td>&gt;100#, ≥50# frequently</td>
<td>Unspecified</td>
</tr>
</tbody>
</table>

This chart is based on the United States Department of Labor System for classifying general strength of the occupations described in the Dictionary of Occupational Titles (1977 and supplements).
FIGURE 10-2. Job analysis and physical demand comparison.

objectify reasons for not doing so, and to render objective opinions on medical IR and permanent work disability. Ethical
concerns include the potential that the physiatrist acting as a
“double agent” might exploit patient candor and trust or inap-
propriately apply medical skills and authority to nonmedical
agendas under workers’ compensation law. The very process
of MMI determination and work disability determination is
potentially counter therapeutic insofar as it conveys a message
that treatment is ended and future recovery is not anticipated.
Ethical dilemmas are posed whenever physicians may face
financial or other incentives (i.e., future IME referrals) to ren-
der IRs that are excessively liberal or conservative and/or to
limit or withhold treatment. Finances should never influence
such decisions (13).

The physiatrist facing these ethical challenges should
recognize that compensable injuries have potential inherent
disincentives to recovery that tend to promote disability. Furthermore, physician actions (regardless of intent) that enable excessive or inappropriate diagnostic and therapeutic efforts, or result in prolongation of claims, may help to confirm a “disability conviction” on the part of the patient, further impeding functional recovery. The difficult questions concerning terminating treatment, rating impairment, and determining work disability must be handled with finesse and dexterity and, above all, with the fairness, objectivity, and consistency that a functionally oriented focus can provide. Such complex decision making should be routinely predicated on a sound understanding of the functional implications of a particular occupational illness or injury, and a firm commitment to achieving the highest possible functional outcome in every case.

GLOSSARY

Disability. An umbrella term for impairments, activity limitations, and/or participation restrictions in an individual with a health condition, disorder, or disease.

Impairment. A significant deviation, loss, or loss of use of any body structure or function in an individual with a health condition, disorder, or disease.

Impairment evaluation. Acquisition, recording, assessment, and reporting of medical evidence, performed by a licensed medical doctor or surgeon, using a standard method such as described in the AMA Guides, to determine permanent impairment associated with a physical or mental condition.

Impairment rating. Consensus-derived percentage estimate of loss of activity, which reflects severity of impairment for a given health condition, and the degree of associated limitations in terms of ADLs.

Independent medical examination (IME). A usually one-time evaluation performed by a licensed medical doctor or surgeon who is not treating the patient or claimant, to answer questions posed by the party requesting the IME.

REFERENCES

INTRODUCTION

Functioning is the lived experience of people (1). It is a universal human experience (1,2) in which body, person, and society are intertwined (3,4). Over the life span, people may experience a variation in the level of functioning associated with congenital disorders, injuries, acute and chronic health conditions, and aging. The experience of a limitation of functioning or disability thus is part of the human condition (1). The WHO estimates that as many as 500 million healthy life years are lost each year due to disability associated with health conditions. These are more than half the years that are lost annually due to premature death (5). This figure is increasing not only due to population growth, medical advances, and the ageing process but also due to malnutrition, war, violence, road-traffic, domestic, and occupational injuries, and other causes often related to poverty (6).

With the International Classification of Functioning, Disability and Health (ICF) approved by the 54th World Health Assembly in 2001 (4), the WHO provides a universal and internationally accepted framework and classification (7). The ICF is a promising starting point for the integrative understanding of functioning, disability and health, and the overcoming of Cartesian dualism of body and mind as well as both sociological and biomedical reductionism (8). It is also a promising starting point for the development of rehabilitation practice and research (7,9).

The objective of this chapter is to introduce the reader to the ICF and how it relates to rehabilitation.

In the first section, we review the history and development of the ICF and describe its structure and validity. In the second section, we illustrate how to use the ICF for the classification and measurement of functioning. Finally, we discuss the current state of the implementation and application of the ICF with a focus on rehabilitation.

THE ICF

The ICF in the Historical Perspective

Clinicians have relied on classifications for the diagnosis of health conditions for over 100 years (10,11). The International Classification of Diseases (ICD) was first published as a classification of causes of death in 1898 (12). In the meantime, the ICD is undergoing its 11th revision. The ICD was initially used for actuarial reasons to document death. It was later adopted for epidemiology and by public health to monitor health and interventions. Lately, it was used for clinical purposes, mainly driven by the need to classify diagnoses in the context of reimbursement systems including diagnostic-related groups.

By contrast, the first classification of disability, the International Classification of Impairment, Disabilities and Handicaps (ICIDH) (13) was published and released in 1980 for trial purposes only. The ICIDH and other models like the Institute of Medicine model (14,15), Nagi’s model (16,17), and the Quebec model (18) have influenced the definitions of rehabilitation (9), the development of rehabilitation practice and research (9), and legislation and policy-making (7,15). The ICIDH model of disablement represented a real breakthrough in that disability was disentangled from disease by removing the disability section from ICD-8 and creating a separate classification.

Particularly in Europe, there was considerable interest in the application of the ICIDH as a unifying framework for classifying the consequences of disease during the last 20 years of the 20th century. For example, the Council of Europe launched its Recommendation No. R (92) 6 on “a coherent policy for people with disabilities” based on the ICIDH (19). Other publications by the Council of Europe, for example, about the use and usefulness of the ICIDH for health professions (20) document this interest.

However, the ICIDH, which was never approved by the World Health Assembly as an official WHO classification, did not find worldwide acceptance (1,15). It was criticized by the disability community over time for the use of negative terminology, such as handicap, and for not explicitly recognizing the role of the environment in its model. In the reprint of the ICIDH in 1993, WHO thus expressed its intention to embark in the development of a successor classification.

The ICF in the WHO and the UN Perspective

The endorsement of the ICF by the 54th World Health Assembly in May 2001 mirrors an important shift in the understanding of health and disability by the WHO. The ICF acknowledges that every human being can experience a decrement in health and thereby experience some disability. With the ICF, WHO responds to the need for a unified, international, and standardized language for describing and classifying health
and health-related domains. The ICF is WHO’s framework for health and disability. It is the conceptual basis for the definition, measurement, and policy formulations for health and disability. The ICF thus complements the ICD that is used to classify deaths and diseases (11). To complement mortality or diagnostic data on morbidity and diseases is important since they alone do not adequately capture health outcomes of individuals and populations (e.g., diagnosis alone does not explain what patients can do, what their prognosis is, what they need, and at what treatment costs) (21,22).

As an international standard ICF contributes to various WHO’s efforts related to the measurement of health and disability. For example, the ICF served as a framework for WHO’s World Health Survey conducted in 70 countries (23,24). The WHA resolution 58.23 on “Disability, including prevention, management and rehabilitation” approved in May 2005 by the 58th World Health Assembly recalls the ICF framework (8). For the upcoming WHO World Report on Disability and Rehabilitation, the ICF provides the basis for the conceptualization of disability and reference framework for the disability statistics presented in the report. The International Society of Physical and Rehabilitation Medicine (ISPRM) which is the International Physical and Rehabilitation Medicine (PRM) organization in official relation with WHO, is represented on the advisory board of the report and is supporting the DAR team in this development.

While the ICF has been developed by WHO, the specialized agency responsible for health within the United Nations (UN) system, the ICF has been accepted as one of the UN social classifications (4). Therefore, the ICF has influenced the characterization of disability in the UN Convention on the Rights of Persons with Disabilities (25) approved on 13 December 2006 at the UN Headquarters in New York. While the convention does not establish new human rights, it does define the obligations on states to promote, protect, and ensure the rights of persons with disabilities. Most importantly, it sets out the many steps that states must take to create an enabling environment so that persons with disabilities can enjoy inclusion and equal participation in society. However, the ICF provides a more comprehensive approach defining disability than it is used in the UN Convention. Hence, there is the need for a common agreement on the meaning of disability (26).

Development of the ICF

The ICF was developed by the WHO in a worldwide collaborative process involving the active participation of some 65 countries and a network of WHO Collaboration Centers for the Family of International Classifications (WHO-FIC). After three preliminary drafts and extensive international field testing, including linguistic and cultural applicability research, the successor classification which was first tentatively named ICIDH-2, the ICF was finalized in 2000 (4). So far, the ICF has been translated into 37 languages.

The ICF not only was derived from Western concepts but has worldwide cultural applicability. The ICF follows the principle of a universal as opposed to a minority model. Accordingly, it covers the entire lifespan. It is integrative and not merely medical or social. Similarly, it addresses human functioning and not merely disability. It is multidimensional and interactive, and rejects the linear linkage between health condition and functioning. It is also etiologically neutral which means functioning is understood descriptively and not caused by diagnosis. It adopts the parity approach which does not recognize an inherent distinction or asymmetry between mental and physical functioning.

These principles address many of the criticisms of previous conceptual frameworks and integrate concepts established during the development of the Nagi model (16,17) and the Institute of Medicine model of 1991 (14,15). Most importantly, the inclusion of environmental and personal factors, together with the health condition, reflect the integration of the two main conceptual paradigms that had been used previously to understand and explain functioning and disability, that is, the medical model and the social model.

The medical model views disability as a problem of the person caused directly by the disease, trauma, or other health conditions and calls for individual medical care provided by health professionals. The treatment and management of disability aim at cure and target aspects intrinsic to the person, that is, the body and its capacities, in order to achieve individual adjustment and behavior change (27,28).

By contrast, the social model views disability as the result of social, cultural, and environmental barriers that permeate society. Thus, the management of disability requires social action, since it is the collective responsibility of society at large to make the environmental modifications necessary for the full participation of people with disabilities in all areas of social life (29–32). The ICF and its framework achieve a synthesis, thereby providing a coherent view of different perspectives of health (1).

ICF Update and Future Developments

The ICF published in 2001 will—similar to the ICD—undergo updates and ultimately a revision process. The WHO coordinates the update process, in collaboration with the Network of the Collaboration Centers for the Family of International Classifications (WHO FIC CC Network).

Recognizing the importance of personal factors, which are included in the ICF conceptual model, the WHO is also exploring the possibility of developing a taxonomy of personal factors.

To meet the requirements of health and disability information systems in the 21st century, the digitalization of analogue information standards as used with the ICF is essential. This is why the work on an ICF Ontology (defining classification entities with their attributes and value sets) is regarded as a priority for future ICF development.

The Structure of the ICF

As shown in Figure 11-1, the ICF is organized into two parts. Part 1 classifies functioning and disability formulated in two
components: (a) body functions and structures and (b) activities and participation. Part 2 comprises the contextual factors which include the following two components: (a) Environmental factors and (b) Personal factors (currently not classified).

Definitions of some of the key terms used in ICF are given below.

Health condition is an umbrella term for disease (acute or chronic), disorder, injury, or trauma. A health condition may also include other circumstances such as pregnancy, ageing, stress, congenital anomaly, or genetic predisposition. Health conditions are coded using ICD-10.

Functioning is an umbrella term for body functions, body structures, activities, and participation. It denotes the positive aspects of the interaction between an individual (with a health condition) and that of an individual’s contextual factors (environmental and personal factors).

Disability is an umbrella term for impairments, activity limitations, and participation restrictions. It denotes the negative aspects of the ICF that provides a detailed classification with definitions: body functions are the physiological functions of body systems (including psychological functions); body structures are anatomical parts of the body such as organs, limbs, and their components; activity is the execution of a task or action by an individual; participation is involvement in a life situation; environmental factors make up the physical, social, and attitudinal environment in which people live and conduct their lives.

The component of body functions and structures refers to physiological functions and anatomic parts of the body system, respectively; loss or deviations from normal body functions and structures are referred to as impairments. The second component of activities and participation refers to a single list of life domains (from basic learning or walking to composite areas like interpersonal relationships or employment). The component can be used to denote activities or participation or both. “Activity limitations” are thus difficulties the individual may have in executing activities (7). “Participation restrictions” are thus problems the individual may experience with such involvement (7). The components of body functions and structures and activity and participation are related to and may interact with the health condition (e.g., disorder or disease) and contextual factors.

Contextual factors include the components of environmental factors and personal factors. Since in the current ICF, an individual’s functioning and disability occurs in a context, ICF also includes a classification of environmental factors. The components of body functions and structures, activities and participation, and environmental factors are classified based on ICF categories. It is conceivable that a list of personal factors will be developed over the next years. The ICF contains a total of 1,495 meaningful and discrete or mutually exclusive categories. Taken together, the ICF categories are cumulative exhaustive and hence cover the whole spectrum of the human functioning. The categories are organized within a hierarchically nested structure with up to four different levels as shown in Figure 11-2. The ICF categories are denoted by unique alphanumeric codes with which it is possible to classify functioning and disability, both on the individual and population level.
An example of the hierarchically nested structure is as follows: “b1 Mental functions” (first/chapter level); “b130 Energy and drive functions” (second level); and “b1301 Motivation” (third level). Based on the hierarchically nested structure of the ICF categories, a higher-level category shares the attributes of the lower-level categories to which it belongs. In our example, the use of a higher-level category (b1301 Motivation) automatically implies that the lower-level category is applicable (b130 Energy and drive functions).

Because the ICF categories are always accompanied by a short definition and inclusions and exclusions, the information on aspects of functioning can be reported unambiguously. Examples of ICF categories, with their definitions, inclusions, and exclusions are shown in Table 11-1.

### ICF-BASED CLASSIFICATION AND MEASUREMENT OF FUNCTIONING

#### ICF Categories: Building Blocks and Reference Units

ICF categories with their definitions are the discrete, meaningful, and universally shared and understood entities, which allow users to comprehensively classify functioning of individuals and populations.

Qualifiers allow us to indicate the functioning level in a particular category. In the case of the component of environmental factors, they show the degree of positive or negative impact of environmental factors on the individual’s functioning. A most simple rule is binary coding. It assigns a “0” if a functioning problem is not present and “1” if it is present. A slightly more complex coding rule would be to assign a “0” if a functioning problem is not present, a “1” if it is somewhat present and “2” if a functioning problem is fully present. All three components classified in the ICF (body functions and structures, activities and participation, and environmental factors) are quantified using the same generic five-point scale. (“0” = no problem, “1” = mild problem, “2” = moderate problem, “3” = severe problem, “4” = complete problem). The qualifiers for the categories for the activities and participation component are capacity and performance. The qualifiers for the environmental factor categories are facilitators and barriers.

In general, coding based on specified rules is a form of measurement. According to Nunnally (33), measurement is the assignment of numbers to attributes, which represent particular features of an entity. In the specific case of the ICF, coding of the ICF qualifier, therefore, is a form of measurement since...
it involves the assignment of numbers to attributes of ICF categories.

Beyond coding, it is possible to measure attributes based on a scaling model using a scale. An example is the use of a visual analogue scale (scale) based on a metric model (scaling model) for the measurement of intensity of pain (attribute) in relation to pain (ICF category b280). Another example is the use of a self-administered questionnaire (scale) based on a Rasch model (scaling model) for the measurement of a capacity limitation (attribute) in walking (ICF category d450). The advantage of using a scale over coding is the possibility to place a person or a so-called stimulus on a continuum as defined by a suitable scaling model. A Rasch-based scale allows us to place a person on a continuum in relation to the experience of other persons (the population). It also allows us to place, for example, a questionnaire item (a stimulus) on the continuum in relation to any other item (e.g., items included in item banks) representing this attribute of an ICF category.

Based on this understanding, it becomes clear that ICF categories serve as building blocks for both the classification and measurement of human functioning. It also becomes clear how ICF categories are distinct from measurement items. ICF categories represent meaningful and universally shared entities. Conversely, items are stimuli, which allow the quantification of attributes in relation to these entities. Different from the limited number of ICF categories constituting cumulative exhaustive entities of the human experience represented in the ICF, there are a virtually infinite number of measurement items conceivable.

Therefore, ICF categories are the building blocks for the construction of ICF-based tools such as the ICF checklist (34) and the ICF Core Sets (35–38) as well as clinical measurement instruments such as the ICF Core Set Index currently under development for Ankylosing Spondylitis (39) and generic health status measurement instruments such as the WHO Disability Assessment Schedule 2.0 (WHODAS 2.0) (40–43).

ICF-Based Practical Tools: ICF Checklist, ICF Core Sets, and WHODAS 2.0
The ICF classification system contains 1,495 categories. “A clinician cannot easily take the main volume of the ICF and consistently apply it to his or her patients. In daily practice, clinicians will only need a fraction of the categories found in the ICF” (42). Therefore, “to be useful, ICF-based tools need to be tailored to the need of the prospective users without forgoing the information needed for health statistics and health reporting” (42). In response to this need, WHO and a wide range of partners including ISPRM have been collaborating in the development and promotion of ICF-based tools.

ICF Checklist
The ICF checklist provides a user-friendly display of the most relevant ICF categories and allows the user to identify and qualify the individual’s functioning profile in a simple but comprehensive and time efficient manner. The inclusion of diagnostic information in the checklist enables the user to study the relationship between a health condition and the associated functioning problems. The listing of environmental codes and the possibility to record information on personal factors permits the user to document and understand the impact of contextual factors on the person’s functioning.

The ICF checklist has been applied in a wide range of surveys and in studies in the process of developing ICF Core Sets (Table 11-2). As a generic tool for recording and documenting an ICF-based functioning profile, the checklist has a proved utility and feasibility (44–46). In situations where more detailed functional status information is needed, the ICF checklist was found to be too generic and the need for more condition-specific ICF tools (e.g., ICF Core Sets) was noted (47,48).

ICF Core Sets
The ICF Core Set Project
The goal of the ICF Core Set project is to systematically develop parsimonious and hence practical sets of ICF categories for clinical practice, service provision, and research and to link the ICF to health conditions as coded with the ICD (35,38,42). The ICF Core Sets serve first as tools for the documentation of functioning and second as international reference standards for the reporting of functioning (7), irrespective of which measurement instruments were used. They are also the starting point for the development of clinical and self-reported measurement instruments (39,49,50).

The ICF Core Set Project is a joint project of the ICF Research Branch of the WHO FIC CC Germany (DIMDI) at the Institute of Health and Rehabilitation Sciences at the Ludwig-Maximilian-University in Munich, Germany (http://www.ICF-research-branch.org), together with WHO, ISPRM and a large number of partner organizations and associated institutions as well as committed clinicians and scientists (35,38,42).

Conceptual Approach
The conceptual approach for the development of the ICF Core Sets was derived from two perspectives: (a) the perspective of people who share the experience of the same condition (e.g., multiple sclerosis) or condition group (e.g., neurological conditions) and (b) the perspective of the health service context along the continuum of care and the life span.

ICF Core Sets for the Acute Hospital and (Early) Post-acute Rehabilitation Facilities
The ICF Core Sets for the Acute Hospital including the ICF Core Sets for neurological, cardiopulmonary, and musculoskeletal conditions are intended for use by physicians, nurses, therapists, and other health professionals not specialized in rehabilitation care provision (37,38). By contrast, the ICF Core Sets for (early) post-acute rehabilitation facilities including the ICF Core Sets for neurological, cardiopulmonary, and musculoskeletal conditions as well as the ICF Core Set for
<table>
<thead>
<tr>
<th>ICF Core Set</th>
<th>Protocol Paper</th>
<th>Preparatory Phase</th>
<th>Consensus Conference</th>
<th>Validation Phase</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Patient perspective</td>
<td>Expert perspective</td>
<td>Patient perspective</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ICF data collection</td>
<td>Literature review</td>
<td>Delphi method</td>
</tr>
<tr>
<td>Psoriasis and psoriatic arthritis</td>
<td>[39]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ankylosing spondylitis</td>
<td>[40]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spinal cord injury</td>
<td>[41]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Systemic lupus erythematosus</td>
<td>[42]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multiple sclerosis</td>
<td>[43]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Head and neck cancer</td>
<td>[44]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bipolar disorders</td>
<td>[45]</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

n.p., not performed.
geriatric patients are intended for use by physicians, nurses, therapists, and other health professionals specialized in rehabilitation or geriatric care provision (37,38). The use of the term early indicates the early part of rehabilitation where patients have both medical needs requiring hospital care and rehabilitation needs.

ICF Core Sets for Chronic Conditions
The ICF Core Sets for chronic conditions are intended for use in the community-oriented (late) phase of rehabilitation and the community (35,36,42). For each chronic health condition, both a Brief ICF Core Set and a Comprehensive ICF Core Set have been developed. While the ICF Core Sets serve as tools for single encounters, minimum data sets for the reporting of clinical and epidemiological studies and health statistics, the Comprehensive ICF Core Sets are intended for use in multidisciplinary settings.

Generic ICF Core Set
While the condition and context-oriented ICF Core Sets are useful when classifying functioning for patients with specific health problems in specific health care situations, a parsimonious set of categories is needed to be able to assess and compare functioning across conditions and contextual factors. The Generic ICF Core Set is currently being developed in an iterative process involving a number of criteria and methodological approaches. A first study in this process examined the explanatory power of determined ICF categories in relation to external standards across the 12 chronic conditions for which condition-specific ICF Core Sets had already been developed (51). The categories identified as candidate categories from this study are shown in Table 11-3.

Development Process
While there are some singularities in the process of developing ICF Core Sets in relation to the context for which they are being developed, the development as illustrated in Figure 11-3 involves an international consensus process based on evidence gathered in a preparatory phase and an international testing and validation phase in the six WHO world regions (Africa, the Americas, the Eastern Mediterranean, Europe, South-East Asian, and the Western Pacific) (36).

The preparatory phase consists of (a) an empirical data collection based on the ICF, reflecting the perspective and the condition of the patient, (b) an expert survey using the Delphi method, (c) a systematic review on outcomes used in observational and experimental clinical studies, which also represents the view of experts, and (d) additionally, for ICF Core Sets now in the preparatory phase, a qualitative study using focus groups or patient interviews, representing the view of patients complement the methods. The results of the preparatory studies are presented at a consensus conference. They represent the starting point for a structured decision-making and consensus process in which clinicians and health professionals, experts in the field for which the specific ICF Core Set is to be developed, participate. Finally, the ICF Core Sets are tested and validated in an international effort in a wide range of contexts.
In these contexts, three different approaches are used to ensure and examine the validity of the ICF Core Sets. In the first approach, patient focus groups determine whether from the perspective of patients all aspects of functioning are covered by the ICF Core Sets (52). In a second approach, the Delphi technique is used to identify the intervention goals that are relevant for health professionals treating patients. In a subsequent analysis it is examined how these intervention goals are represented by the current version of the Comprehensive ICF Core Sets (53,54). In a third approach, the ICF Core Sets are applied to patients in different settings to gather information about the relevance of the categories. The preliminary results from the patients’ and health professionals’ perspective, support the validity of the ICF Core Sets. However, some categories currently not covered by the ICF Core Sets should be considered for inclusion in an updated version of the ICF (52–54).

WHO Disability Assessment Schedule (WHODAS 2.0)
For the standardized cross-cultural measurement of health status, the WHO has developed an ICF-based assessment instrument, the WHODAS 2.0. The WHODAS 2.0 is chosen from a comprehensive set of ICF items that are reliable and sensitive enough to measure what difference a given treatment makes: assessing the same individual before and after an intervention. It has been designed to assess the activity limitations and participation restrictions experienced by an individual irrespective of medical diagnosis. The primary version of WHODAS 2.0 is a fully structured 36-item lay interviewer-administered assessment instrument. It queries difficulties in six domains of life during the last 30 days. For all six domains, the WHO disability assessment schedule provides different profiles and a single summary score. The following domains were selected after a careful review of existing research and survey instruments and the cross-cultural applicability study (55):

Understanding and communicating with the world (cognition); moving and getting around (mobility); self-care (attending to one’s hygiene, dressing, eating, and staying alone); getting along with people (interpersonal interactions); life activities (domestic responsibilities, leisure, and work); participation in society (joining in community activities).

The WHODAS 2.0 has been shown to be useful for assessing health and disability levels in the general population through surveys; and it is an aid for measuring the clinical effectiveness of interventions and productivity gains as a result of treatments (24,56–63). The psychometric properties of WHODAS 2.0 show good to very good internal consistency (alpha = 0.86), a stable factor structure, a high test-retest reliability (ICC=0.98), concurrent validity, conforms to RASCH scaling properties across populations and is sensitive to change (effect size = 0.44–1.38) in different populations (64–67). A study applying the WHODAS 2.0 in rehabilitation using a German version found that it is a useful instrument for measuring functioning and disability in patients with musculoskeletal diseases, internal diseases, stroke, breast cancer, and depressive disorder (68). However, for the domain household and work activities, a clear distinction between work activities versus household activities was apparent in musculoskeletal and internal conditions (68).

Therefore, a separate scoring of the WHODAS 2.0 with and without the items on remunerated work is provided (54). The WHODAS 2.0 has been translated into over 20 languages and comes in several versions: A 36-item, 12-item version and an adaptive 12+24 item version. All three versions are available in interviewer-, self-, and proxy-administered format. The average time to complete the questionnaire for the 12-item version is approximately 5 minutes and for the 36-item version, 20 minutes.

Mapping the World of Measures to the ICF
Applications
Since the ICF is the universal and standardized language to describe and report functioning and health, users need to be able to map the world of measures to the ICF. The mapping of concepts (qualitative mapping) of measurement instruments to the ICF relies on linkage rules (69,70). The quantitative mapping relies on transformations using the RASCH model (41).

Mapping of concepts is applied for the content comparison of measurement instruments, for example, when studying their comparative content validity. The ICF-based comparison of measurement instruments can, therefore, assist researchers and clinicians to identify and select a most suited measurement instrument for a specified purpose. ICF-based comparisons also enable researchers to ensure that all ICF categories of a suitable ICF Core Set are covered by candidate measurement instruments and hence to report functioning according to standards (7) as described in the last section of this chapter. Table 11-4 lists studies which have compared the most widely used measurement instruments for specified health conditions as well as a comparison of generic health status measures.

Mapping of concepts, in combination with quantitative mapping, is used for the identification of items addressing the construct covered by a specified ICF category and the construction of RASCH scales to estimate the level of functioning for this category. As we will describe in more detail in the following paragraph, this involves the identification of items from measurement instruments that address the construct of a specified ICF category within their scope. Another example of qualitative combined with quantitative mapping is the transformation of information from electronic records (71).

Linkage Methodology
The linking methodology allows users to map contents from measurement instruments to the ICF and vice versa. The first step refers to the identification of concepts within the health-related information to be translated to the ICF. The second step refers to linking those concepts to the ICF.

Step One: Identification of Key Concepts
The first step, the identification of key concepts, varies slightly depending on the origin of the information that is to be
<table>
<thead>
<tr>
<th>Context</th>
<th>Health Condition</th>
<th>Reference</th>
<th>Measurements/Instruments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Early postacute context</td>
<td>Neurological conditions, musculoskeletal conditions, cardiopulmonary conditions, geriatric patients</td>
<td>[1]</td>
<td>Functional Independence Measure; Functional Assessment Measure; Barthel Index</td>
</tr>
<tr>
<td>Long-term context</td>
<td>Obesity</td>
<td>[2]</td>
<td>Bariatric Analysis and Reporting Outcome System; Bariatric Quality of Life Index; Life, Impact of Weight on Quality of Life Questionnaire; LEWIN-TAG Questionnaire; Obesity Adjustment Survey-Short Form; Obesity-Related Distress Questionnaire; Obesity Eating Problems Scale; Obesity-Related Problems Scale; Obesity-Related Well-being Questionnaire; Short-Specific Quality of Life Scale; Obesity and Weight-Loss Quality of Life; Weight-Related Symptom Measure</td>
</tr>
<tr>
<td></td>
<td>Osteoarthritis</td>
<td>[3]</td>
<td>Health Assessment Questionnaire (HAQ); Australian/Canadian Osteoarthritis Hand Index; Cochin scale; Functional Index of Hand OA; Score for Assessment and Qualification of Chronic Rheumatoid Affections of the Hands questionnaire; Arthritis Impact Measurement 2 Short Form questionnaire</td>
</tr>
<tr>
<td>Low back pain</td>
<td></td>
<td>[5]</td>
<td>North American Spine Society Lumbar Spine Outcome Assessment Instrument; Oswestry Low Back Disability Questionnaire; Roland-Morris Disability Questionnaire</td>
</tr>
<tr>
<td>Osteoporosis</td>
<td></td>
<td>[6]</td>
<td>Quality of Life Questionnaire of the European Foundation for Osteoporosis; Osteoporosis Assessment Questionnaire; Osteoporosis Assessment Questionnaire Short Version</td>
</tr>
<tr>
<td>Stroke</td>
<td></td>
<td>[7]</td>
<td>Stroke Impact Scale; Stroke-Specific Quality of Life Scale; Stroke and Aphasia Quality of Life Scale; Quality of Life Index-Stroke Version; Stroke-Adapted Sickness Impact Profile-30; Burden of Stroke Scale; Quality of Life Instrument for Young Hemorrhagic Stroke Patients</td>
</tr>
<tr>
<td>Ankylosing spondylitis</td>
<td></td>
<td>[8]</td>
<td>Bath Ankylosing Functional Index; Dougados Functional Index; HAQ modified for the spondylarthropathies; Revised Leeds Disability Questionnaire</td>
</tr>
<tr>
<td>Chronic obstructive pulmonary diseases</td>
<td></td>
<td>[9]</td>
<td>St. George’s Respiratory Questionnaire; Chronic Respiratory Questionnaire, Standardized Version; Pulmonary Functional Status and Dyspnea Questionnaire, Modified Version; Pulmonary Functional Status Scale; Breathing Problems Questionnaire; Seattle Obstructive Lung Disease Questionnaire; Quality of Life for Respiratory Illness Questionnaire; Airway Questionnaires 20; London Chest Activity of Daily Living Scale; Maugeri Foundation Respiratory Failure Questionnaire; Clinical COPD Questionnaire.</td>
</tr>
<tr>
<td>Generic</td>
<td>Different conditions</td>
<td>[10]</td>
<td>Medical Outcomes Study 36-Item Short-Form Health Survey (SF-36); Nottingham Health Profile (NHP); Quality of Life Index; World Health Organization Quality of Life Scale; World Health Organisation Disability Assessment Schedule II; European Quality of Life Instrument (EQ-5D)</td>
</tr>
</tbody>
</table>
translated. In health-status questionnaires, the concepts refer to the different contents addressed in each of its items. A single item may contain more than one concept. For example, item 8 of the SF-36 “During the past 4 weeks, how much did pain interfere with your normal work (including both work outside the home and housework)” contains three different concepts “pain,” “work outside the home,” and “housework” (69). It is also possible to identify concepts from qualitative data such as interviews (Table 11-5), clinical assessments, and clinical interventions.

**Step Two: Linking of Concepts to the ICF**

After the key concepts have been identified, the second step involves the linking of those concepts to the ICF according to ten rules. An example of the linkage of concepts to the ICF is shown in Table 11-5.

Both steps of the linking methodology should always be performed by two trained health professionals independently of each other. Thus, after the second step, two independent results of the linking process exist. These results are compared. The reliability of the linking process is evaluated by calculating kappa coefficients (72) and nonparametric bootstrapped confidence intervals (73,74) based on the two independent linking results in order to indicate the degree of agreement between the two health professionals. Well-trained health professionals showed a considerable degree of agreement (69) when linking health-status measures to the ICF.

### Table 11.4: Mapping of Measurement Instruments to the ICF (Continued)

<table>
<thead>
<tr>
<th>Context</th>
<th>Health Condition</th>
<th>Reference</th>
<th>Measurements/Instruments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Occupational</td>
<td>Different conditions</td>
<td>[11]</td>
<td>SF-36; Reintegration to Normal Living Index; Sickness Impact Profile; EQ-5D; LHS London Handicap Scale; NHP; Dartmouth COOP Charts; 15-Dimensional Measure of Health Related Quality of Life Test; Assessment of Life Habits; Assessment of Quality of Life; Craig Handicap Assessment and Reporting Technique; Health Utilities Index Mark II; Health Status Questionnaire; Lancashire Quality of Life Profile; Quality of Life Index; World Health Organization Quality of Life Scale Canadian Occupational Performance Measure; Assessment of Motor and Process Skills; Sequential Occupational Dexterity Assessment; Jebsen Taylor Hand Function Test; Moberg Picking Up Test; Button Test; Functional Dexterity Test</td>
</tr>
</tbody>
</table>

### Table 11.5: Illustration of the Linkage Procedure (70) with Parts of a Conversation Recorded During a Focus Group Interview. The Information has been Divided into Meaning Units, Concepts Have Been Identified Within the Meaning Units and They Have Been Linked to the ICF

<table>
<thead>
<tr>
<th>ID</th>
<th>Transcription Divided According to Meaning Units</th>
<th>Identified Concepts</th>
<th>ICF Categories</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>My nails break more. I used to have long, strong nails, but now they break easily. Also, my thumbnails split quickly.</td>
<td>Breaking nails, Thumbnails split, No power in hands, Things drop</td>
<td>b860—Functions of nails, b860—Functions of nails, b7300—Power of isolated muscles and muscle groups</td>
</tr>
<tr>
<td>2</td>
<td>My hands; they are not painful but I have no power. Things often drop.</td>
<td>Nails are not strong, Hair falling out due to medication</td>
<td>d440—Fine hand use, b860—Functions of nail, b850—Functions of hair</td>
</tr>
<tr>
<td>1</td>
<td>For the past couple of years I have noticed that my nails are not strong.</td>
<td>Nails are not strong, Hair falling out due to medication</td>
<td>e1101—Drugs</td>
</tr>
<tr>
<td>3</td>
<td>I have always had bad nails. That is why I cannot judge whether they have become worse. But my hair has been falling out. Could be due to the medication. It is hard to say. It is awful.</td>
<td>Stopping dyeing hair, Strong medication (+) without hair loss</td>
<td>d5202—Caring for hair, e1101—Drugs</td>
</tr>
<tr>
<td>4</td>
<td>I have not lost any hair, but I stopped dyeing it. I thought that, since I already have to take such strong medication, I should do without hair dye and let the natural color grow in again.</td>
<td>Stopping dyeing hair, Strong medication (+) without hair loss</td>
<td>d5202—Caring for hair, e1101—Drugs</td>
</tr>
</tbody>
</table>
Principles of ICF-Based Measurement of Functioning

Measuring a Single ICF Category

In principle, there are two approaches to measure a specified ICF category, that is, to quantify the extent of variation therein. The first is to use the ICF qualifier as a rating scale ranging from 0 to 4 (Table 11-6). The second is to use information obtained with a clinical test or a patient-oriented instrument and to transform this information into the ICF qualifier.

Direct Coding of the ICF Qualifier

With this approach, a physician or health professional integrates all accessible and suitable information from the patient’s history, clinical, and technical exams to code a specified category according to established coding guidelines (75). To ensure quality in a specific setting, it is advisable to regularly assess the reliability of coding (76). Figure 11-4 shows a simple and informative graphical approach to assess the interobserver reliability of ICF qualifier codes (76). The rating of certain ICF categories may be facilitated by complementary instructions provided, in addition to the descriptions of the ICF categories as provided in the ICF reference material. Table 11-7 shows an additional instruction developed by the American Psychological Association (2007) (77) for the ICF category b130 Energy and drive functions for which the original description in the ICF reference material is given in Table 11-6. Similar instructions have been developed by the American Psychological Association for a large number of ICF categories (77).

Transformation of Information Obtained with a Clinical Test or a Patient-Oriented Instrument

With the second approach, the ICF qualifier serves as a reference scale. The results from a clinical test or a patient-oriented measurement instrument are transformed into the ICF qualifier.

For many ICF categories there are suitable clinical tests which include standardized expert and technical exams to code a specified category according to established coding guidelines (75). To ensure quality in a specific setting, it is advisable to regularly assess the reliability of coding (76). Figure 11-4 shows a simple and informative graphical approach to assess the interobserver reliability of ICF qualifier codes (76). The rating of certain ICF categories may be facilitated by complementary instructions provided, in addition to the descriptions of the ICF categories as provided in the ICF reference material. Table 11-7 shows an additional instruction developed by the American Psychological Association (2007) (77) for the ICF category b130 Energy and drive functions for which the original description in the ICF reference material is given in Table 11-6. Similar instructions have been developed by the American Psychological Association for a large number of ICF categories (77).

Transformation of Information Obtained with a Clinical Test or a Patient-Oriented Instrument

With the second approach, the ICF qualifier serves as a reference scale. The results from a clinical test or a patient-oriented measurement instrument are transformed into the ICF qualifier.

For many ICF categories there are suitable clinical tests which include standardized expert and technical exams to code a specified category according to established coding guidelines (75). To ensure quality in a specific setting, it is advisable to regularly assess the reliability of coding (76). Figure 11-4 shows a simple and informative graphical approach to assess the interobserver reliability of ICF qualifier codes (76). The rating of certain ICF categories may be facilitated by complementary instructions provided, in addition to the descriptions of the ICF categories as provided in the ICF reference material. Table 11-7 shows an additional instruction developed by the American Psychological Association (2007) (77) for the ICF category b130 Energy and drive functions for which the original description in the ICF reference material is given in Table 11-6. Similar instructions have been developed by the American Psychological Association for a large number of ICF categories (77).

Transformation of Information Obtained with a Clinical Test or a Patient-Oriented Instrument

With the second approach, the ICF qualifier serves as a reference scale. The results from a clinical test or a patient-oriented measurement instrument are transformed into the ICF qualifier.

For many ICF categories there are suitable clinical tests which include standardized expert and technical exams to code a specified category according to established coding guidelines (75). To ensure quality in a specific setting, it is advisable to regularly assess the reliability of coding (76). Figure 11-4 shows a simple and informative graphical approach to assess the interobserver reliability of ICF qualifier codes (76). The rating of certain ICF categories may be facilitated by complementary instructions provided, in addition to the descriptions of the ICF categories as provided in the ICF reference material. Table 11-7 shows an additional instruction developed by the American Psychological Association (2007) (77) for the ICF category b130 Energy and drive functions for which the original description in the ICF reference material is given in Table 11-6. Similar instructions have been developed by the American Psychological Association for a large number of ICF categories (77).

Transformation of Information Obtained with a Clinical Test or a Patient-Oriented Instrument

With the second approach, the ICF qualifier serves as a reference scale. The results from a clinical test or a patient-oriented measurement instrument are transformed into the ICF qualifier.

For many ICF categories there are suitable clinical tests which include standardized expert and technical exams to code a specified category according to established coding guidelines (75). To ensure quality in a specific setting, it is advisable to regularly assess the reliability of coding (76). Figure 11-4 shows a simple and informative graphical approach to assess the interobserver reliability of ICF qualifier codes (76). The rating of certain ICF categories may be facilitated by complementary instructions provided, in addition to the descriptions of the ICF categories as provided in the ICF reference material. Table 11-7 shows an additional instruction developed by the American Psychological Association (2007) (77) for the ICF category b130 Energy and drive functions for which the original description in the ICF reference material is given in Table 11-6. Similar instructions have been developed by the American Psychological Association for a large number of ICF categories (77).

Transformation of Information Obtained with a Clinical Test or a Patient-Oriented Instrument

With the second approach, the ICF qualifier serves as a reference scale. The results from a clinical test or a patient-oriented measurement instrument are transformed into the ICF qualifier.

For many ICF categories there are suitable clinical tests which include standardized expert and technical exams to code a specified category according to established coding guidelines (75). To ensure quality in a specific setting, it is advisable to regularly assess the reliability of coding (76). Figure 11-4 shows a simple and informative graphical approach to assess the interobserver reliability of ICF qualifier codes (76). The rating of certain ICF categories may be facilitated by complementary instructions provided, in addition to the descriptions of the ICF categories as provided in the ICF reference material. Table 11-7 shows an additional instruction developed by the American Psychological Association (2007) (77) for the ICF category b130 Energy and drive functions for which the original description in the ICF reference material is given in Table 11-6. Similar instructions have been developed by the American Psychological Association for a large number of ICF categories (77).

Transformation of Information Obtained with a Clinical Test or a Patient-Oriented Instrument

With the second approach, the ICF qualifier serves as a reference scale. The results from a clinical test or a patient-oriented measurement instrument are transformed into the ICF qualifier.

TABLE 11.6  ICF Qualifier With Percentage Values Provided by the WHO

<table>
<thead>
<tr>
<th>ICF Qualifier</th>
<th>Percentage of Problem (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 No problem (none, absent, negligible,…)</td>
<td>0–4</td>
</tr>
<tr>
<td>1 Mild problem (slight, low,…)</td>
<td>5–24</td>
</tr>
<tr>
<td>2 Moderate problem (medium, fair,…)</td>
<td>25–49</td>
</tr>
<tr>
<td>3 Severe problem (high, extreme,…)</td>
<td>50–95</td>
</tr>
<tr>
<td>4 Complete problem (total,…)</td>
<td>96–100</td>
</tr>
</tbody>
</table>

aHaving a problem may mean an impairment, a limitation, a restriction or a barrier, depending on the construct; that is, depending on whether we are classifying body functions and structures (impairments), activity and participation (limitations or restrictions), or environmental factors (barriers or facilitators).
TABLE 11.7  Additional Instructions for ICF Categories Illustrated with the Example \textit{b130 Energy and Drive Functions}

\textbf{b130: Energy and drive functions}
General mental functions of physiological and psychological mechanisms that cause the individual to move towards satisfying specific needs and general goals in a persistent manner.

Inclusions: functions of energy level, motivation, appetite, craving (including craving for substances that can be abused), and impulse control

Exclusions: consciousness functions (b110); temperament and personality functions (b126); sleep functions (b134); psychomotor functions (b147); emotional functions (b152)

\textbf{Additional information}
This code includes general behavioral tendencies including \textit{Energy level b1300} and \textit{Motivation b1301} to move toward goals. It also includes the constructs of \textit{Appetite b1302} and \textit{Craving b1303}, which may be general tendencies or relate to specific substances or behaviors (e.g., psychoactive substances, food, gambling). In addition, this code includes \textit{Impulse control b1304}, which may refer to impulses in general or relate to more specific impulses to engage in particular behaviors. This code and its subcodes should be used only to refer to characteristics or behaviors that are consistent or occur frequently over time, not to single behaviors or transitory states.

These codes may be useful in a variety of settings. Motivation, craving, and impulse control are often a part of motivational assessment in relation to substance abuse treatment or other treatments that have the goal of reducing, avoiding, or abstaining from particular behaviors (e.g., substance use, overeating, gambling). In such cases, impairments related to these factors may be a part of the disorder. Energy level and motivation may also be important in cases of CNS injury or disease (e.g., stroke), where concerns related to “lack of initiation” or “mental fatigue” may be present, and in patients with psychological disorders such as depression and bipolar disorder. Impairments in impulse control are, by definition, a part of substance abuse and impulse-control disorders, and may also be a central part of a variety of other psychological disorders including attention deficit hyperactivity disorder, conduct disorder, and bipolar disorder.

Generally, \textit{Energy level b1300} and \textit{Motivation b1301} should be reserved for cases in which abnormal levels or significant changes in energy level and motivation occur as a direct result of a disorder, disease process, or injury, or as an effect of treatment (e.g., decreased energy level is a side effect of some medications).

Motivation is considered to be particularly important in relation to the success of treatment for many health conditions. However, caution should be exercised in assigning this code. Body functions are meant to be coded with the ICF to the extent that impairments are attributable to a health condition or health-related state, which will not be to the extent that high or low motivation is a general personality characteristic of the individual. This is not to say that it will not be highly relevant to treatment, only that it would correspond more closely in this case to what the ICF identifies as \textit{Personal Factors} rather than to \textit{Body Functions}. In addition, \textit{Motivation b1301} should not be used to describe an individual’s motivation to comply with a specific treatment, such as physical therapy in rehabilitation programs. Finally, lack of motivation may be used by health care personnel or others in the patient’s social environment as a pejorative explanation for a patient’s lack of progress in treatment, one that attributes the problem to the patient. It is important not to attribute lack of motivation to patients who are physically or mentally unable to perform particular tasks or actions, or who are not receiving the most appropriate treatments to help them progress.

\textbf{Case Examples}
Following a stroke, a 67-year-old woman has difficulty selecting or getting started on projects, and often complains of feeling “too tired” and “mentally worn out.”

A 45-year-old man with an alcohol abuse disorder refuses all attempts at treatment, indicating that although he recognizes the negative consequences of substance use in his life, he is not willing to stop drinking.

\textbf{Other codes within this section}
- b1300: Energy level
- b1301: Motivation
- b1302: Appetite
- b1303: Craving
- b1304: Impulse control
- b1308: Energy and drive functions, other specified
- b1309: Energy and drive functions, unspecified

category. Figure 11-5 illustrates the construction of an interval reference scale using the RASCH model to estimate the level of functioning for \textit{b130 Energy and drive functions} (41). Sixteen of the 19 items linked from three instruments did fit the RASCH model and could be integrated in an ICF category interval scale. Based on this principle, clinicians can estimate the level of \textit{b130 Energy and drive functions} by adding the responses to the 16 items. In clinical practice, one would obviously need only a subset of possibly five items to reliably estimate the level of functioning in \textit{b130 Energy and drive functions}. 

Following a stroke, a 67-year-old woman has difficulty selecting or getting started on projects, and often complains of feeling “too tired” and “mentally worn out.”

A 45-year-old man with an alcohol abuse disorder refuses all attempts at treatment, indicating that although he recognizes the negative consequences of substance use in his life, he is not willing to stop drinking.
functions. Alternatively, one may increase efficiency by using computer adaptative testing (CAT). Whatever method is used, the obtained raw scores can then be transformed into the ICF qualifier which serves as a reference scale.

A major advantage of the second approach is that the original format of the items used to construct the ICF category interval scale remains unchanged. Thus, it is possible to use the information provided by items within the context of their original instruments and, at the same time, within the context of the ICF. This application can be extremely useful, given the increasing use of the ICF and the ICF qualifier as references when documenting and reporting functioning and disability (79,80).

Measuring Across ICF Categories

**ICF-based Measurement**

The WHODAS 2.0 provides a generic measure of health status and disability as described in the previous section (56,68).

For specific conditions and/or settings, one may want to use a specific measurement instrument. A suitable starting point for the development for such measurement instruments are the ICF Core Sets. The ICF Research Branch of the WHO FIC CC Germany at the University of Munich is thus cooperating with and supporting research groups in the process to develop self-reported questionnaires based on the ICF Core Sets (www.icf-research-branch.org).

Clinician’s ratings of the ICF Qualifier (Table 11-6) across a number of ICF categories, for example, across the categories of an ICF Core Set, can be reported in the form of a categorical profile. A categorical profile across a valid set of ICF categories such as an ICF Core Set provides an estimation of a person’s functioning state. The functioning state is the central information for clinicians when planning and reporting the results of a health care intervention. Table 11-8 shows the example of functioning states at the start and the end of a rehabilitation program.
# Table 11.8: ICF-based Assessment and Evaluation Including Goal Setting and Goal Achievement in a Patient After Spinal Cord Injury. The Functioning States at the Start of Rehabilitation and After 4 Weeks are Shown as Categorical Profiles Based on Expert Ratings of the ICF Qualifier

<table>
<thead>
<tr>
<th>Assessment (12 weeks posttrauma)</th>
<th>Evaluation (16 weeks posttrauma)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global goal: Complete independence, university entry</td>
<td>Not evaluated yet</td>
</tr>
<tr>
<td>Service program goal: Independence in activities of daily living</td>
<td>Not evaluated yet</td>
</tr>
<tr>
<td>Cycle goal 1: d4 Mobility</td>
<td>1</td>
</tr>
<tr>
<td>Cycle goal 2: d5 Self-care</td>
<td>0</td>
</tr>
<tr>
<td>Cycle goal 3: d9201 Sports</td>
<td>2</td>
</tr>
</tbody>
</table>

## Body functions, Body structures, Activities and Participation

<table>
<thead>
<tr>
<th>ICF qualifier</th>
<th>Problem</th>
<th>Goal achievement</th>
</tr>
</thead>
<tbody>
<tr>
<td>b28013 Pain in back</td>
<td>0</td>
<td>✓</td>
</tr>
<tr>
<td>b415 Blood vessel functions—at risk</td>
<td>0</td>
<td>✓</td>
</tr>
<tr>
<td>b420 Blood pressure functions</td>
<td>0</td>
<td>✓</td>
</tr>
<tr>
<td>b7101 Mobility of several joints</td>
<td>0</td>
<td>✓</td>
</tr>
<tr>
<td>b735 Muscle tone functions</td>
<td>1</td>
<td>✓</td>
</tr>
<tr>
<td>b755 Involuntary movement reaction functions</td>
<td>0</td>
<td>✓</td>
</tr>
<tr>
<td>b7603 Supportive functions of the arms—resource</td>
<td>0</td>
<td>✓</td>
</tr>
<tr>
<td>b7800 Sensation of muscle stiffness</td>
<td>0</td>
<td>✓</td>
</tr>
<tr>
<td>s810 Structure of areas of the skin—at risk</td>
<td>0</td>
<td>✓</td>
</tr>
<tr>
<td>d410 Changing basic body position</td>
<td>0</td>
<td>✓</td>
</tr>
<tr>
<td>d4153 Maintaining a sitting position</td>
<td>0</td>
<td>✓</td>
</tr>
<tr>
<td>d4200 Transferring oneself while sitting</td>
<td>1</td>
<td>✓</td>
</tr>
<tr>
<td>d465 Moving around using equipment</td>
<td>1</td>
<td>—</td>
</tr>
<tr>
<td>d4751 Driving a car</td>
<td>0</td>
<td>✓</td>
</tr>
<tr>
<td>d510 Washing oneself</td>
<td>0</td>
<td>✓</td>
</tr>
<tr>
<td>d520 Caring for body parts</td>
<td>0</td>
<td>✓</td>
</tr>
<tr>
<td>d5300 Regulating urination</td>
<td>0</td>
<td>✓</td>
</tr>
<tr>
<td>d5301 Regulating defecation</td>
<td>0</td>
<td>✓</td>
</tr>
<tr>
<td>d540 Dressing</td>
<td>0</td>
<td>✓</td>
</tr>
<tr>
<td>d9201 Sports</td>
<td>0</td>
<td>✓</td>
</tr>
</tbody>
</table>

## Influence of Environmental Factors on Functioning

<table>
<thead>
<tr>
<th>Facilitator</th>
<th>Barrier</th>
<th>Facilitator</th>
<th>Barrier</th>
</tr>
</thead>
<tbody>
<tr>
<td>4+</td>
<td>2+</td>
<td>4+</td>
<td>2+</td>
</tr>
<tr>
<td>3+</td>
<td>2+</td>
<td>3+</td>
<td>2+</td>
</tr>
<tr>
<td>1+</td>
<td>2+</td>
<td>1+</td>
<td>2+</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

## Influence of Personal Factors on Functioning

<table>
<thead>
<tr>
<th>Positive</th>
<th>Neutral</th>
<th>Negative</th>
<th>Positive</th>
<th>Neutral</th>
<th>Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>0</td>
<td>–</td>
<td>+</td>
<td>0</td>
<td>–</td>
</tr>
</tbody>
</table>

*ICF Qualifier range from 0 = no problem to 4 = complete problem in the components of body functions (b), body structures (s), activities and participation (d) and from −4 = complete barrier to +4 = complete facilitator in the environmental factors. In personal factors, the sign + and – indicates to what extent a determined personal factor has a positive or negative influence on the individual’s functioning.
The aggregation of information obtained from a categorical profile using the RASCH model results in a summary score (39,50). In the case of aggregation of information across a valid set of categories such an ICF Core Set, the summary score provides an estimation of a person’s functioning status. If using an electronic clinical chart, the creation of a score from a categorical profile created based on an ICF Core Set does not require additional work. Functioning status information provides clinicians with an intuitive, overall understanding of a patient’s general level of functioning. It can be used by clinicians, service program providers, and payers, for example, for the assignment of patients to suitable rehabilitation service programs, to monitor and manage patient’s functioning along the continuum of care and across service program providers, to evaluate service programs, to predict resources and hence costs, and to derive payment schemes.

The principle of how to develop one- or multidimensional Clinical Measurement Instruments based on clinicians ratings of ICF Core Sets has been recently demonstrated (39,50). It could also be demonstrated how to apply such scores across countries by adjusting for differential item function. It is thus possible to compare functioning status information across countries and world regions.

Validity of the ICF
A wide range of studies across world regions and user perspectives have been examined and have provided empirical and theoretical evidence supporting different aspects of the validity of the ICF framework. They include exhaustiveness, or width, and precision, or depth, of the classification.

Exhaustiveness or Width
A classification needs to be exhaustive by its very nature. In relation to the ICF and its categories, exhaustiveness refers to the coverage of the complete spectrum of health and health-related domains that make up the human experience of functioning and disability, and the complete spectrum of environmental factors that influence that experience of functioning and disability. Exhaustiveness is thus closely related to the concept of width, which refers to the number of distinct health and health-related domains at the same level of specification included in the classification. Based on results of published studies, the ICF appears to fulfill the formal criteria of exhaustiveness, especially in relation to the bandwidth of covered domains. In this respect, the results of the studies conducted in the context of the ICF Core Set development (35–37) (Table 11-2) can be considered “proof of concept.”

The ICF has been shown to be a highly comprehensive classification covering virtually all aspects of the patient experience. More specifically, the ICF has covered the spectrum of problems encountered in people with a wide range of conditions and along the continuum of care. Ongoing validation studies for the ICF Core Sets from the patient and health professional perspectives (Table 11-2) have shown that the ICF broadly covers patient problems and aspects of functioning treated by occupational therapists and physiotherapists (53), and psychologists (53), for example, in patients with rheumatoid arthritis. The results also show that health professionals from different professions differ greatly in their intervention goals, reflecting the importance of validating the ICF from the perspective of many different health professions.

Further proof as to the comprehensiveness of the ICF is the finding that items of a wide range of measurement instruments (Table 11-4) can be mapped to the ICF. Most importantly, the ICF broadly represents the contents of health-related quality of life measures (81).

Precision or Depth
The second consideration for a classification is its depth, or precision. Depth, or precision, can be defined as the number of distinct levels of specification differentiated within a health or health-related domain. Most importantly, the level of specification of ICF categories was established in relation to the human experience of people across a wide range of health conditions, along the continuum of care, along the life span and across the WHO regions. Since the ICF categories are intended to be discrete and meaningful entities, they reflect the intuitive level or the level of informed “lay experts” but not the level of “professional experts” in a specific area.

Few studies have so far explicitly addressed this issue. A study that linked health-related quality of life measures to the ICF found that items with different content are linked to the same ICF category (82). This can be seen as an indication that the ICF does not differentiate these categories adequately. One example is the category b152, Emotional functions. In a review of the items of the SF-36 and the Nottingham Health Profile, different items of these instruments were linked to the same ICF category b152, even though they referred to different emotions. Based on this and other results, the most common emotional functions that could be specified in a future version of the ICF are: sadness, happiness, anxiety, and anger (82).

ICF Framework
Jette has rightly argued that for “scientific investigation, a crucial aspect of any conceptual framework is its internal coherence and its ability to differentiate among concepts and categories within the framework (83). Without empirical differentiation, conceptual frameworks cannot be investigated and validated. One of the common criticisms of the original ICDH was that it was difficult to ascertain the boundaries between the basic concepts; each lacked the clarity and distinctness necessary for useful empirical testing (84–88). Thus, for the ICF to be useful as a framework for research, it is critical that the classification be clear about the phenomena it classifies with distinct and measurable definitions of each component. Without distinct and measurable dimensions, researchers will have trouble using the ICF for hypothesis development, study design and measurement construction” (89).

An important question with regard to the components of the ICF framework is the differentiation of activities and participation (90,91), and, the relationship with capacity and performance. The activities and participation component consists of a single list of functioning domains indicating actions...
and life areas. In annex 3 of the ICF, four alternative options for structuring the relationship between activities and participation in terms of the domain list are given. The main options described in annex 3 of the ICF are (a) Distinct sets of activities domains and participation domains (no overlap); (b) Partial overlap between sets of activities and participation domains; (c) Detailed categories as activities and broad categories as participation, with or without overlap; (d) Use of the same domains for both activities and participation with total overlap of domains. In option (d) all categories can, however, be rated in terms of capacity and performance.

In the ICF publication, the WHO thus noted that “with the continued use of ICF and the generation of empirical data, evidence will become available as to which of the above options are preferred by different users of the classification. Empirical research will also lead to a clearer operationalization of the notions of activities and participation” (4). The respective results from current studies analyzing the differentiation between activities and participation are inconclusive (89,92). In addition, a recent review on the ICF concluded that there is little consensus regarding the distinction between activities and participation (93). However, the insights gained from these studies can contribute to the further development of the ICF.

Another approach to study the validity of the ICF framework is its reflection from the perspective of theoretical or professional models in relation to functioning. For example, occupational therapy models which focus on occupations and activities of daily living in the context of the environment can be expected to be closely related to the ICF. In a paper exploring the link of conceptual occupational therapy models to the ICF, the majority of the concepts from three conceptual occupational therapy models could be linked to the ICF (53,94,95). The ICF also proved to be useful as a framework for comparing the similarities and differences of the three conceptual occupational therapy models. The findings of the study also demonstrated that there are strong conceptual connections between the ICF and occupational therapy models, which encourage occupational therapists to use the ICF in their practice (94).

**IMPLEMENTATION AND APPLICATION OF THE ICF**

**Official Implementation**

With the approval of the ICF, the WHO member states are called upon to implement the ICF in the health, education, labor, and social sector. The implementation of the ICF is coordinated by WHO, in collaboration with the WHO FIC CC Network in which members of WHO and members of WHO Collaborating Centers from all WHO world regions are represented. An ongoing effort is to promote the use of the ICF and improve international comparability of functioning and disability data by establishing standardized procedures and implementation guidelines for different applications of the ICF. Examples of current implementations include the use of the ICF in health and disability surveys, clinical services, disability certification, services use, and education. Especially in the field of rehabilitation, the ICF is exemplary implemented for the rehabilitation management in rehabilitation facilities (96). The ICF and the WHODAS 2.0 have been used in multiple health and disability surveys at global, regional, and national levels (23,24,97–100).

**Implementation and Applications in the Health Sciences**

Parallel to the official implementation activities, the ICF has found immediate interest in the health sciences and particularly rehabilitation (7,43). By 2008 there have been over 600 ICF-related publications reflecting the interest, relevance, and impact of its application in health and particularly rehabilitation research worldwide.

The ICF itself has become the focus of interest of scholars worldwide. It has been critically discussed in a number of papers in recent reports by the Institute of Medicine on the future of disability in America (90,91). ICF-related applications include the use of ICF for the classification and measurement of functioning, as presented in the previous section. In the next paragraph, we discuss applications of specific relevance for rehabilitation. In this context it is important to recall that the ICF is relevant not only to PRM but also to allied professional groups, including physiotherapy, occupational therapy, psychology, and social work. In physiotherapy and occupational therapy, many curricula are now already based on or have integrated the ICF (101). Also, following reports on the application of the ICF in rehabilitation (43,102–104), there are now also reports on the application of the ICF in other medical specialties in which rehabilitation is a major health strategy (19) including psychiatry (105–107) and rheumatology (103). OMERACT, an international group committed to the standardization of outcome measures in rheumatology now uses the ICF as their reference framework (108,109).

**The ICF, a Unifying Framework for the Conceptualization of Rehabilitation**

Rehabilitation is the public health strategy which focuses on functioning and complements the preventive, curative, and supportive strategy (19). The ICF, therefore, is the relevant framework and universally applicable taxonomy for conceptual descriptions and definitions of rehabilitation (19). Initiated by the Journal of Rehabilitation Medicine, the official journal of ISPRM, a process towards universally agreed conceptual descriptions of rehabilitation and the rehabilitation professions has been started in 2007. A brief version of the current conceptual description as shown in Table 11-9 describes rehabilitation as the “health strategy applied by PRM and professionals in the health sector and across other sectors which aims to enable people with health conditions experiencing or likely to experience disability to achieve and maintain optimal functioning in interaction with the environment” (19). A brief version of the conceptual description for PRM as shown in Table 11-10 describes our field as “the medical specialty that based on the assessment of functioning and including the
The ICF, the Basis for Rehabilitation Management

The ICF has the potential to importantly contribute to the quality of rehabilitation care delivery. It can serve as a starting point to structure clinical assessment and rehabilitation management (7,103,111,112) which generally involves the four steps: assessment, assignment, intervention, and evaluation (43). Respective training materials are provided on the internet by Swiss Paraplegic Research which uses the example of spinal cord injury (http://www.icf-case-studies.org).

In the context of a case management situation, the assessment step includes the identification of patients’ problems and the definition of long-term goals and the service program goals of an envisioned rehabilitation program that is specified in the assignment step. The evaluation step refers to the evaluation of service program goal achievement.

In the context of a rehabilitation service program, the assessment step includes the identification of patients’ problems, the review and potential modification of the assigned service program goal, and the definition and modelling of the first cycle goals and intervention targets. Patients’ problems can be assessed in a non-systematic or in a systematic way using the applicable ICF Core Sets. To use a systematic approach using an ICF Core Set either alone or in addition to the nonsystematic approach is advisable to ensure that all potentially relevant problems have been addressed. The systematic approach is particularly useful in the training situation of team members. The structured approach also has the advantage that different team members can take primary responsibility for defined categories.

The assignment step refers to the assignment to health professionals and intervention principles. The intervention step refers to the specification of the intervention techniques, indicator measures, and target values to be achieved in a predefined time period and the monitoring of the intervention. The evaluation step refers to the evaluation of goal achievement with respect to the specified cycle goals and intervention targets.

The ICF framework is useful in all steps and can, for example, in the assessment step be used as a structure for an ICF sheet, with an upper section for the patient perspective and a lower section for the professional perspective as illustrated with an example in Figure 11-7 (102,103). This is a practical way to structure and document patient’s problems and clinical and technical examinations. Also, important personal and environmental factors can be denoted.

The use of the ICF may enhance a structured approach to rehabilitation management and ease the communication of the team with respect to problems, goals, intervention categories, and interventions. Most importantly, it may improve the communication between settings, with insurers or case managers. An ICF-based rehabilitation management approach is also most useful in the context of clinical quality management and assurance, research and evidence-based rehabilitation, training and best practice, and in the formulation, implementation, and evaluation of guidelines. In the context of disability evaluation, the ICF provides a comprehensive framework for assessment and modelling of the determinants of work incapacity.
TABLE 11.10  Conceptual Description of PRM (79). ICF Terms in the Proposed Description are Marked in Bold.

PRM is the medical specialty, which, based on:
- the World Health Organization’s (WHO) integrative model of human functioning and rehabilitation as its core health strategy,
- applies and integrates the biomedical and engineering approach to optimize a person’s capacity with approaches which build on and strengthen the resources of the person, provide a facilitating environment and develop performance in the interaction with the environment,
- involves the diagnosis and treatment of health conditions, assesses functioning in relation to health conditions, personal and environmental factors including prognosis; the potential to change the prognosis; identification of long-term goals, intervention program goals, rehabilitation-cycle goals and intervention targets; and the evaluation of disability
- performs or applies biomedical and engineering interventions to optimize capacity including physical diagnostics, e.g. electro-neurophysiologic testing, assessment of endurance, force and co-ordination; physical modalities such as mechano-therapy including massage, exercise, strengthening and mobilization techniques, heat and cold, water and balneology, light and climate, electric currents including functional electro-physiologic stimulation; neuropsychological interventions; acupuncture, nerve root blockades and local infiltrations; nutritional and pharmacological interventions; rehabilitation technology including, for example, implants, prostheses and orthoses, aids and devices suitable to stabilize, improve or restore impaired body functions and structures including e.g. deconditioning; incontinence; sleep and swallowing disturbance; joint instability; and the minimization of pain, fatigue and other symptoms
- prevent impairments, medical complications and risks including, for example, depression, sleep disturbance, skin ulcers, thrombosis, joint contractures and muscle atrophy, osteoporosis and falls, compensate for the absence or loss of body functions and structures including, for example, amputation; vision and hearing impairments
- leads and coordinates intervention programs to optimize performance in a multi-disciplinary iterative problem-solving process performing, applying and integrating biomedical and engineering interventions; psychological and behavioral; educational and counseling; occupational and vocational; social and supportive; and physical environmental interventions
- provides advice to patients and relevant persons in their immediate environment, service providers and payers over the course of a health condition along and across the continuum of care ranging from the acute hospital to rehabilitation facilities and the community and across sectors including health, education, labor and social affairs
- manages rehabilitation, health and multisectorial services informs and advises the public and decision-makers about suitable policies and programs in the health sector and across the other sectors which provide a facilitating larger physical and social environment; ensure access to rehabilitation services as a human right; and empower PRM specialists to provide timely and effective care with the goal to enable people experiencing or likely to experience disability to achieve and maintain optimal functioning in interaction with the environment

ICF-Based Standards for the Planning and Reporting of Studies

Currently, rehabilitation studies are often difficult to interpret for the reader. A main reason is the widely varying taxonomy due to a lack of a generally accepted framework and taxonomy for functioning, disability, and health before 2001. Authors who all have the same construct in mind referred to and still refer to “function,” “physical function,” “physical functional disability,” “physical disability,” “disability,” “functional limitation,” or “quality of life.” The reader is thus often left wondering what the study is all about. Readers may also find that primary and secondary study endpoints are not explicitly denoted.

The endpoint may be defined in terms of a measurement instrument totally unknown to the reader when there is a wide variety of them in use. And because measurement instruments typically contain a wide variety of constructs, the reader may wonder which constructs are actually measured.

Similarly, the exact targets of rehabilitation interventions, especially in the case of complex and multidisciplinary interventions, are hardly mentioned. Often, only the therapeutic modality (e.g., exercise) or the involved therapeutic professions (physiotherapy, nurse practitioner) are denoted. The authors thus leave it to the readers’ creativity to imagine which impaired body functions, limited activities, or restricted participation
CHAPTER 11 | APPLYING THE ICF IN REHABILITATION MEDICINE

Comprehensive perspective of human functioning based on the integrative model

Human Functioning Sciences

to Society

from the Cell

Biosciences in Rehabilitation

Integrative Rehabilitation Sciences

Professional Rehabilitation Sciences

Biomedical Rehabilitation Sciences and Engineering

Focused perspective of the biomedical aspects of human functioning

**BASIC SCIENCES**
The Human Functioning Sciences aim to understand human functioning and to identify targets for comprehensive interventions with the goal to contribute to the optimization of functioning and the minimization of the experience of disability in the population. The Biosciences in Rehabilitation aim to explain body injury and repair and to identify targets for biomedical interventions to improve body function and structure.

**APPLIED SCIENCES**
The Integrative Rehabilitation Sciences design and study comprehensive assessments and interventions which integrate biomedical, personal factor and environmental approaches suited to optimize people’s performance. The Biomedical Rehabilitation Sciences and Engineering study diagnostic measures and interventions suitable to minimize impairment, including symptom control, and to optimize people’s capacity.

**PROFESSIONAL SCIENCES**
The Professional Rehabilitation Sciences study how to provide best care with the goal of enabling people with health conditions experiencing or likely to experience disability to achieve and maintain optimal functioning in interaction with their immediate environment.

**FIGURE 11-6.** Graphical depiction and short descriptions of a proposal to organize human functioning and rehabilitation research into five distinct scientific fields (110) based on the ICF framework.

**FIGURE 11-7.** ICF assessment sheet for a patient with rheumatoid arthritis.
were targeted by the interventions provided. Statistical analyses often provide more insight into statistical methods than information about the exact handling of the study variables (e.g., about which variables were included or left out in a multivariate analysis).

Also, most current rehabilitation studies do not analyze the mechanism of action, especially in the case of multidisciplinary interventions. Only rarely are the changes in the intervention targets (e.g., improvement in strength and mobility) explicitly denoted and specifically examined to show that they indeed contribute to the explanation of the change in the study endpoint (e.g., physical function). As recent studies have shown, what explains improvements in the study endpoint may often be surprisingly different from what was initially expected and may relate to confounders (e.g., to changes in anxiety and depression) rather than changes in the targeted body function parameters (113).

It seems, therefore, that the systematic use of the ICF taxonomy and ICF-based standards in the description of the study population, the intervention targets, and the analyses may contribute importantly to the quality, readability, and comparability of rehabilitation studies (114). It is thus suggested that researchers report a categorical profile of functioning of the population under study using the appropriate ICF Core Set. To do this, one may use either the ICF qualifier or transform measurements made with a clinical test or self-reported instrument into the ICF qualifier using the described qualitative and quantitative mapping methods as described in the previous section (69,70). It is then suggested to report interventions by denoting the relevant intervention target according to the established linkage rules (69,70). With respect to the analysis section, it is suggested to report the variables used in multivariate analyses using again the applicable ICF code.

**SUMMARY**

A comprehensive understanding of human functioning and the development of programs to optimize functioning of individuals and populations is provided by the conceptual framework of the International Classification, Disability and Health (ICF). The ICF was approved in 2001 by the 54th World Health Assembly as the shared model and classification of functioning, disability, and health. The acceptance and use of the ICF as a reference framework and classification has been facilitated by its development in a worldwide, comprehensive consensus process and the increasing evidence regarding its validity.

Health professionals all over the world need also practical instruments to apply the classification. In this chapter, we have described the development process of the ICF and different tools related to the ICF. After an introduction in the structure of the ICF, we showed how the ICF categories can serve as building blocks for the measurement of functioning. We then described the current state of the development of ICF-based practical tools and standards such as the ICF Core Sets and WHODAS 2.0. We illustrated how to map the world of measures to and from the ICF. We also outlined the methodological principles relevant for the transformation of information obtained with a clinical test or a patient-oriented instrument to the ICF as well as the development of generic and specific ICF-based measures. Finally, we reviewed the validity of the ICF and its implementation and application.

The increasing number of publications reflects the interest and relevance of the ICF in research and clinical practice. There is a wide range of putative applications for the ICF. The ICF serves as a unifying framework for the conceptualization of rehabilitation since it is the relevant framework and universally applicable taxonomy for conceptual descriptions and definitions of rehabilitation. In addition, the ICF has the potential to importantly contribute to the quality of rehabilitation care delivery. It can serve as starting point to structure clinical assessment and rehabilitation management. Finally, the systematic use of the ICF taxonomy and ICF-based standards in the description of the study population, the intervention targets, and the analyses may contribute importantly to the quality, readability, and comparability of rehabilitation studies.

**ACKNOWLEDGMENT**

The authors would like to thank Prof. Jerome Bickenbach and Dr. Somnath Chatterji from WHO for the inspiring discussions of the ICF and related concepts and Gisela Immich, Susanne Stucki, and Jan Reinhardt for their help in the preparation of the manuscript.

**REFERENCES**


20. Heerden YS. *Use and Usefulness of the ICDH for Health Professions (Nursing, Physical Therapy, Speech Therapy, Occupational Therapy, Chiroprapy and Orthopedics)*. Strasbourg: Council of Europe; 1996.


REFERENCES FOR TABLE 11-2

REFERENCES FOR TABLE 11-4


INTRODUCTION

Demands for accountability, improved quality, and the delivery of expected outcomes have grown throughout health care. As the Institute of Medicine’s (IOM) *Crossing the Quality Chasm* states: “The frustration levels of both clinicians and patients have probably never been higher. Health care today harms too frequently and routinely fails to deliver its potential benefits” (1). The primary motivation for quality and outcomes improvement systems in rehabilitation, however, is not the avoidance of bad care or patient injury, but the provision of high quality services that improve the function and quality of life (QOL) of persons with disabilities. Public accountability, including justification of the cost of rehabilitation, is an intrinsic part of this fundamental motivation. Medical rehabilitation facilities are caring environments, and the great majority of rehabilitation patients clearly improve in function (2). The technical basis of quality and outcomes monitoring systems for rehabilitation needs to be developed by professionals who are trained to understand the evidence basis of rehabilitation services, working with professionals experienced with clinical practice. The interests of other major stakeholders, including patient or consumer representatives, payers, and government agencies, must also be represented, even though their values may differ (3). The influence, prosperity, and even the survival of rehabilitation as a specialty may hinge on its ability to develop and implement evidence-based monitoring and management systems relevant to consumers, payers, administrators, and policy makers.

A generation ago, rehabilitation—like most of health care—was widely regarded as an applied art rather than a science. It was commonly argued that professionals could recognize quality if they saw it, but it could not be objectively predefined. The possibility of scientific measurement of outcomes was disputed. Progress has been made since then. It is now widely recognized that scientifically valid instruments are a necessary basis for monitoring the quality and outcomes of rehabilitation programs. A large number of instruments and scales are now available in rehabilitation health care, and they are widely applied in clinical and community settings. For example, inpatient rehabilitation facilities (IRFs) use a standardized assessment protocol, the Inpatient Rehabilitation Facility-Patient Assessment Instrument (IRF-PAI), which includes items from the functional independence measure (FIM) (4). Other current outcomes monitoring systems include the outcomes assessment and information set (OASIS), used in home health settings, and the minimum data set (MDS), used in nursing homes (5). The Centers for Medicare and Medicaid Services (CMS) is presently developing monitoring systems that can be applied across multiple postacute rehabilitation settings, for example, the continuity assessment record and evaluation (CARE). Ensuring the validity of these instruments to evaluate the variety of needs of rehabilitation patients across postacute settings is a current challenge.

The broad thesis running through this chapter is that quality outcome monitoring and improvement efforts in medical rehabilitation must be based on the best available evidence. This evidence must be integrated with clinical experience, and interventions need to be varied and highly sensitive to individual variations in beliefs, values, and circumstances. Assertions that personal opinions alone should rule will neither advance rehabilitation as a profession nor the welfare of people with disability. While we emphasize the need to ground quality and outcomes monitoring evidence-based practice (EBP) and systematic reviews of the scientific evidence, research evidence of effectiveness is typically neither strong nor unequivocal in rehabilitation. Multiple strategies are useful for assuring and improving the quality and effectiveness of medical rehabilitation programs. Monitoring systems need to include measures of both process and outcomes.

In this chapter, we first present basic concepts and principles, necessary for rational communication about quality, outcomes, effectiveness, and evidence. We then review rehabilitation’s long experience with program evaluation (PE) and outcomes monitoring systems. Quality improvement (QI), quality assurance (QA), and Joint Commission approaches are then discussed. Because improving the quality and outcomes of rehabilitation programs will require information systems, such systems are then discussed. Other necessary approaches—including professional education, patient-centeredness, and
“clinical practice improvement” (CPI)—are then discussed. Finally, we will discuss key public issues. We hope the chapter is a useful guide and reference work for physicians, administrators, QI specialists, policy makers, and disability advocates concerned with quality and outcomes in medical rehabilitation.

BASIC TERMS AND CONCEPTS

To discuss quality and outcomes improvement, certain basic terms and concepts need to be understood.

Evidence-Based Practice

The EBP movement has significantly influenced thinking about quality and outcomes monitoring and related benchmarks, guidelines, PE, and performance indicators. The principles of EBP were originally introduced in 1992 under the term evidence-based medicine (6). The concepts and techniques rapidly evolved from a focus on medicine and are increasingly integrated into virtually all health care quality monitoring and accreditation systems. Rehabilitation too needs to adopt principles of EBP. It is now increasingly recognized that measurement of functional gain itself provides only very weak evidence of program quality and effectiveness, as patients may improve even without specialized rehabilitation programming. Much stronger evidence—including evidence from well-controlled clinical trials—is required as a basis to infer provision of effective treatment.

Sackett’s classic definition of evidence-based medicine is “the integration of best research evidence with clinical expertise and patient values” (EBM, 2nd ed., p. 1) (6). In this chapter, we define EBP as the use of best research evidence in clinical and community practice, both in making decisions about individuals and at the level of policy and procedures, integrating this evidence with clinical experience and clients’ values. Best evidence is no longer a matter of unfettered opinion: it is evaluated by systematic application of a predefined hierarchy of research quality. Key features of high-quality intervention studies include randomization or other methods of controlling for selection bias and case severity, blinding and avoidance of measurement biases, and minimization of attrition biases (7–9). Widely accepted standards also exist for evaluation of the quality of diagnostic, screening, and predictive (7) studies; standards also exist for measurement studies (10–12). EBP and systematic review are core to the modern view of performance improvement systems presented in this chapter. More complete information on EBP is found in Chapter 80.

Quality and Outcomes Monitoring

PE is the systematic collection and analysis of information about some or all aspects of a health service program to guide judgments or decision about that program. An effective PE involves procedures that are useful, feasible, ethical, and accurate (13). QA can be defined as all activities that contribute to defining, designing, assessing, monitoring, and improving the quality of health care. These activities can be performed as part of the accreditation of facilities, supervision of health providers, or other efforts to improve the performance of health providers and the quality of health services (14). The term quality assurance (QA) has fallen out of vogue, perhaps because it at one time led to reliance on external policing of clinicians, peer review alone, and other limited techniques. However, QA activities of some type continue to be needed to assure that standards of care are met.

Although PE and QA differ in focus, they are complimentary. PE examines programs in relation to stated objectives and is concerned with identifying and evaluating the structure, efficiency, process, effectiveness, relevance, and impact of the program. QA generally focuses on patient-specific practices of health providers and evaluates these practices with regard to standards expected by the peer group or benchmarks of exemplary practice agreed upon by the profession. In addition to program objectives and professional benchmarks, consumer-focused and outcomes-oriented performance/QIs have received increased public and professional attention.

QI involves applying appropriate methods of evaluation and outcomes assessment to close the gap between current and expected levels of quality as defined not only by professional standards but also by consumers and other stakeholders. The Joint Commission on Accreditation of Health Care Organizations (JCAHO) states that the most important reasons to establish an outcome-oriented assessment initiative are to (a) describe in quantitative terms, the impact of routinely delivered care on patients’ lives; (b) establish a more accurate and reliable basis for clinical decision making by clinicians and patients; and (c) evaluate the effectiveness of care and identify opportunities for improvement (15, p. 25).

Quality of care can be defined in many different ways. “Quality” is always positive connoting activities that benefit the person served in the short- or long-term. The IOM has defined quality as the “degree to which health services for individuals and populations increase the likelihood of desired health outcomes and are consistent with current professional knowledge” (14,16). In other words, quality involves achieving desired health outcomes to a degree that is consistent with current knowledge of diagnosis and effective treatment.

In addition, quality care also requires treating patients with dignity and sensitivity to their individual needs, expectations, and circumstances. Communication, concern, empathy, honesty, sensitivity, and responsiveness to individual patients have long been recognized as necessary attributes of quality health care (17). Patient involvement is particularly important in rehabilitation and chronic care because engaging motivations is essential to the success of activity and behavioral therapies that work to enhance meaningful functional capacities of people served. Individuals served in rehabilitation not only want to be informed about what is going on but also want to be involved in selection of treatment goals (18).

Performance indicators. The terms “performance indicator” and “performance measure” are commonly used to designate key outcomes and processes that need to be measured and reported to judge the effectiveness and efficiency of service
Quality and Outcomes of Medical Rehabilitation Programs

CHAPTER 12

delivery. The choice and implementation of performance indicators are central concerns to performance monitoring and improvement, as objective data are needed as a basis for evaluation. Performance indicators are used in QI and reported to stakeholders such as consumers, payers, governing boards, accrediting organizations, and the public. Joint Commission of Accreditation of Health Care Organizations (JCAHO), National Committee for QA, and government agencies have devoted great effort to develop performance measurement systems over the last two decades.

The term “outcome” is used in different ways: life outcomes, pertaining to role restoration or QOL; health-related QOL, pertaining to aspects of life, experience, or function that are logically related to physical health or recognized mental disorders; and the outcomes of care, here, rehabilitation outcomes. Rehabilitation improves many quality aspects of patients’ lives; however, it would be naïve to suggest that medical rehabilitation can routinely produce or assume responsibility for total or all encompassing improvements in patients’ lives. Although we are concerned with the person’s QOL as a whole (large circle in Fig. 12-1), medical rehabilitation is primarily directed at health-related QOL (smaller oval). Medical rehabilitation professionals are primarily responsible for those valued aspects of patients’ lives that they can affect, namely treatment outcomes (small triangle in Fig. 12-1).

The Joint Commission has defined outcomes as “restoration, improvement or maintenance of the patient’s optimal level of functioning, self-care, self-responsibility, independence and QOL” (19). The term also connotes connection to preceding rehabilitative treatments; the outcome in some sense is due to rehabilitation. It is essential to realize that outcomes due to rehabilitation are not directly measured: they are inferred from prior evidence and theory and estimated using a data set permitting adjustment for case severity and confounding factors that influence measured outcomes.

Benchmarking is basic to both quality and outcomes monitoring. A benchmark is a target value of a performance indicator. Joint Commission requires that facilities compare their processes and outcomes with those known to be attainable elsewhere (20). The Commission on Accreditation of Rehabilitation Facilities (CARF) also writes of benchmarks. Continuous quality improvement (CQI) assumes a scoreboard of process or outcome measures (21–24). Severity or risk adjustment is usually needed to develop accurate benchmarks for evaluating outcomes and processes.

Treatment Effectiveness

Knowledge of treatment effectiveness ties together processes and outcomes. Effectiveness may be defined as the sustained improvement in patient function produced by a care intervention beyond the natural healing and adjustment that occurs with less intensive or specialized care. Assertions that an intervention is effective require evidence, ideally from prior well-controlled studies. Effectiveness is assumed by use of the term “rehabilitation outcomes” and is the core professional attribute of quality medical rehabilitation. Effectiveness encompasses the appropriateness of care, the technical competence with which procedures are carried out, risks, and intended as well as unintended consequences. Both QI and outcomes monitoring systems should be based on best knowledge regarding effective treatment.

Surrogate Indicators of Effectiveness

In practice, PE and outcomes management in rehabilitation are commonly based on implicit beliefs and practical but flawed surrogate estimates of effectiveness such as functional gain. In PE, the term “effectiveness” is often used to mean how successful a program is in accomplishing its goals or the average amount of functional gain by patients. Higher rates of functional improvement may suggest greater effectiveness in some facilities than others but are far from proving it (25).

Older QI publications have defined effectiveness as “the degree to which the care is provided in the correct manner, given the current state of the art” (19,26). In this traditional view, “correctness” is associated with adherence to normatively based standards and methods of care. Studies in acute hospitals, for instance, have provided correlational evidence that better adherence to the expert-defined best practices would improve patient outcomes (27–30). By contrast, an evidence-based approach would ask about the strength of evidence for various expert recommendations, accepting expert opinion tentatively when there are no strong empirical studies.

Efficiency and Value

Efficiency means delivering appropriate, effective care within cost constraints. Straight cost considerations must be distinguished from cost-effectiveness, which involves evaluating costs of treatment against gains in patient outcomes. Measures of service use (e.g., length of stay [LOS], treatments units) are often useful surrogates for detailed computations of cost. In a managed care environment, we especially need to know whether the imposed limitations have compromised patient outcomes. In a prospective payment system (PPS) environment, if one spends too much effort on one patient, there

FIGURE 12-1. QOL, health-related QOL, and rehabilitation outcomes.
will be fewer resources available for others. Data are needed to guide us in determining where the optimum lies.

Rehabilitative care not only needs to be provided efficiently, but it must also be of value to its customers, including most of all patients and also payers (31). Patients, payers, and society commonly demand robust improvement in the patient’s functioning and/or QOL that endures in everyday life after discharge. QALYs—quality adjusted life years—are, in principle, applicable to rehabilitation (32,33), but current methods of computing QOL over time have not yet been shown to be sensitive to rehabilitative interventions (e.g., functional gain) or to the decisions that rehabilitation professionals and people with disability must make (see Chapter 18).

Defining and improving the quality of care would be easy if money were not an issue. Resources, however, are always limited. The difficulty of defining affordable value is greatly complicated when available funding varies enormously across patients and no societal decision has been made regarding standards for valuation of the quality of human life over time (e.g., QALYs) (33). As economic constraints have increased and funding for rehabilitation has become more variable, assuring the provision of high-quality care has become increasingly challenging, at least in the United States.

Levels of Health and Functioning

Health and functioning are rich concepts that involve a number of levels that need to be understood given their central importance to systematic QI. Chapter 19 explains components of health outcomes as defined by the newer International Classification of Functioning, Disability, and Health (ICF) (34). In brief, distinctions among functioning at the biological level, the level of the individual per se, and the level of the individual in society and the environment (previously designated as impairment—disability—handicap) (35) have been replaced by body systems and structure, activity, and participation. Measures of pathology—dysfunction at the cellular or biochemical level—and disease are also needed in QI and outcomes management in medical settings. The term functional limitation is also valuable to denote specific limitations or activity restriction of the person, compared to a normative average and measured in a controlled environment (36,37).

Impairments are of focal importance in medical rehabilitation treatments. Performance monitoring systems in medical rehabilitation must at least group patients by their primary etiology or impairment group, and ideally, severity adjustment is in terms of the primary diagnosis or impairment (e.g., severity and level of paralysis in spinal cord injury) (38). If an impairment is used as an outcome measure, there should be evidence—not merely an assumption—that the impairment is significantly related to functional outcomes or QOL. Numerous medical and nursing conditions treated in medical rehabilitation—infection control, reduction of decubitus ulcers, control of blood pressure, prevention and treatment of deep vein thrombosis, diabetes management, pain relief—meet this criterion. In many circumstances in rehabilitation, however, a pathology or impairment can be treated

and reduced without alteration of the primary disease or the functional status of the patient (39). Range of motion and even spasticity reduction, for instance, are poorly correlated with functional outcomes (40), probably because they are not the primary barriers to improved function for many patients. Discriminating more worthwhile from less worthwhile but still technically effective interventions is a challenge for rehabilitation providers.

Medical rehabilitation deals with all these levels but has long focused on diminishing impairment and improving basic capabilities of persons with disabilities (e.g., reducing assistance requirements in activities of daily living [ADL]). There are currently a number of reasonably reliable and valid scales of functioning that have been developed for various rehabilitation settings (5,41). The FIM, for instance, has been widely studied, and its utility and basic validity for assessment of physical functioning are well established, as are its limitations in the areas of speech, language, and cognition (42). In assessing function, it can be important to realize that activities and disabilities are determined not only by impairments but also by the extent of compensating strengths. Moreover, disability and even indicators of participation may have a loose connection to life satisfaction (38,43).

Health Status Measures

Medical rehabilitation outcome measures may be considered to be a subcategory of health status and QOL measures. Books summarizing different scales of health-related QOL are now available (38,44–48). These sources are filled with scales relevant to rehabilitation outcomes assessment, though their sensitivity and logical applicability to medical rehabilitation require verification. Perhaps the most commonly used measure of general health is the short form 36 (SF-36) (49), although many other measures are also used (48). Some of the subscales within these instruments appear to be too broad or are otherwise not directed at problems treated in medical rehabilitation, but many dimensions—such as pain relief, general feelings of health and well-being, and physical function—are relevant.

QOL measures are relevant to rehabilitation outcomes assessment. Subjective or affective QOL is so important that it deserves assessment, despite its (nonqualified) exclusion from the World Health Organization (WHO) scheme. Subjective well-being and life satisfaction have been increasingly studied for use as ultimate rehabilitation outcomes measures (50). Subjective well-being is statistically associated with health and function, community participation, and a loving and satisfying social life, but the inconsistency of these associations demonstrates that subjective well-being cannot be reduced to indicators of objective health and circumstance (43,50): the person’s own expectations or implicit standards regarding his or her own life are critical. Chapter 12 discusses QOL assessment in greater detail. At this point, indices of patient well-being are sufficiently well validated to be used in research but have not yet been validated for use as routine indicators of the performance of individual rehabilitation programs.
Criteria for Choice of Measures

Criteria for choice of measures for performance and outcomes monitoring systems include relevance of content, reliability, other internal psychometric characteristics, or alternatively, biometric validity, and evidence of predictive validity, including evidence of utility in practice (10,13,51–53). While ease of administration and expense are critical considerations in practice, justifiable expense depends on the benefits of the measurement system.

Standards and Levels of Validity and Reliability

Scales employed in performance monitoring should meet recognized standards of reliability and validity in rehabilitation or other fields (10). While reliability and validity are commonly treated as catch phrases, they in fact subsume a set of interrelated criteria for the quality and utility of measurement. The validity of functional and QOL measures can be understood as a set of criteria for (a) sensible internal structure—that is, whether the set of items or procedures has the needed content, low intrinsic error, desirable internal psychometric characteristics and (b) desired external validity characteristics, including whether it generally “behaves” as it should according to one’s understanding of the construct, including convergent and divergent validity, and evidence of utility in practice, also known as consequential validity—that is, whether the measure leads to correct inferences and verifiable benefits, at least in its major application or use (52,54). If the construct to be measured and the main application are clearly specified, it is possible to grade the quality of measurement evidence (52,54,55).

One needs to know the reliability—that is, the stability, agreement, and reproducibility—of measures to interpret them. Without reliability information, one may not be able to distinguish between actual objective differences in scores and mere subjective or chance fluctuations. Error-prone indices of the appropriateness of medical care have been shown to overstate the frequency of inappropriate care (56).

Validity is a concept associated both with the construct being measured and with its application. Evaluation of a measurement procedure involves consideration of whether it has necessary internal characteristics (e.g., homogeneity, hierarchical structure) as well as external predictive characteristics, including validity for some purpose or construct. Accuracy is the relevant criterion to evaluate the validity of a measure when a true “gold standard” is available. Sensitivity is the probability of detecting a condition that a person actually has. Specificity is the probability that the test gives a negative result among people without the condition. When the question concerns whether an individual actually has a specified condition, given a positive result of a test, the needed statistic is positive predictive validity, which requires also knowledge of base rates (57,58).

When summing items to provide a meaningful summary number, one should know the degree to which the items are internally consistent, that is, additive and unidimensional (58–60). The FIM instrument, for instance, consists of at least two dimensions: motor ADLs and cognitive-psychosocial function (61).

Limited range of item difficulty has also been a problem with some scales used in rehabilitation, since rehabilitation deals with a great range of human performance—from coma or total paralysis through independent living and paid employment. Many existing scales are sensitive to the typical range of improvement seen in medical rehabilitation hospitals (37,42,62,63) but still have ceiling or floor problems, that is, they may be insensitive to very real improvements that occur in some patients who remain at a “total assist” level in ADLs or in individuals who are independent in ADLs but need speed, endurance, or higher level skills to sustain a productive lifestyle in the community (64).

To employ parametric analysis techniques (e.g., reporting means, t-tests, or Pearson correlations), the scales employed should have equal-interval characteristics. Measures developed using Rasch analysis have probabilistic equal-interval properties (58–60,65). They can identify and lessen floor and ceiling limitations of older measures, as can other forms of item response theory (IRT) (58,60). The method also has the advantage of identifying “misfitting” persons, that is, individuals whose pattern of functioning is so atypical that the conventional method of scoring their outcomes may be misleading. Different methods of scoring functional scales may be needed for different diagnostic groups. Walking, for instance, is relatively easy for a person with brain injury but is near impossible for a person with complete paraplegia; its significance as a marker of progress is radically different between the two persons. Chapter 11 presents additional criteria for choice of measures.

Sensitivity to Change and Evidence Bases

Sensitivity to change is a basic criterion for choice of outcome measures. This is true in the sense that outcome or performance measures unrelated to actual treatment objectives and valued outcomes should not be employed. At the same time, a measure can be too sensitive, so that improvement is of little value to patients or can fluctuate due to factors unrelated to treatment. With modern IRT and other metric analysis, it is possible to quantify the degree of sensitivity of a measure, that is, the degree of error of measurement. Previous controlled research provides a superior basis for choice of outcome measures; as such, research can identify attainable outcomes and linkages between outcome and needed treatment processes.

Severity Adjustment and Statistical Consideration

The need for severity or risk adjustment of performance data can hardly be overemphasized. QI and outcomes monitoring systems that are unadjusted or poorly adjusted for disease severity, functional limitations, and other factors that affect outcomes are likely to provide misleading reports. While all factors cannot be controlled statistically, outstanding confounding factors can be measured and their effect projected. Finally, knowledge of at least basic statistical principles is needed to interpret performance monitoring data. Sample size is always a consideration: a single bad outcome may well be a fluke; a pattern of outcomes below severity-adjusted norms indicates a possible process problem needing further investigation.
PROGRAM EVALUATION AND OUTCOMES MANAGEMENT

This section discusses systems of measurement, monitoring, and interpretation focused on the outcomes attained after care. We begin by discussing PE and associated schema in rehabilitation. Rehabilitation facilities—largely under the aegis of the CARF—now have several decades of experience with PE, and the resulting knowledge provides a basis for current program monitoring and clinical management activities.

Rehabilitation is provided in many settings, including transitional care facilities, nursing homes, outpatient clinics, homes, and hospital-based IRFs. Most of our examples will deal with IRFs and the most common outcome measure currently employed in IRFs—the FIM—in order to provide a focus and to limit length. Principles and concepts apply to other settings in which health-related rehabilitation is provided.

Program Evaluation

PE refers to a variety of information-gathering activities designed to aid in program development or functioning (i.e., formative evaluation) or to decide whether a program, as a whole, is worthwhile (i.e., summative evaluation). Many approaches to PE have been employed over the last three decades (66). “Performance monitoring” is a more current term that includes both PE and monitoring of key processes. Accountability to the public and internal management are overarching purposes regardless of rubric. These systems have multiple uses, including marketing, profitability, program planning and development, research, prognosis, utilization review, and improved clinical planning and treatment.

CARF and the Program Evaluation

Leaders in rehabilitation have long realized that the field needs to demonstrate its benefits to the public. Beginning in the 1970s, CARF assumed leadership, providing a forum that led to standards that required established rehabilitation facilities to develop PE systems that measure outcomes (67), implemented in numerous rehabilitation facilities over past decades.

PE has been described as “a systematic procedure for determining the effectiveness and efficiency with which results are achieved by persons served following services” (18). These results are collected on a “regular or continuous basis” for all patients or for a systematic sample of patients (18,68). PE and outcomes management involve setting goals and expectancies. If goals are not attained, reasons should be determined and action should be taken. In its usual form, PE does not provide answers to specific problem areas but merely identifies that a problem or strength exists. Answers are identified through more in-depth investigations involving further analyses of data, chart review, examination of quality measures or monitors, and discussions with the knowledgeable staff (18,21,69–72). PE systems are used to help make clinical management decisions and improve program operations.

Realizing the need for objective comparative data, most medical rehabilitation programs have joined large data systems. Accreditation standards state that organizations should compare their results and/or processes to some benchmarks, such as pooled data systems, the organization’s own larger network, and/or the published literature.

CARF has long emphasized meaningful, sustained outcomes in the real world after discharge. The goal is to maximize patient functioning and QOL in the community after discharge. Medical outcomes are noted when these may affect functional or general health outcomes. CARF standards ask that rehabilitation programs assess outcomes in terms of the WHO’s ICF (34) and emphasize the patient’s goals, desired activities, community participation, and satisfaction with services (18).

Experience with PE systems resulted in a shift of emphasis away from choice of measures and formal design of the PE system, and as early as the 1980s, use of the PE—not details of system design—became the key point. In the mid-1990s, CARF standards changed to use the terms “outcomes measurement and management” rather “PE” in order to emphasize the need for more operationally oriented approaches.

The Standard Rehabilitation Program Evaluation Model

In the 1970s and 1980s, medical rehabilitation programs developed their own tradition in PE (68). These PE systems were designed to provide an overview of program outcomes. In effect, they were designed to assure outcomes to the public, that is, to be summative evaluation systems. In operation, however, these systems functioned as formative evaluation systems (70). Information on outcomes is given primarily to program staff, who constitute the main audience for reports. Improved program management was, in fact, a primary expectation, leading to the relabeling as “outcomes management.”

Standard PE systems in rehabilitation have three components: design, goals and objectives, and reports. While this basic model is still widely used, updating has occurred as part of CARF’s strategic outcomes initiative to incorporate notions of outcomes management. CARF offers training, guidance, and materials on outcomes management. Anyone developing, implementing, or using a PE or QI system in rehabilitation should consult a CARF standards manual or Web site (www.carf.org) for references to the most recent information (18). Major components of a standard PE system are summarized below.

Program Purpose and Description

The PE design is based on a mission statement describing who the organization serves, what services it provides, and what goals it expects to accomplish. Goals should be anchored in the concerns of the persons served and other stakeholders—groups or entities with an interest in the success of the program. The special programs that constitute the organization are then described (e.g., stroke program, brain injury, spinal injury, pain program, general rehabilitation, independent
living center). Key influencers are listed to ground the statement in reality. These are external agencies that constrain and direct the rehabilitation program, such as the rehabilitation market and clients, referral sources, patients, staff, Medicare, third-party payers, and key government agencies.

Each program within an organization and the population it serves are to be described:

1. General program objectives. Defining a PE system requires defining program objectives. Foremost among these are anticipated results to the primary clients, but indicators of efficiency are also typically needed.

2. Admission criteria or definition of the population served in the program. Both inclusionary (e.g., cerebrovascular accident) and exclusionary (e.g., free from communicable disease, over 18 years of age, noncomatose, dependent in ADLs and ambulation, medically stable for 3 hours per day of therapy, likely to survive at least 6 months) criteria are defined.

3. Persons served, described with regard to diagnosis, functional issues and problems to be addressed, and relevant demographics.

4. Services provided or readily available to the patient, such as medical care (e.g., physiatry), physical therapy, occupational therapy, speech/language pathology, psychology, social services, nursing, or attendant care.

**General Program Objectives**

CARF standards require the measurement of program performance in the domains of effectiveness (results or outcomes for persons served), efficiency (relationship between outcomes and resources used), service access (e.g., number of days from referral to admission, convenience of the hours and location of operation), and satisfaction (experience of the persons served and other stakeholders) (18). Effectiveness, efficiency, and satisfaction have been in the CARF standards for at least three decades, with service access being added to reflect the challenging and dynamic aspects of today’s health care environment. Data elements to assess these domains are measured at admission, discharge, and follow-up, depending on the appropriate time for each data element. Outcomes are assessed after discharge. Follow-up data collection usually takes place 3 months after discharge but other periods can also be justified.

Also needed are progress objectives or intermediate outcomes in terms of patient improvement in the clinical setting toward outcomes such as improved independence in mobility, self-care, communication, or medical self-management. These are similar to (but less specific than) our concept of treatment objectives.

**Efficiency** objectives are also needed. Resources consumed such as staff time, LOS, number of treatment sessions, and dollars should be monitored and related to the results achieved. For example, the functional gain for a given LOS can be monitored to ensure that outcome is not sacrificed with resource restriction.

**Program Evaluation and Outcomes Monitoring Systems in Rehabilitation**

For several decades, rehabilitation programs have employed a model of PE. This model is only one of several alternatives; textbooks provide lessons in the variety of approaches and issues encountered in PE more generally (66,73). We describe this standard or classic model because examples of it are comparatively well-defined and tested and because lessons from it provide the basis for current and future performance monitoring systems.

Typical constituents of a rehabilitation inpatient hospital PE system are shown in Figure 12-2. The sparseness of measures (italicized) in the process box and the larger set of admission (i.e., input) and outcomes (i.e., discharge and follow-up) measures show the emphasis of conventional PE systems. Scales of independence in ADLs such as the FIM constitute the primary input (e.g., admission, baseline) and output (e.g., discharge, follow-up) measures. Cost and LOS are classified here as process or input measures because they indicate the degree of effort or resource use devoted to benefiting the patient. PE systems also address quality of routine nursing care, hotel services, and patient satisfaction (74,75).

![Figure 12-2](https://example.com/figure122.png)

**FIGURE 12-2.** Basic conventional PE framework for rehabilitation programs. Items from the IRF-PAI are shown in italics.
The FIM instrument is the most commonly used functional outcomes measure in inpatient medical rehabilitation. It is an 18-item scale that rates each item on a scale ranging from 1 (total assist) to 7 (completely independent). The FIM consists of two overall factors (motor function and cognition) and recent reports indicate acceptable-to-good reliability (65). It became the basis for the PPS for medical rehabilitation hospitals in the United States beginning January 1, 2002.

The most widely used set of rehabilitation performance indicators at present is found in the IRF-PAI data set. This data set contains information on impairment group, FIM at admission and discharge, demographic information, and LOS. General purpose data sets categorize patients so that they can be grouped by estimate expense for the PPS. For purposes of PE or clinical performance monitoring, however, the information system needs to be tailored to diagnostic and functional groups.

PE systems also need supplementary measures used for general descriptive or comparative purposes (see Fig. 12-2). Demographic variables (e.g., age, gender, race) are needed as input or independent variables. Although they typically are not good measures of case severity, demographic variables do help segment the population for other analyses (e.g., access to care, service type).

Data on and reasons for rehospitalization and death are also essential supplementary measures in PE in medical rehabilitation, which treats aged, infirmed, and chronically ill patients. Although the main purpose of medical rehabilitation is not to decrease mortality, certain medical rehabilitation programs have been shown to substantially increase survival (76).

Even though accreditation standards have allowed completely local measures and standards, the flimsiness of completely local, subjective expectancies has been recognized. Acknowledging this, rehabilitation programs have voluntarily created regional and national outcomes data systems such as the Uniform Data System for Medical Rehabilitation, eRehabData, ITHealthTrack, and other firms. The use of normative benchmarks is highly valuable to performance monitoring but provides a challenge when the available benchmark data do not correspond exactly to the program’s objectives or population.

**Needed Specification and Additional Design Points**

Defining a useful PE system requires forethought, including

- Specification of whom measures are applied to. While traditional program objectives were applied to all patients in the program, newer approaches recognize that important and expected outcomes vary across groups. CARF standards after 1998 require analysis of outcomes in meaningful groupings rather than in all patients.
- Specification of how measures are implemented and when they are applied. Most programs measure function at admission and discharge. Assessment of function 1 to 6 months after discharge gives a more valuable picture of patient outcomes. Follow-up of outcomes has become common and is required by CARF standards. The person who does the measurement should also be specified.
- Specification of expectancies—specific statements of the expected level or range for objective performance indicators. The classic PE model involves specifying a range of performance expectancies: minimal, optimal, and the maximal, thought to be attainable under ideal circumstances. Outcomes were not to fall below the minimum. If they did, action was to be taken (67,68). Expectancies are commonly based on a combination of internal trends and targets, and if known, regional or national norms.

Consideration of the relative importance of objectives. In the traditional PE model, program success was to be summarized in a single number. Actual objectives attained were multiplied by weights and expectancies chosen so that optimal attainment of outcome was signified by 100. The weighting system is no longer required, but the concept of weighting outcomes can still be useful.

Additional points for design of outcomes monitoring systems are as follows:

- Cases that stay only a few days are not comparable to full-stay cases and need to be looked at as a separate group.
- Long-stay outliers also need to be examined.
- Outcomes monitoring systems center on episodes of illness rather than on administratively convenient units such as a stay in rehabilitation. Readmissions need to be collapsed or analyzed separately. Efficiency cannot be achieved by cycling difficult cases back and forth between facilities.
- Some rehabilitation programs distinguish between cases admitted for different reasons. Some patients, for instance, are admitted largely for care of certain medical-nursing problems that rehabilitation hospitals are particularly adept at treating (e.g., decubitus ulcers, urinary tract infections, weaning a patient from a ventilator). Incorporating measures relevant to the reasons for admission and for rehabilitative treatment enhances the meaningfulness of outcomes monitoring reports.

**Outcome Monitoring Models for Different Populations**

Patient populations need to be divided into major groups, usually by etiology or impairment group and functional severity. References are available on how to tailor a PE system for

- General inpatient medical rehabilitation, including stroke (37,63,77)
- Spinal cord injury (37,63,78)
- Traumatic brain injury (38,79–81)
- Chronic pain management programs (82)
- Outpatient rehabilitation clinics (67)
- Postacute community reentry (63,81,83) and vocational programs (63,81)
- Other conditions requiring rehabilitation (63,84).

Because inpatient rehabilitation programs must contend with numerous mixed-diagnosis cases, comorbidities, and rare diagnoses, mixed-diagnosis evaluation systems are a necessity if outcomes (and processes) are to be monitored for all patients. Functional improvement is a meaningful, if imperfect, way of quantifying the
benefits in mixed-diagnosis groups. Mixed-diagnosis systems that focus on functional and participation-level outcomes appear to be relatively successful for later stages of rehabilitation, including transitional living, community integration, vocational rehabilitation, and long-term nursing home care. Both function and diagnosis are critical in evaluation of processes and outcomes of inpatient, outpatient, and at-home medical rehabilitation programs.

**Outcomes Measurement**

This section examines issues distinctive to the measurement of outcomes for rehabilitation PE.

**Generality of Measures**

PE goals have been designed to credit the program with the larger benefits it produces, such as independence from assistance (68). These goals are more general than treatment or case management objectives. Measures of long-term outcomes in the community are valuable for marketing and are ultimately needed for policy and accountability (69,70). Data on how a program has reduced the frequency with which patients are institutionalized in nursing homes and hospitals after discharge, for instance, are meaningful and even influential with boards of directors, government officials, insurers, families, and referral sources. While reports of such benefits are useful in communication of the benefits of rehabilitation to the public, more proximal outcome measures are usually more closely related to interventions and hence are more likely to be related to action to improve clinical processes.

**Performance Versus Ability**

The standard and usual practice is that primary outcomes are measured in terms of actual patient performance, preferably measured in the community after discharge, rather than in terms of capability demonstrated or judged in the clinic (63,68,84). This is because actual performance is usually a more reliable and objective measure than judged ability, and activities in the community are more meaningful and valued than activities in the clinic. Abilities that are used in practice prima facie provide greater benefit than those used in artificial situations. Exceptions exist when dealing with performance capabilities that are important though infrequently needed (e.g., safety skills) or if there is evidence that the clinical performance has high validity as a proxy for real-world outcomes.

**Timing of Outcomes Measurement: Follow-up**

Outcomes for persons served are best measured following discharge (68). Measurement at discharge is less expensive but may be less informative, as clinical staff are already aware of patient function at discharge. Information on durability of outcomes is valuable. Patterns of under-preparation, or of long stays by patients otherwise ready for discharge, should be actionable. Whatever time is chosen, data need to be obtained from all persons served or from a representative sampling (18).

There is no perfect time for follow-up, as there are contrasting advantages to both short-term and long-term follow-up. Three months has been the most common period for follow-up of rehabilitation outcomes, but periods of 1 to 6 months after discharge are also found. Rehabilitation involves enhancing healing and adaptation processes, so recovery processes should ideally be measured repeatedly over time.

Outcomes after discharge are usually assessed by telephone calls or clinic visits. PE systems in rehabilitation have long employed telephone follow-up. A great deal of research has shown that telephone follow-up using structured questionnaires of demonstrated reliability and validity (10) provides a good balance of reliability, low rate of missing data, and modest-to-moderate costs. The number of self-report scales for assessing health and function with basic knowledge of reliability and validity is now large (37,42,44–47,62,85,86). In-clinic follow-up methods are required to objectively assess medical problems. Missing data, however, can be a problem if patients do not return for their follow-up visit in the clinic. Tele-rehabilitation technologies may improve our capacity to provide objective patient assessment following inpatient discharge.

Practical difficulties of follow-up include its expense, funding restrictions on continuing outpatient care, a lack of payment for educational or evaluative follow-up, and the fact that continuing outpatient care may involve a different provider than inpatient care. Nonetheless, rehabilitation programs can be improved by ongoing knowledge of whether new, unexpected problems or complications arise after discharge, and if so, to whom and why. Monitoring of long-term outcomes is also needed to assess whether changes in health care designed to control costs have compromised the health or functioning of patients undergoing rehabilitation.

**Benchmarking Functional Outcomes**

The availability of benchmarks or standards of comparison is basic to systematic QA and QI. While they may be obtained from many sources, including the published literature, contemporary benchmarks are most commonly obtained from shared data systems that pool data from a number of facilities. The typical outcomes benchmark in rehabilitation has been average functional outcome or gain for major diagnostic groups. Accurate adjustment for case mix and severity is essential for meaningful comparison of raw quality and outcome indicators across patient groups and programs.

**Severity Adjustment for Functional Outcomes**

To compare a program’s outcome or improvement scores to a benchmark, one should examine major factors that drive these scores. There are a number of factors that generically affect functional outcomes across many diagnostic groups in rehabilitation (25,37,60,87):

1. Functional severity at admission. Improvement may not be equally likely or meaningful across all levels of an admission measure. Some studies have reported curvilinear relationships, that is, greater improvement among patients admitted at intermediate levels of severity (25,87).
2. Chronicity (i.e., onset-admission interval). After the acute phase of many severe injuries, there is a period of relatively rapid recovery, followed by increasingly slow improvement and eventual asymptote, at least on a group basis. Control for natural history recovery curves is needed.

3. LOS. Improvement in rehabilitation tends to be correlated with LOS.

4. Differences in comorbidities and severity of illness or injury (38). Differences in improvement across facilities may be due to differences in medical-nursing severity or case mix. Diagnostic complexity and comorbid conditions adversely affect outcomes and increase LOS in rehabilitation (88). Further development of indices and models of such factors is needed to identify patients with high medical-nursing needs and to establish clinically useful performance benchmarks for them.

Longitudinal research has identified relatively powerful outcome predictors within diagnostic groups. General severity of disease or impairment is typically a major predictor (e.g., severity of spinal paralysis and American Spinal Cord Injury Association [ASIS] motor scores in spinal cord injury [SCI] (38), Glasgow Coma Scale and duration of unconsciousness or posttraumatic amnesia for traumatic brain injury [TBI] (42), severity of paralysis as measured by Fugl-Meyer Motor Scores in stroke (89)). Premorbid factors can be powerful predictors of long-term community outcomes after rehabilitation, even more powerful than severity of injury (90). A great deal of research has been done on predictors of outcome following rehabilitation, and this research is applicable to quality outcomes improvement.

There are several methods of case mix or severity adjustment for medical rehabilitation (91). As methods of risk or severity adjustment, all these are approximate and typically predict a minority of the variance of LOS or functional gain. Rankings of acute hospital outcomes are sensitive to the method of adjustment employed (92). One would expect similar results for rankings of rehabilitation hospitals by functional gain.

**Function-Related and Diagnostic Groups for Prospective Payment**

Function-related groups (FRGs based on the FIM) were developed to adjust inpatient medical rehabilitation caseload for case-mix factors affecting LOS (93). Relabeled case-mix groups (CMGs) are now used as a basis for the Medicare’s PPS for patients admitted to inpatient rehabilitation programs in the United States. CMGs group patients based primarily upon admissions FIM and impairment group. Average LOS can be projected. FIM-FRGs predict about 31% of the variance of LOS in rehabilitation, which is similar to the performance of diagnosis-related groups (DRGs) for acute hospital LOS. FRGs and CMGs are more detailed than previous PE systems that reported by broad etiologies. Strokes, for instance, were grouped into multiple diagnostic-functional subgroups (94). FRGs classify rehabilitation patients into groups that are more clinically homogeneous and interpretable than groupings by primary diagnosis alone. FIM-FRGs have been used to investigate the “efficiency” of rehabilitation, that is, the relationship of functional gain to cost or LOS (95). The main use of FRGs/CMGs is as case-mix adjusters to identify groups whose costs are higher or lower than expected. They are used to identify patients whose LOS exceeds the average for the FRG. They are, however, potentially applicable to analysis of efficiency and QI in rehabilitation, defining patient groups whose gains in function are unexpectedly low given LOS (94,95).

**Functional Gain as an Indicator of Quality**

It was once thought that functional gain would provide a robust indicator of the quality of rehabilitation programs. While greater gain in function is undoubtedly desirable, research connecting functional gain of actual ongoing rehabilitation programs to indicators of care processes or program characteristics is scarce. Recent, relatively large studies have failed to find an appreciable correlation between staffing intensity and other characteristics of inpatient rehabilitation programs and severity-adjusted functional gain (87). Functional outcomes and LOS, however, are relatively predictable, and managed care clearly constrained LOS in rehabilitation hospitals. “Relationships between rehabilitation practices and functional gains by patients do not appear to be either simple or overt” (87). With continued research, one may expect that reliable connections will be identified between characteristics of certain kinds of rehabilitation programming and certain severity-adjusted outcomes for selected patient groups.

**Measurement and Statistics: Summary**

Medical rehabilitation has reached agreement on basic typical domains for inpatient programs (e.g., mobility and self-care ADLs in the FIM), but measures of other critical domains still have to be developed or agreed upon (e.g., measures of treatment objectives clearly linked to therapies prescribed, extended or instrumental ADLs, ecologically valid measures of communicative and cognitive outcomes, patient satisfaction, and family and other environmental factors) (10,38,42,44). Methods of statistical control for severity of disease, comorbid conditions, and environmental factors that affect outcomes have been inadequately developed.

**Outcomes Management**

The term “outcomes management” has become increasingly popular. The term is loosely associated with outcomes measurement, PE, case management, and managed care (84). Paul Ellwood provided the original conceptualization of outcomes management as “a technology of patient experience designed to help patients, payers, and providers make rational medical care–related choices based on better insight into the effect of these choices on the patient’s life” (96). Outcomes management is based on the increasingly scientific basis of medical care, including the increasing ability to predict outcomes, and on advances in measurement of health and function at the level of patient experience rather than at the level of mortality.
or disease rates. Professional analyses of huge databases were to provide estimates of the effectiveness and efficiency of medical services in practice. Rather general outcomes measures were to be used, so implications would be primarily at the aggregate level—for systems of intervention or programs rather than for management of individual patients.

Outcomes management has also been used to refer to systems that manage individual patients. We define a clinical outcomes management system as one that involves routine monitoring of the treatment objectives for individual patients and indicators of patient responsiveness to treatments. Clinical outcomes management, as defined here, differs from traditional PE in that standard goals are not routinely applied across a group but are modified to fit the individual. Treatment objectives and outcomes need to be risk adjusted, necessitating a computerized system. Objective measures are possible for important, high-frequency objectives and outcomes.

CARF’s Standards Manual treats information management and service delivery improvement—also called outcomes management—in general terms:

CARF-accredited organizations are proactive. The organization continually collects data from a variety of internal and external sources. These data are analyzed and the results are used to make decisions (18).

Human service organizations exist to make a difference in the life of the person served. Evidence that an organization does make a difference comes from information about the outcomes achieved and satisfaction with service delivery from a variety of perspectives. This collection of information is based on the establishment of a level of performance to be achieved prior to measure indicators (18).

CARF has fostered discussion toward indicators for outcomes management (97) but has not specified details (18).

Outcomes management in medical rehabilitation involves four techniques:

1. The use of treatment guidelines (or standards) to help clinical professionals to evaluate patients and choose appropriate treatments.
2. Routine and systematic measurement of both indicators of disease and of patients’ functioning, health, and well-being and of changes in these related to the likely effects of treatments.
3. Combining data on inputs (severity), processes, and outcomes into large databases to permit scientific analyses.
4. Analysis and dissemination of results in a form useful to different stakeholders.

Although U.S. health care as a whole is far from operationalizing Ellwood’s grand vision for outcomes management, elements of it are being implemented on a piecemeal basis as integrated health organizations create their own clinical databases and smaller organizations voluntarily join health outcomes databases. Rehabilitation, with its long experience with outcomes-oriented PE systems, should be ready to operationalize outcomes management.

**CARF Performance Indicators**

CARF has historically been oriented toward outcomes, but CARF standards increasingly use the term “performance indicator.” CARF has worked to develop performance indicators for rehabilitation programs (97), attempting to meet the concerns of the major stakeholders to the rehabilitation process. Core questions have been addressed: What does a stakeholder want to know about a program’s performance in order to assess its quality and choose among programs? How should these concerns be quantified? Key concerns and draft indicators have been identified for further development.

Many of the proposed indicators are outcomes oriented (97). Outcome concerns include percentage of clients reporting improvement after service provision; durability of outcomes; functional gain; increased productivity, participation, and activities; reduced impairment and disability; satisfaction with processes of care; satisfaction with results of care; QOL after care; efficiency concerns; reduced downstream costs; and value (outcome relative to cost). Examples of possible outcome indicators include:

- For “durability of outcomes”—percent of persons in a program who maintain or increase level of motor function from discharge to follow-up or percent of persons who have not been rehospitalized for the same impairment within 6 months of discharge.
- For “satisfaction with results of care”—percent of persons with mobility impairments who respond “satisfied” or “very satisfied” to the question “How satisfied are you with the amount you are able to move around outside following your rehabilitation?”

Structure and process indicators are also proposed:

- Structure concerns—types and numbers of persons served, staff credentials and education, legal status of organization, accreditation status.
- Process concerns—cost and resource use, communication effectiveness, involvement of persons served in rehab decisions, collaboration of providers with payers.

Specific performance indicators and data set are not yet required by CARF standards, and CARF does not endorse or specify any particular measure, tool, or vendor for outcomes management or PE. Indicators or measures employed should be implemented in such a way that stakeholders can know they are looking at comparable information from different programs (97). Though still incomplete, CARF publications and standards continue to serve as a guide for the kinds of indicators needed.

**Critique of Program Evaluation and Outcomes Management**

The standard PE model has a number of strengths. It provides an overview of primary patient outcomes, progress, and cost. If used with a shared national or regional data system, standard PE systems provide an index or benchmark of the
effectiveness of the program in improving patient function and placing patients in community settings. Efficiency, or at least an operational utilization review system, is demonstrated if the facility shows a direct correlation between cost or LOS and improvement (25,67) and if improvement/day rates are similar to those in other rehabilitation facilities for similar diagnostic-functional groups. PE data have numerous administrative and clinical uses (84). PE systems have begun to tell us whether rehabilitation programs attain an outcome for their patients.

A drawback of the classic PE system has been that when performance expectancies were not met, a common response was simply to change the expectancy. Without accurate statistical methods of adjustment for case mix and severity, outcome comparisons can be misleading. Without transparency—the ability of outside scientists or professionals to check results and inference—the integrity of results cannot be verified. Without public release, motivation for the system is weakened. In sum, the evaluation and especially the link to action have been weak in PE model, and relabeling such as “outcomes management” has not cured the problem. Nonetheless, the process has taught rehabilitation much about clinical data systems and realistic outcome expectations.

QUALITY ASSESSMENT AND IMPROVEMENT

To ensure quality in medical care, it should meet standards that are in some sense predefined (20,98,99). Although efforts to systematically ensure quality in medical care go back to the first quarter of the 20th century, pressure for accountability has increased in recent decades, driven by explosive growth of costs and by higher expectations of medical care (98). The federal government and the Joint Commission have been major institutional forces behind quality assessment for hospital care, and historically, CARF has played an important role in defining quality in rehabilitation (18). This section will begin by discussing total quality management (TQM) and CQI techniques and their application to medical rehabilitation. Discussion of QI terminology and then Joint Commission accreditation will follow.

Total Quality Management and Continuous Quality Improvement in Health Care

An important insight of TQM and CQI (22,23), based on experience, is this: quality and effectiveness primarily depend on the routine system. The root causes of problems are more commonly at the level of the system or the sequences of care processes than at the individual or even department level. The causes of error or undesirable variation in the sequence of activities must be identified and rooted out. The aim is to improve systems not blame individuals. Improved protocols for activities and processes need to be developed and implemented as a key element of QI (74,100). Global organizational commitment is the dominant requirement (22,24,74). The philosophy has moved health care toward improving routine processes.

In practice, TQM and CQI emphasize knowledge of effective processes and involvement of the staff directly involved in the process. They involve fact finding, emphasize prevention of problems, and use measures of processes or of shorter- or longer-term results, depending on the problem. In Deming’s terms, “profound knowledge”—detailed, expert, first-hand understanding—of what is really happening in the organization and of the complex processes involved in producing a product is required to improve quality; knowledge of “general variation” (i.e., statistics and scientific measures) alone is insufficient (22,23,101). The emphasis is on understanding the total system and involvement of everyone to diagnose, plan, and fix problems or improve systems. Both the specific problem and systems in which it is embedded need attention. CQI emphasizes review of systems and sequences rather than discrete inspections (21). One must acquire extensive knowledge about the system, not just identify errors or outliers as in traditional QA. When variations exceed normally observed limits, knowledge of the system is needed to infer the cause and mend problematic processes.

To improve medical care, knowledge of diagnosis and treatment alternatives as well as an evidence basis is required. Clinical practice guidelines are, if well-developed and based on evidence, invaluable tools for QI, so valuable that we treat them separately in subsequent sections. Guidelines need to be integrated with and adapted to actual team processes to assure and improve the quality and effectiveness of care.

The superior effectiveness of improving routine processes, compared to simply trying to eliminate the worse problems or the worst performers, is graphically displayed in Figure 12-3. This conventional display assumes that measured quality or results are distributed normally. An approach aimed at eliminating unacceptably poor care would, if successful, eliminate poor care for only a small fraction of patients (the small left tail of the distribution). An approach aimed at improving the process of care and eliminating inappropriate variations in the process (see Fig. 12-3) would improve results for most patients. As a by-product, the fraction of results or care that is clearly below the old threshold is also greatly diminished.

As with CQI in other industries, the majority of problems, especially remediable ones, in medical settings are usually problems with systems or procedures rather than with the incompetence of individuals. Many works give examples of systems problems (e.g., nursing care or hospital pharmacy) (1,21,102). By contrast, the claims history of individual physicians, for instance, only weakly predicts future claims proneness so use of such data to target individual physicians is problematic (103). Malpractice claims data have been used to identify problem-prone clinical processes and to suggest improvements to reduce the likelihood of negligence (30).

Another insight is that QI systems that depend on mass inspection of discrepancies in outcomes are often ineffective or inefficient (21,74). If QI simply counts errors and points them out to staff, QI can be perceived as an unpleasant policing activity, and the substantial effort to detect outcome anomalies may not be paralleled by efforts to improve production or treatment processes. The problem is not that outcome measures are wrong
but how they are used: data should function as scoreboards for team efforts to improve their own processes (23). Multiple statistical approaches are needed, including process and outcome measures to both identify defects and verify improvements. CQI integrates knowledge, processes, at least short-term outcomes, and action to improve them.

The literature now reports many examples of successful system improvement in health care organizations, especially hospitals. Limited quantitative evidence of effectiveness exists (104), but CQI and TQM are difficult to test as they transcend any particular setting or method. TQM has often been implemented by managers and applied to administrative organization so that clinicians have difficulty seeing its use to improve clinical processes. CQI/TQM teams led by clinical professionals may have a different experience.

Experience has shown that approaches to quality and outcomes improvement from manufacturing or provision of hotel services need major modifications to be applied to health care. Patients are not uniform material input to a manufacturing process, and customer satisfaction is not the only, or even the primary, relevant outcome in health care. Rather than being assumed or predicted on the basis of processes, patient responses to treatment must be tested in controlled research and monitored in practice. Comorbid conditions and idiosyncratic patient characteristics and needs alter ordinary patterns of treatment and response. "Context" has fundamental effects on activity therapies in rehabilitation (105). The principle of reducing variance in processes is relevant to parts of rehabilitation, but quality rehabilitation also involves the tailoring of treatments to the priority needs of the individual.

Professional Quality Improvement Terminology

A few basic terms assist in professional assessment and improvement (98,106). Norms are measures of actual clinical practice. Examples are average LOS, average improvement in FIM scores, and average hours of physical therapy. Norms are most clinically useful when they are specific to a patient diagnosis or otherwise graded to patient characteristics. Benchmarks apply to processes as well as outcomes. Rehabilitation professionals need to have benchmarks against which to compare their staffing, education, costs, initial evaluation, intervention types and intensities, patient satisfaction, and short-term outcomes as well as long-term ones. Norms and benchmarks have greater authority when they are based on large samples or when they tell us what is done and achievable by the "best" or at least better programs.

Criteria are statements that define appropriate or correct clinical care (98,107). Criteria are typically developed on the basis of professional experience and scientific literature. Some distinguish between a criterion and a standard (108), using the former as the more general dimension and the latter as the specific numeric cut-point. We will not rigidly distinguish the two, because a general dimension separate from a quantitative decision point is of little use (98). For instance, the statement that "stroke patients will have a blood level of Coumadin in the therapeutic range" is useless without specification of what the range is (e.g., prothrombin time of 1.2 to 1.5× control). Another example of a criterion or standard is the assertion that inpatients in medical rehabilitation should receive 3 hours per day of combined therapy (physical, occupational, and speech and language pathology). Criteria and standards may describe structure, process, or outcome, and in practice involve all three.

Sentinel Events

In practice, action to maintain quality of care frequently depends on sentinel events (109)—single occurrences that are highly problematic or socially unacceptable. Litigation following patient injury, staff quitting over unacceptable quality or ethical issues, and cockroaches on the walls are not definitive evidence of global quality problems, but they should motivate a review to determine whether there are remediable problems. Sentinel events require a response. The point of systematic QI is to go beyond concern for negative outliers alone.

An indicator condition is a frequent, treatable clinical situation (98,101). In Joint Commission terms, an indicator is a specific instrument to measure an aspect of care to guide the assessment of performance (110). Clinical indicators point to clinical processes or procedures that need further analysis to determine if improvements can be made. Improved clinical procedures should lead to improved outcomes.

A threshold indicates a preestablished point in an indicator that should trigger more in-depth investigation to determine whether a problem or opportunity to improve care...
exists (98). As an example, a threshold of 5% might be set for rehabilitation patients discharged back to acute care and 15% to a nursing home. In the past, thresholds have been either rather arbitrary or set by expert judgment; some have suggested statistical criteria (98,111,112). The Joint Commission has had difficulty in setting and implementing thresholds. Benchmarks based on both regional and local experience are needed to set thresholds.

There are situations where a 0% or 100% threshold is needed (98,113). For sentinel events, such as death or suicide within rehabilitation or within 7 days of discharge, a threshold of 0% would be justifiable: every case needs to be individually reviewed. In general, however, thresholds of 100% success or 0% problems are unrealistic. Setting thresholds at less-than-perfect levels avoids disproportionate use of time to evaluate a few discrepant cases that will probably be found to be clinically justified (98,114). QI in rehabilitation usually requires discrimination and amelioration of frequently occurring or significant problems, not undiscriminating compulsiveness.

The term “monitor” has been commonly used to describe any routinely collected measure on a group of patients. Staff engaged in activities to improve or oversee care monitor aspects of these processes or their outcomes. Because formal indicators have yet to be validated, QI efforts for rehabilitation must be undertaken with the use of ad hoc monitors. Specific Joint Commission indicators are discussed below.

Statistical Issues
Statistical control principles are as relevant to quality monitoring as outcomes monitoring. Sample size needs to be specified to set a threshold in QI. A 20% rate of apparent error with five patients is very different from the same rate with 50 patients. Patient groups for which indicators are applied must be well-defined. Interpretation of process and intermediate outcomes data is greatly facilitated by severity adjustment, just as with monitoring of long-term outcomes.

Joint Commission Approaches and Standards
Joint Commission standards have evolved over the years, and there are lessons in this evolution. Structural and process indicators of quality were first propounded. Care had to be provided by licensed practitioners with certain staffing patterns and authority relationships. Extensive record keeping, facility, and equipment standards were prescribed. Although certain structural indicators have been retained today, emphasis has shifted. The aim now is to improve the actual provision of high-quality, effective care, not just to assure the capacity to do so.

In past decades (1950s through 1970s), great reliance was placed on peer review, methodologies for which became increasingly elaborate and focused. Although peer review methods remain useful in certain circumstances, the method came to face increasing criticism. There was little evidence that it improved the actual process or effectiveness of care provided, although it did improve medical records. Current health data systems also face the challenge of proving that they improve care and outcomes rather than record keeping.

Ideas of objective indicators or monitoring, introduced into Joint Commission standards during the 1980s, are now increasingly operationalized in terms of systems of standardized severity-adjusted, objective performance indicators (98), arbitrarily labeled ORYX. Only implicit standards and thresholds exist for many conditions in rehabilitation, but the movement is clearly and strongly toward objective ones.

Prior to 1992, standards were organized around departments and services, distinguished processes and outcomes, and defined specific leadership responsibilities. Influenced by TQM and CQI, the departmental focus was de-emphasized, and the chapter on “Improving Organizational Performance” treats both quality and outcomes issues. “Quality assessment and improvement” was replaced by “QA” in the 1992 Accreditation Manual for Hospitals (AMH). Joint Commission now emphasizes continuous improvement of organizational performance, the identification of processes that most significantly impact care effectiveness or patient outcomes and their improvement, and integration of systems and processes across departments and functions.

Current Standards
“Helping Health Care Organizations Help Patients” is the Joint Commission’s new tagline, encapsulating a commitment to the public and to organizations that it accredits (115). It also reflects ongoing efforts to enhance the value of accreditation and its utility. Joint Commission standards today emphasize care processes as well as structure. Objective indicators of care processes and linked indicators of care effectiveness (outcomes) are increasingly required. Current Joint Commission standards relate to care and assessment of patients, patient education, consistency across the continuum of care, environment of care, management of human resources (staff competency assessment and training), management of information, infection control, patient rights and organizational ethics, leadership, governance, nursing care, medical care, and improving organizational performance (20). The Joint Commission evaluates quality of care in terms of the degree to which health services for individuals and populations increase the likelihood of desired health outcomes and are consistent with current professional knowledge (115). More specifically, they look at what is done, whether care is appropriate and efficacious for the patient, and how well it is done—the degree to which care is available in a timely manner, is safe and effective, is respectful and caring, and is continuous with other care.

The Joint Commission requires that “the leaders establish a planned, systematic, and organization-wide approach(es) to performance improvement” (115). The performance improvement plan is to be based on the organization’s mission, vision, and values. The plan is to reduce variation in processes and outcomes, increase patient satisfaction, decrease or control the cost per patient, and increase the effectiveness of information management. Performance improvement efforts should be prioritized according to (a) expected impact on performance; (b) high-risk, high-volume, or problem-prone processes; (c) relationship of potential improvement to Joint Commission dimensions; and (d) organizational resources.
There are several methods that health care organizations can use to implement performance improvement. These methods include a ten-step method outlined by Joint Commission in the past and the Deming cycle or the plan-do-check-act (PDCA) methodology (116). The focus PDCA method specifies (F) finding a process to improve, (O) organizing a team, (C) clarifying the current knowledge, and (U) uncovering root cause of variation, before (S) starting the PDCA cycle. These methods include the following (20):

- **Plan.** Planning is multidisciplinary, includes input from relevant stakeholders, and includes an understanding of the current process and outcomes.
- **Implement.** Implementation involves developing potential solutions, benchmarking best practices, and pilot testing new processes.
- **Check.** New processes and processes that involve risks must be monitored and measured. To measure performance, a hospital collects data on processes; outcomes; a comprehensive set of performance measures (indicators); high-risk, high-volume, and problem-prone processes; and other sensors of performance.
- **Assessment.** Assessment is defined as transforming data into information by analyzing it (20). Benchmark data, trends over time, and adherence to regulatory requirements are all useful in determining whether a process improvement has been successful. Appropriate statistical quality control techniques are to be used.
- **Improve.** Once improvements have been realized, process changes need to be standardized and gains maintained.

The Joint Commission has defined general elements of inpatient rehabilitation (e.g., a requirement for assessment of “functional rehabilitation status”) (standard tx.6 (20)). Detailed standards and indicators for inpatient or outpatient medical rehabilitation programs have not been established at the time of this writing, but standards and indicators for related areas of health care may be informative (e.g., standards manuals for home medical equipment, respiratory therapy and rehabilitation technology (117), ORYX indicators for home care (118)).

**Data Collection Recommendations and Standards**

The Joint Commission considers measurement and data collection to be the foundation for performance improvement activities (20). The organization’s leadership is responsible for establishing an information system to monitor quality-related events. The Joint Commission requires that

1. Organizations collect data to monitor performance (20).
2. Data are systematically aggregated and analyzed.
3. Information from data analysis is used to make changes that improve performance and patient safety and reduce the risk of sentinel events (115, p. PI-7-9).

Arbitrary thresholds and standards particular to a program are not enough. Benchmark values may come from an external multifacility database, the published research literature, and other sources.

To maximize the impact of limited resources, QI and data collection efforts should focus on high-volume, high-risk, and problem-prone processes. Additional foci include

- Patient outcomes (e.g., functional improvement in rehabilitation)
- Targeted areas of study (e.g., a new or redesigned process)
- Comprehensive performance measures
- Client needs, expectations, and feedback (e.g., patient satisfaction)
- Infection control measures (e.g., urinary tract infection rates)
- Safety of the environment (e.g., hazard surveillance monitoring)
- Quality control and risk management indicators (e.g., medication incidents, patient falls)

The Joint Commission has developed more stringent patient safety standards in recent years. Prevention of errors is a major focus. Effective in 2002, health care organizations were mandated to perform at least two Failure Mode and Effect Analyses (FMEA) annually (20). An FMEA analysis is similar to a root cause analysis but is proactive rather than reactive. For example, rather than reacting to a sentinel event, an FMEA analysis is performed before a negative outcome occurs. The organization first identifies a high-risk process (e.g., maintaining security of medication carts) or population (e.g., admitting and monitoring ventilator-dependent patients). “Sentinel Event Alerts,” published by JCAHO, may be used to identify patients or processes at high risk. The next potential “failure modes” or hazards are identified, and the process is redesigned to minimize potential risk. “The 2008 National Patient Safety Goals have been expanded to address early recognition and response to significant changes in a patient’s condition, as well as the complex challenge of managing anticoagulation therapy” (115, p. FW-1).

**Performance Indicators**

The Joint Commission has devoted considerable attention to developing and identifying appropriate performance measurement sets or systems. Difficulties have been encountered, and development continues after more than a decade of work. Priority was given to high-volume, high-risk, or problematic clinical practices. Most performance indicators developed thus far are short-term clinical indicators for acute medical conditions rather than long-term outcomes, structure, or processes measures. Indicators developed have been subsumed into what is now called the ORYX initiative (described below). Health care organizations are required to send data to an approved data organization on a continuing basis.

The Joint Commission’s *National Library of Healthcare Indicators* describes over 200 measures of clinical conditions, functional health status, or satisfaction in a standard format (110). Key hospital-wide processes—infecition control and prevention of medication errors—have been a special focus. The Agency for Health Care Research and Quality now operates the National Quality Measures Clearinghouse.
and validation of indicators, still underway, should ultimately enable accreditation decisions to be based more directly on the actual performance of the health care organizations.

**Indicators for Rehabilitation**

Joint Commission standards describe rehabilitation outcomes in general terms as involving “improvement of functioning, self-care, self-responsibility, independence, and QOL” (20, p. cx.7). Actual rehabilitation-specific indicators, however, are still under development. Many of the indicators discussed above, such as satisfaction with care, infection control, and medication monitoring, are also appropriate to medical rehabilitation (120,121). Quality indicators for nursing practice are becoming increasingly well-defined, with available benchmarks (122).

The FIM instrument is currently an accepted health status measure for rehabilitation. Rehabilitation facilities currently submit FIM data as indicators (e.g., change in FIM rating from admission to discharge). FIM ratings at discharge, LOS, and number of days from onset to admission might also be submitted. The rate of unplanned program interruptions, acute hospital discharges, and discharges to nursing homes might also be employed as problem indicators, as they have been employed for decades in rehabilitation PE.

Comprehensive rehabilitation is required, at minimum, to meet the most frequently encountered physical and psychosocial needs of the patient. The services of physical rehabilitation encompass rehabilitation medicine, rehabilitation nursing, physical therapy, occupational therapy, speech and language pathology, psychology, and social work or case management. Older Joint Commission publications have discussed quality monitoring, goals, and evaluation for specific physical rehabilitation service departments (123). Possible “monitors” for physical rehabilitation services have included increase in self-care ADLs, mobility, reduction in pain, patient satisfaction with services overall and with specific aspects of services (e.g., satisfaction with fit and functioning of prosthetic and orthotic services); cognitive and emotional adaptation of the patient and family to disability; improvement in communication skills; health maintenance; and reduction or prevention of preventable complications commonly seen in the impairment group. Such goals and indicators resemble the objectives and goals in PE systems in medical rehabilitation. A difference is that PE goals were typically for the entire program or a broad diagnostic group rather than for a clinically well-defined group of patients. Interdisciplinary team objectives are preferable to departmental goals where such a team is routinely critical to achievement of outcomes for the patient. In comprehensive rehabilitation, the attainment of functional goals for the patient typically involves such a team.

**GUIDELINES AND EVIDENCE-BASED PRACTICE**

Changes in the health care environment have resulted in increasingly systematic and critical examination of routine treatment throughout health care. No longer are health care...
consumers and payers content to make decisions on the basis of work-of-mouth reputation alone. Rather, quality and outcomes data are increasingly checked, and each episode of care is subject to both cost-containment policies and usually implicit standards of “best practice.” Clinical practice guidelines—also called practice guidelines or merely guidelines—explicate best care practices.

Typically, practice guidelines are detailed descriptions of how different types of patients should be evaluated and treated in different circumstances. They aim to define standard, quality clinical practice for commonly seen and well-understood patient problems. Because they define care quality, guideline development is intrinsic to performance monitoring and actual improvement in clinical practice. Guidelines should include enough detail to specify ordinarily appropriate decisions and processes. They typically involve a sequence of initial measures, alternative clinical processes, at least some decision rules, and subsequent assessment of patient responsiveness or clinical outcome. Current practice guidelines typically synthesize scientific evidence with knowledge of normal practice as well as information on patient values. Well-developed guidelines are more applicable to clinical practice than review articles (124), and they are particularly useful for clinical education.

There is a growing subset of practice guidelines called integrative care pathways (ICPs) (125). Originally termed anticipated recovery pathways, ICPs are often referred to as multidisciplinary pathways of care, care protocols, care maps, or collaborative-, coordinated-, or critical-care pathways. As with any guideline, ICPs aim to have the right people provide the right care in the right sequence and in the right place to achieve the right outcomes. ICPs are designed to help clinicians and facilities reach or exceed existing quality standards and decrease unwanted practice variation (126). ICPs involve tracking of deviations (termed variances) from the expected care plan or outcomes. Information obtained from these variances is used to alter care plans or adjust outcome expectations so that future patients may experience better and more predictable results (127,128).

### Evidence and Bases for Guidelines

Guidelines have traditionally been based on knowledge of usual practice, and their implementation has focused on identification of practice variations. Increasingly, guidelines are based on systematic evidence review methods (7). The measures and criteria employed in guidelines and quality and outcomes monitoring systems need to be based on best-available evidence. Medical rehabilitation, like all health care disciplines, needs to strive to base its practices on stronger evidence than is provided by custom or observational data. Criteria for such evidence have become increasingly well-defined as the movement for evidence-based medicine and EBP has grown. It is now possible to systematically grade the strength of available evidence for the care received by rehabilitation patients (53). Systematic reviews and meta-analyses (see http://www.cebm.net/) provide the essential basis for quality monitoring and improvement. Chapter 80 explains EBP in rehabilitation more fully.

High-quality guidelines are developed using increasingly standardized, formal methods of synthesizing information from expert judgment as well as from research studies (129). The combination of systematic evidence review and expert consensus can result in an authoritative and reliable guideline that can give clinicians greater confidence in making treatment decisions (130). Randomized clinical trials (RCTs) are the most widely accepted standard for evaluation of treatments. When strong evidence of treatment effectiveness is synthesized into guidelines, clinical applicability is clear (27,100,121).

Direct evidence for the effectiveness of many types of rehabilitative care is limited, but relevant RCTs and a larger number of lesser controlled studies do exist, with the strength of evidence varying across diagnostic groups and interventions (53).

Synthesizing evidence on the effectiveness of rehabilitation is not easy. Rehabilitation is multifaceted and involves an array of medical and nursing interventions as well as therapies that work through the activity of the person, such as physical therapy, occupational therapy, and speech/language pathology. Strong evidence is unlikely to be available for this entire array of interventions. Nonetheless, interventions based on learning and physical conditioning surely “work,” as do many environmental modifications, prostheses, orthoses, nursing interventions, and pharmaceutical interventions from general medical care. Guidelines can be written that integrate evidence and experiences from many therapies and methods. Adherence to such guidelines has been associated with greater functional gain and patient satisfaction (131,132). In addition, carefully developed evidence-based guidelines provide a good basis for quality-outcomes monitoring.

### Grading Strength of Recommendations to Strength of Evidence

Strength of recommendations should be graded according to the strength of evidence. Table 12-1 presents the most widely used method for connecting level of evidence to strength of recommendation. The method is used by several medical societies and evidence review organizations. Though often overlooked, the point is critical to rational and successful implementation of evidence-based guidelines.

<table>
<thead>
<tr>
<th>Level</th>
<th>Recommendation</th>
<th>Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Should be done</td>
<td>Established as effective: two consistent class I (rigorously controlled) studies</td>
</tr>
<tr>
<td>B</td>
<td>Should be considered</td>
<td>Probably effective: one class I study or two consistent class II (e.g., cohort) studies</td>
</tr>
<tr>
<td>C</td>
<td>May be considered</td>
<td>Possibly effective: class III studies, e.g., natural history controls or own control</td>
</tr>
</tbody>
</table>

Source: (7).
When evidence is unequivocal, and the human outcomes involved are highly important, the treatment or procedure in question should be considered to be a standard. Compliance should be monitored. But evidence in rehabilitation typically is not unequivocal (53). Recommendations based on weak evidence are more likely to be overturned by future research, and there are likely to be exceptions that are currently unknown. When evidence for a clinical procedure is weak or merely suggestive, it is difficult to find a rational justification for policies requiring all clinicians to do it. At the same time, even weak evidence may be the best available, and such evidence should be seriously considered and presented as a treatment option. Consideration of best-but-weak evidence could be a quality indicator, but rigid adherence, exaggerating the strength of knowledge, would not have a scientific basis. The importance of the procedure, patient preferences, and the human outcomes involved—whether transient discomfort or total disablement hang in the balance—also affect the strength of recommendation.

Guidelines for Common Conditions
Initially led by the IOM and the Agency for Health Care Policy and Research (now the Agency for Health Care Quality and Research), work to develop clinical practice guidelines and clinical paths has spread throughout the health care industry (100,125). Thousands of clinical practice guidelines have been developed. An outstanding source for these is the National Guidelines Clearinghouse (http://www.guideline.gov/). At the time of this writing, the site lists 102 guidelines relevant to medical rehabilitation (excluding psychiatric and substance use rehabilitation, purely preventive guidelines, nonreconstructive plastic surgery and dental problems, and developmental learning disorders). Some of the more rehabilitation-relevant guidelines include those for

- Poststroke rehabilitation (133)
- Several pain syndromes, including low back problems
- Medical complications following SCI, including depression, autonomic dysreflexia, and prevention of thromboembolism
- Brain injuries (an authoritative National Institutes of Health review, but so general it should hardly be called a guideline)
- Cardiac rehabilitation (134)
- Treatment of lower-limb osteoarthritis, with rehabilitative implications (as with other guidelines, they vary in quality and use (135))
- Deep vein thrombosis
- Treatment of depression in primary care
- Treatment of pressure ulcers (136) and their prediction and prevention (137)
- Acute and chronic management of urinary incontinence (138)
- Fall prevention
- Exercise and strength training
- Several other neurological and neuromusculoskeletal conditions

More information on ICPs and examples of existing tools is available on the National Library for Health Web site (http://www.library.nhs.uk/pathways/). There is also a journal explicitly devoted to exploring all aspects of ICPs: the *Journal of Integrated Care Pathways*, published by the Royal Society of Medicine (http://www.rsmpress.co.uk/jicp.htm). Other sources also provide information on guidelines, care pathways, and related forms for medical rehabilitation (139,140).

Implementation and Impact of Guidelines on Practice
Research on the impact of guidelines on clinical practice and outcomes is emerging. Initial hopes that guidelines would greatly improve the quality of medical care, improve health outcomes, and constrain the growth of medical care expenditures as a whole have been dashed. Adherence to guidelines has often been limited, and many guidelines are little used after publication. Nonetheless, some well-developed guidelines have been widely accepted. Guidelines have repeatedly been associated with small-to-moderate improvements in care processes (141,142). Effects on health outcomes have been less studied and are less clear. As examples, researchers have documented 30% reduction in adverse events from antibiotics, a 27% decline in mortality, and decreased costs when using computer programs that help physicians choose antibiotic treatment (143). A detailed treatment protocol for mechanical ventilation can reduce unwarranted variations from good practice and substantially improve survival (144). Randomized trials have shown that, when combined with feedback on performance and education by respected peers, practice guidelines can improve medical care processes and outcomes (130); however, success varies among settings, procedures, and health conditions.

Developing and implementing guidelines for stroke rehabilitation are particularly difficult, given the multiplicity of presenting problems and consequent complexity of needed treatment. Clinical pathways have been shown to reduce LOS and complications in acute stroke care, but the results from stroke rehabilitation are more equivocal (145). Adherence to stroke rehabilitation guidelines was associated with greater functional recovery (146) and patient satisfaction (132) in a Veterans’ Administration study. Conversely, a randomized trial comparing a clinical pathway with traditional stroke-team rehabilitation (147) found no advantage in either cost or patient gains; it even decreased patient satisfaction. A possible explanation is that specialized stroke-team rehabilitation is already highly developed and consistent, and it would be difficult to write guidelines that improve functional gain, complication rates, patient satisfaction, and LOS attained by existing experienced and coordinated rehabilitation teams. Similarly, guidelines designed to explicate only basic care requirements likely define what virtually everyone is already doing.

Research on guidelines for depression deserves further mention, not only because persons with disability exhibit depressive symptomatology more frequently than the general population but also because research on their implementation is instructive. A substantial body of research has elucidated effective methods for recognizing depression and improving its treatment in ordinary care settings (148). Research has shown that clinical information and education are necessary...
but insufficient. To achieve levels of treatment and outcome recommended by the guideline, patients too must be educated and encouraged. “QI efforts that focus resources on improving systems of care and the active participation of patients offer the best evidence of improved patient outcomes” (149). Knowledge of organizational tactics and factors enhances the likelihood of successful implementation of guidelines (150).

Computer decision support systems have similarly demonstrated improvements in clinical practices (151). Using a touch screen, for instance, patients or staff can complete a questionnaire, and responses can be processed against evidence-based guidelines at the time of the clinical encounter. Care suggestions, targeted to the patient’s history; comorbid conditions; and current symptoms can also be provided. Reminder and suggestion systems have only begun to be developed, but both patients and health care practitioners have been generally satisfied with them (152).

Cost is a constant concern in health policy and in terms of actual rehabilitative care. Implementation of cost-conscious guidelines has been reported to save substantial sums in certain circumstances (145,153). However, methods of ethically incorporating costs into guidelines are not yet well developed or widely accepted (154). The cost-effectiveness of alternative approaches to assuring and improving QOL, including EBP, QI, and patient partnerships, needs to be examined (155). Even with all these shortcomings, we can expect the development and use of clinical guidelines to increase in the future.

The ultimate goal of introducing clinical guidelines and paths is to establish the means for a multidisciplinary team to provide coordinated, timely, outcome-focused, and patient-centered care while using the minimum necessary resources (125,156). Standardized protocols facilitate this process by revealing detours from established best practice, identifying patients who are not responding as expected, and generating data sources to stimulate research (157,158). Furthermore, guidelines can be educational for new team members.

Multifaceted approaches to guideline implementation, involving provision of written information, talks with local consensus discussions, and individualized contact by colleagues (academic detailing), appear to be most effective in dissemination (141). Guidelines and attempts to implement them should be pilot tested and built into normal, ongoing channels for improving care. Obstacles to implementation vary across organizations, so a preliminary diagnosis of the organization itself is logically needed to identify barriers and optimal implementation strategies—a topic on which more research is needed.

**Critique of EBP and Guidelines**

Evidence-based guidelines are now widely accepted by clinical professionals and policy makers as essential bases for clinical practice. Expert, best-practice guidelines can improve clinical performance and serve as a tool for clinicians and a source of information for patients. Evidence-based guidelines are a necessary and potentially powerful tool for accountability and QI, as well as to achieve effective and efficient care, though limitations are becoming clearer (159,160).

Guidelines can be rigid, based on the assumption that one single method of care is best when several approaches may be effective. Proliferation of forms and applicability to varying local circumstances remain concerns. Patient acceptance has not been considered in development of some guidelines, and though patient involvement may be recommended, specifics of how this is to occur are often not given. Guidelines need to be tested to evaluate their use in practice and whether they actually improve clinical processes and patient outcomes. They will also need recurrent revision.

While the feasibility of guidelines in rehabilitation—at least for certain better understood problems—has been established, most existing guidelines focus on medical/nursing aspects of care. Many conditions and clinical practices in rehabilitation are not covered by guidelines or by direct efficacy research. However, practice guidelines and care paths can be beneficial in chronic conditions wherein the focus is not on curing so much as maximizing functional independence and QOL (125,161–164).

Guidelines themselves vary in their quality. Some are well based on evidence, while others are based largely on expert opinion. The few guidelines that have been produced for rehabilitation involve substantial expert opinion (e.g., for stroke rehabilitation) or target highly specific conditions (e.g., prophylaxis of deep vein thrombosis). Criteria now exist for appraisal of the quality of clinical practice guidelines (165,166). Guidelines need to be developed to the point that they are useful for professional and patient education and should be sufficiently detailed and clear that one can objectively evaluate the degree to which patients have received the needed care and experience the expected health benefits.

**INFORMATION SYSTEMS FOR PERFORMANCE MONITORING**

Until recently, “health care delivery has been relatively untouched by the revolution in information technology that has been transforming nearly every other aspect of society” (1, p. 15). The mass of information on personal health collected in encounters with health care professionals is a great potential resource for improved quality of care, but that information is dispersed in poorly organized, sometimes illegible, paper records at a variety of sites, with fears of security breaches inhibiting access to the information needed to understand and manage patients with chronic illnesses and disabilities. Systematic improvement of care, quality, and outcomes in health care involves attention to information flow, the medical record, and the clinical data system. A sophisticated system for data retrieval, analysis, and reporting is required to convey information to clinicians in a useful, timely way, improving quality and outcomes.

**Data System Structure: Inputs, Processes, and Outcomes**

Performance monitoring systems involve measures of three types—inputs, processes, and outcomes. Figure 12-4 shows...
two schemata for outcomes-oriented monitoring systems. In traditional PE systems in medical rehabilitation, the emphasis has been on outcomes, and the process box is sparse. There is no explicit theory or guideline to determine what aspects of intervention should be measured, so that rehabilitative processes are treated virtually as a black box (at the top of the figure). The approach has been useful to characterize patient gains associated with comprehensive medical rehabilitation programs, where processes are so multiple and complex their complete explication would appear to be an intractable task. Many current data systems in medical rehabilitation are like this, including data systems for the PPS based in the IRF-PAI. A minimal ability to analyze the effect of general factors such as intensity of treatment, primary impairment, payer, and demographic factors for somewhat similar patient groups (e.g., FRGs or CMGs) may be provided. However, the limitations of this model have been increasingly recognized (13,51,112,167,168). When needed processes are not stipulated and measured in the data system, it is extremely difficult to identify what might be done to improve the appropriateness or outcomes of care.

More sophisticated data systems—tied to the actual process of rehabilitation planning and treatment provision—have been attempted in rehabilitation (79,80,169). As displayed at the bottom of Figure 12-4, treatment objectives and progress are represented by arrows. While some functional and medical measures may be constant across all patients, others vary to permit needed individualization. Treatment objectives are chosen to fit the priority needs of the individual (79,80). Ongoing patient reassessment is part of quality rehabilitation (20). In the Joint Commission Standard 6.2, so a more adequate data system incorporates change in patient functioning and treatment objectives. Whether medical and functional goals have been attained (and possible reasons if not) is determined in patient follow-up. Requirements for attainment of a productive, independent lifestyle are evaluated. Given the importance of discharge planning to quality rehabilitation (20), Joint Commission Standard 6.1.1), environmental, family, and other requirements for discharge to a maximally independent living arrangement need to be evaluated.

The provision of rehabilitation services should be guided by an interdisciplinary plan (20). When objectives in a performance monitoring system are based on actual rehabilitation plans for individuals, it is a true clinical outcomes management data system. With such a system, the team can be provided with specific and potentially valuable feedback on their ability to choose objectives, implement effective interventions, and attain outcomes for individuals served.

**Information System Design and Integration**

A fully automated medical record is not needed to achieve QI (1). Automated order entry systems can reduce errors in prescription and delivery of medications (170). Reminder systems have recurrently been shown to improve compliance with clinical practice guidelines (171). Although current examples are few and limited to specific structured problems, computer-assisted diagnosis and management promise to improve quality in the future (1,131).

Information has traditionally been entered in free form, sometimes illegible notes, into paper medical records.
If entered into computers, new natural language search engines make retrieval of information from free text notes possible. Even so, natural language recording yields results of limited reliability, given individual variations in style and completeness of recording.

Information systems, whether on computer or paper, can present clinicians with standardized, relevant considerations and information to assist them in their diagnosis and treatment decisions and in monitoring response to treatment. Common evaluation and intervention processes can be presented in a format (e.g., prompts or a flow sheet) that reminds clinicians of standard best practices and simultaneously facilitates recording of whether these are done. Additional screens or pages can facilitate standardized measurement of patient responsiveness. The importance, appropriateness, simplicity, and transparency (understandability) of items presented to clinicians are critical to a usable clinical data system. While structured input is needed for systematic QI and outcomes monitoring, free text notes are still required to record individual variations and for the numerous clinical situations for which validated, structured guidelines do not exist.

The idea of integrating information systems is hardly new, but it is still common for health care organizations to have multiple poorly integrated recording systems. Clinical data may be recorded in paper, medical records, professional files, or in pharmacy and other departmental computers. Billing records may be in an entirely separate system. The integration of clinical, financial, case management, QI, and outcomes data increases the potential use of these systems (75,79).

Computerization assists QI by increasing the amount of data available but does not in itself provide more relevant or useful information. A mass of information can enhance decision making, but it can also confuse the process. The bottleneck may not lie in collection of data but in analyzing data and synthesizing it with knowledge, based on evidence and clinical experience, to highlight meaning. One could argue that the performance monitoring literature has concentrated too much on measurement and too little on how to use the data. Routine performance reports give rise to hypotheses about problems in the program or why outcomes are or are not attained. In-depth analyses are needed to test hypotheses about alternative explanations about why performance is better or worse than expected. The information system must facilitate professional statistical investigation and interpretation.

Information systems are justified not by single use but by their multiplicity of uses. The following sections will discuss clinical and then management uses.

Clinical Reporting and Uses

Reports of functional gains by patients have been used in rehabilitation for decades, not only to evaluate the rehabilitation program as a whole and to assist in QI but also to organize communication and set goals in the rehabilitation team conference (172). Patient functioning and improvement are key indicators in utilization review and assessment of readiness for discharge in rehabilitation (37,172). Data on functional history and gains have also been used to assist case managers by predicting outcomes.

Rehabilitation information systems can produce informative, graphic displays of individual patient progress used in team conference, case management, and reports to referral sources or payers. Figure 12-5 presents an example of such a report (63,172). Functional tasks are ordered by difficulty: the easiest, first-to-recover activities are presented to the left with more difficult activities to the right. The case displayed is a possible discharge candidate. Team discussion of the patient’s functioning, lifestyle preferences, special needs, and family

![Figure 12-5](status-and-goal-profile-of-a-possible-discharge-candidate)
support is needed for quality rehabilitative management and a safe, high-quality community placement (173).

Well-designed reports tell us whether desired, attainable results are in fact routinely attained and whether interventions are being applied to patients who are most likely to benefit from them. There are several common, avoidable traps in interpretation of clinical outcomes data. First is the common tendency to assume a massive effect of treatments. Rehabilitative interventions will tend to lessen complications and augment healing and adjustment, but quick cures are not ordinarily expected. A second common error is assuming that improvement is due entirely to rehabilitation. Improvement can be due also to natural healing and family and environmental processes that can occur at home. Conversely, failure to improve may not reflect provision of inadequate treatment. Reports need to incorporate sophisticated severity adjustment to identify patients with anomalously poor outcomes. Even then, detailed clinical data are needed to judge whether failure to improve is associated with provision of substandard care (81).

**Case Management Databases and Their Uses**

Difficulties in use of traditional routine clinical monitoring data systems are not entirely due to technical deficiencies, although these exist. When there is a concrete motive (e.g., a financial motive), clinical data systems are used. Managed care organizations, for instance, have developed their own databases to track the appropriateness of medical care, including medical rehabilitation. These databases concentrate on financial factors but increasingly include detailed clinical data. If the case management agency does not like the care or outcomes of patients sent to a rehabilitation provider, it does not send new cases to that provider. Only a minority of expensive cases are managed with the use of such databases, but for this minority, the external party’s data system is essentially the operational outcomes management database. These case management systems are typically proprietary, but their adequacy and impact on patient care and outcomes are of public interest.

**Management Reporting and Uses**

Management reports are the routine product of performance monitoring systems. Patient progress, indicators of clinical outcomes and processes, goal attainment rates, efficiency, case-load, service intensity or frequency, and trends are reported periodically. Reports may go to the governing board and staff as well as clinicians and managers. Both frontline staff and key decision makers should receive findings and use them. Performance reports should engage staff at all the levels of the organization involved in actually implementing possible improvements based on the findings. Data on patient progress or outcomes may also be formatted for release to purchasers of services and the public (18,69,70).

**Managing Length of Stay and Cost-Effectiveness**

Management of LOS and readiness for discharge is critical in rehabilitation facilities in the United States and is a major use of clinical information systems. Additional uses include marketing (174), refined profitability analyses, planning, accreditation, and estimation of patient acuity for determination of staffing requirements (71,72,75,79–91). LOS, however, is not an indicator of quality. Shortening LOS in rehabilitation has been accompanied by an increase in the frequency of rehospitalization after discharge (175). Discharges of hip fracture patients from acute hospitals with active clinical issues (e.g., temperature, blood pressure, other vital signs) or with new impairments have been associated with increased rates of rehospitalization and mortality, and patients who develop new impairments have worse functional mobility (176). Information systems need to present data on patient readiness for discharge.

Interpretation of cost-effectiveness data requires understanding of basic relationships. More severe cases tend to receive and require longer care (25,61). While outcomes per se often have little relation to effort or even an inverse one, improvement in medical rehabilitation hospitals is and should be probabilistically related to LOS (25,60). Strategies for cost-benefit and cost-effectiveness analysis in rehabilitation are presented in other works (177).

**Outcomes-Focused Quality Improvement**

Given the lack of evidence for a single best way to deliver rehabilitative care, it makes sense to monitor functional outcomes, giving programs leeway in how they produce these outcomes. Shaughnessy et al. have shown that outcomes-based QI can substantially improve patient outcomes (178). In a study of over 300,000 patients receiving services from home health agencies, outcomes-based QI reduced rehospitalization for targeted conditions by 22% to 26% over 3- to 4-year demonstration projects, compared to a 1% reduction in matched non–outcomes-based QI agencies. The risk-adjusted rates of improvement in outcomes-based QI target outcomes measures of health status averaged 5% to 7% per year in demonstration trials and were significantly greater than analogous improvement rates for nontarget comparison outcomes, which averaged about 1% per year. There is no reason why outcomes-focused approaches should not work for other forms of rehabilitation as well.

In general, when a program or patient group has higher or lower outcomes than projected, this by itself does not constitute evidence for high or low quality or effectiveness of care (87). Although FRGs/CMGs can be modified to predict up to 63% of the variance of discharge FIM scores, current severity-adjustment methods typically predict only a minority of the variance of patient gains, so that much is left unexplained (87,179). Outcomes after rehabilitation are due to many factors beyond the control of even high-quality care. As a consequence, a discrepancy between expected and actual outcomes is validly usable (only) as a first-stage screen—an indicator of a possible problem or opportunity to improve operations (Fig. 12–6). The discrepancy indicates that further investigation is needed.

Second-stage screens or investigations are required to determine whether there are deviations in care processes or whether
differences in outcomes reflect unmeasured patient or environmental factors. Many studies show that appropriate use of outcomes data to improve clinical processes "virtually always requires detailed clinical data" (119, p. 868). Peer review of charts may be employed to determine reasons for deviations but are expensive. Chart review may focus on patients whose outcomes are particularly discrepant, but ad hoc methods may fail to detect remediable problems.

A system for in-depth analysis and review of detailed clinical data, accessing well-developed guidelines and evidence regarding desirable clinical procedures, is needed to confirm whether a discrepancy in process or outcomes implies actual deficiencies or not. The difficult process of determining reasons for suboptimal outcomes should not be left to harried local QI committees or part-time PE staff without well-developed tools to assist and guide their efforts. Objective review criteria that connect outcomes to process deviations are feasible and have begun to be developed for home health services (144) and stroke rehabilitation (132, 146). Perhaps the greatest technical problem with current approaches to QI in rehabilitation is that the evidence-based guidelines and methods of operationalizing them in second-stage analyses have not been developed and validated.

When a problem or opportunity for improvement in care processes is identified, the quality task force should direct effort toward solving the problem. In this third stage, detailed practical planning should occur, followed by implementation and change in care processes. Continued monitoring will reveal the impact on patient outcomes. In principle, performance improvement efforts can and should have the capacity to focus on several points in the causal chain connecting clinical processes to intermediate and long-term outcomes. QI analyses, for instance, can identify patients who experience variances in care processes and determine if their severity-adjusted outcomes are in fact below expectation. If so, action is indicated. Evidence-based clinical guidelines and methods of operationalizing them are needed to improve performance, whether the initial examination focuses on processes or outcomes.

FIGURE 12-6. The outcomes-focused QI cycle.

OTHER APPROACHES TO PERFORMANCE IMPROVEMENT

Professional Education and Development

Education of professional and other staff is an essential component of QA and QI. The quality of care provided in practice is dependent on professionals’ pride, ownership, and self-regulation. In appropriate environments, clinical professionals will strive toward improved competence and improved practice. These assumptions underlie current professional licensure, testing, and continuing medical education, although classic continuing medical education approaches involving courses, provision of written materials, and conferences have demonstrated only limited effectiveness (160, 180). A variety of more interactive, realistic, and timely educational methods may prove to be more effective. Educational outreach and personally tailored approaches are promising; small group learning and physician peer review are also useful (160). Bottom-up approaches, involving clinical education and formation of groups of clinicians charged with improving quality, have been shown to improve clinical practice (160). A variety of new and revised educational techniques will be needed to provide the skills required to enhance the quality of health care, including skills in interpretation and applications of evidence and guidelines, using decision support systems, and communication with patients as full partners in decision making (1).

Patient-Centered Approaches

Although details may vary, patient-centeredness and assessment of quality from a patient’s perspective are essential components of quality health care (1). We will primarily consider the monitoring of patient satisfaction, a clearly practical matter, and then comment on broader issues of patient involvement or empowerment.

Patient Satisfaction

The monitoring of patient satisfaction has deservedly become a standard part of performance monitoring in health care. Patient satisfaction is evaluated as both a process and an outcome. There are persuasive practical reasons for organizations to have
a strong emphasis on patient satisfaction. Patient satisfaction is measurable at a modest cost (181). A facility’s reputation is largely based on satisfied clients. Marketing is enhanced by high customer satisfaction, both by word of mouth and presentation of satisfaction information. Patient perceptions of quality have been found to be robustly correlated with hospital profitability (182). Increased patient satisfaction has been associated with decreased risk of malpractice suits (183).

Measuring patient satisfaction helps both clinicians and administrators understand patients’ care experiences. Patient satisfaction assessment can identify problematic processes. Dissatisfaction rates are much higher in facilities that misrepresent the services they offer (27,181,184). Satisfied patients tend to be more compliant to treatment recommendations (185).

Well-validated, reliable questionnaires, with available normative data, are available for assessment of patient satisfaction and should be used. Multipoint rating scales have become standard (e.g., ratings from “very poor” to “very good” or “excellent”). Valid satisfaction questionnaires avoid bias, giving customers freedom to complain if they feel like it. The point is to get information rather than to manipulate a vote of approval. Mail provides more anonymous results than phone calls and at lower cost, but low-response rates can be a problem. Patient satisfaction questionnaires have been published for general medical settings (186) and for inpatient (187) and outpatient (188) rehabilitation programs (23). Many large group health organizations have adopted standard formats for assessment of patient satisfaction based on Ware’s research (189). Third-party firms provide standard satisfaction monitoring systems, including reports of comparative data across settings. Reports are needed for different types of patients and should target different actionable areas of care.

Patient satisfaction measures have limitations and need to be presented to staff in a positive light. Care issues are often viewed positively. Patients tend to express high satisfaction with medical rehabilitation services. “The most consistent finding is that the characteristics of providers or organizations that result in more ‘personal’ care are associated with higher levels of satisfaction” in medical settings (181), while patient satisfaction is poorly or inconsistently related to technical effectiveness and professional standards. Satisfaction measures often elicit comments about food, temperature, billing hassles, and other facility and personal services. Billing and cost issues typically receive ratings that are distinct from care issues and should be summed and reported separately from care issues. Administrative journals provide advice on how to interpret and use patient satisfaction data (190).

Patient Involvement
Increased patient involvement or empowerment has been propounded as a way to improve the quality of rehabilitation programs. CARF accreditation particularly emphasizes and even requires patient involvement (18), and Joint Commission standards now also state that the rehabilitation plan is “developed by qualified professionals, in conjunction with the patients and/or his or her family, social network or support system” ((20), tx.6.3). Patient autonomy is a basic value, and involving “consumers” is seen as essential to enabling them to receive the individualized care they desire. Patients’ valuation of outcomes may differ from those of professionals, and individual patients have differing functional needs. “Consumers” of rehabilitative services may choose cognitive and communicative abilities over physical abilities (191), although conventional outcome measures weight them equally or even give greater weight to physical outcomes. The significance of the same disability can vary among patients (27,81,192), who have differing lifestyles and live in differing cultures.

Differing methods of empowering or involving patients have been employed, including satisfaction surveys, complaint procedures, communication training for professionals, surveys of patient needs, and enhanced provision of educational information. One of the few studies on the topic has reported encouraging results of a system involving a patient care notebook in a rehabilitation setting (193). Patient, family, and team are all involved in the notebook, which also educates the patient in self-care and is used as a resource when the patient goes home.

Many studies have been done on physician-patient communication, and reviews show that various indicators of sensitive caring communication and patient-centeredness lead to greater patient satisfaction and even to better compliance with prescribed treatment (194). While most patients want to receive information on their condition and treatment alternatives, many do not want to make key decisions regarding necessary treatment; seriously ill patients in particular may not want to take responsibility for management of their disease (195).

After decades of simplistic advocacy of patient involvement and professional suspicion of patient opinions, new conceptualizations of patient-centered care have been propounded. “Dialog-centered care,” based on a clearer definition of the rights and responsibilities in communicative process between clinicians and patients, is among the most promising of these (160). Research is needed to clarify methods and circumstances for appropriately involving patients in such a way that satisfaction and outcomes are improved.

Clinical Practice Improvement
Systematic “clinical practice improvement” (CPI) has led to insights and improved care processes in a number of areas of medical care (196,197). Developed by Dr. Susan Horn, CPI is a data-driven “bottom-up” approach involving collection of data on processes of care, relevant outcomes, and patient characteristics, including indicators of severity. Clinicians who choose different practices from those in the protocol are given the opportunity to present their reasoning to the CPI team so that the protocol can be modified or consensus can be reached. The goal of the process is to enable clinicians to improve severity-adjusted outcomes within cost limits or to maintain relevant outcomes while decreasing costs. Sophisticated multivariate statistics are used to statistically control for factors that confound process-outcome relationships for selected subsets...
of patients. Process variations are lessened through ongoing feedback from statistical analyses, discussion, and consensus.

Proponents of CPI argue that clinical practice guidelines are often based on a consensus of experts, and what scientific evidence is available is often limited or based on selected patient groups and programs that vary from those seen in practice. CPI studies have identified actionable factors that associated with severity-adjusted outcomes (e.g., earlier admission of stroke patients to rehabilitation) (198). The validity of CPI depends on the accuracy of the statistical adjusters employed—an uncertain matter. Compared to typical RCTs, CPI has less internal validity but greater external validity, that is, greater generalization to practice.

PUBLIC ACCOUNTABILITY AND HEALTH SYSTEMS ISSUES

Trends in health care as a whole impact and even overshadow rehabilitation, which is a relatively small part of the total health care “system.” Current payment policies in the United States are complex and contradictory and may provide no incentive for improving the quality or true efficiency of care (1). Rehabilitation may face a “market” that disempowers its “customers”—persons with disability (199).

Changes in the Health Care Industry

Cost control mechanisms, including various managed care mechanisms, risk-sharing arrangements, and Medicare’s PPS, have heightened concern about the quality of rehabilitative care. Although reports on the quality of care have not always shown that managed care arrangements provide worse care to persons with disabilities and chronic conditions, many of these firms have incentives to do so, as these patients cost much more to care for than the average patient (31). Under capitated payment arrangements and disabled and chronically ill individuals are money losers, so the incentive is not to enroll such individuals. Severe disabled individuals lose their private insurance and shift to Medicaid or Medicare. The disabled individuals affected may be poor and limited in mobility, and some cannot communicate articulately and so cannot advocate effectively for themselves. Organizations may decline to participate in studies of rehabilitation quality and outcomes. Mortality and apparently also rehospitalization rates following inpatient rehabilitation have increased in recent years (200). In sum, rehabilitation is vulnerable to adverse pressure in the changing health care system in the United States, and quality of rehabilitative care is or should be a public issue as well as a professional one.

Public Reporting of Performance

Fears regarding simplistic misinterpretation of quality and outcomes data are still with us, leading to restrictions on access to and use of data. The data in rehabilitation clinical information systems, like that in other medical information systems, have been the private property of the facility. Consumers and the public, however, are major stakeholders to the rehabilitation process and have a valid claim to deidentified performance information (i.e., information that does not identify any individual patient). Medicare now releases severity-adjusted data on quality of care in nursing homes and home health (see http://www.medicare.gov/). The Center for Medicare and Medicare Services may release similar reports on outcomes in inpatient rehabilitation.

Payers and large health care institutions have developed methods of profiling the cost and case mix of physicians and facilities, but little has been published regarding the reliability or validity of these systems. Because nonspecialist physicians may encounter only a small number of patients with a specific disease, physician profiles could easily be unreliable, and individual physicians could game the system by avoiding just a few severe, high-cost, or poorly adherent patients (201).

Health Plan Report Cards

Consumer Assessment of Health Providers and Systems (CAHPS) was developed with funding from the Agency for Health Care Quality and Research to evaluate quality of care from the viewpoint of consumers or patients. The CAHPS has evolved to a family of standardized surveys on multiple types of care, and reports from the CAHPS are widely disseminated (http://www.cahps.ahrq.gov/default.asp). CAHPS’ surveys on disability and rehabilitation have not yet been developed.

The National Committee on Quality Assurance (NCQA) (see http://web.ncqa.org/) was formed to address quality issues among managed health care plans. Without objective measures and credible data, competition between such plans is possible only on the basis of cost. The NCQA accredits health plans, and its Health Plan Employer Data and Information Set (HEDIS) is now used by the majority of health care plans in the United States to provide data on quality processes, professional resources, and outcomes (see http://www.ncqa.org/). The number of HEDIS quality indicators has increased over time too. The current (2008) version has expanded to include 80 indicators sets, including indicators of effectiveness, access/availability, satisfaction, plan stability, intensity of care utilization, and structural descriptors. A few outcome indicators or surveys are specified (e.g., the Medicare Health Outcome Survey), but most are process indicators. Quality indicators are specified for a number of chronic conditions (e.g., diabetes, cholesterol management after acute cardiovascular events) and for disease management programs but not for the neurological, traumatic, or orthopedic conditions commonly seen in inpatient rehabilitation hospitals. HEDIS data are compiled into a national database, and report cards are issued on health plans.

Whether HEDIS is measuring the most important factors is controversial. HEDIS quality indicator sets, however, continue to evolve, expand, and improve. Case studies attest to the use of the database in organizational QI efforts (http://www.qualityprofiles.org/). A broader issue is that accreditation...
and much reporting are voluntary. Health management organizations (HMOs) with lower quality-of-care scores have tended to stop disclosing their quality data (202). In sum, adequate health plan report cards have not yet been developed for disability and rehabilitation, but CAHPS and HEDIS provide an essential basis for future progress.

Incorporating Wider System-of-Care Factors in Performance Monitoring
The monitoring of the quality of the specific services provided can beg the question of whether a different type of service might provide superior benefits to patients and whether the person served has major unmet service needs. Although monitoring of current care processes is appropriate, performance improvement systems have greater potential use if they consider all interventions known to be effective for patients’ problems, regardless of whether they are currently offered by the program. Studies of quality in acute care hospitals have reported that errors of omission (e.g., a physician not detecting a major diagnostic problem) were more common than errors of commission (not providing the correct treatment for the diagnosis) (29). Errors of omission of needed rehabilitative treatments may be frequent as well, especially given today’s cost and LOS constraints. The results of inpatient rehabilitation, for example, will be compromised if quality follow-up care is not provided. A wider consideration of patient needs might lead to QI efforts that address patient referral, adding or changing service mix, and education of payers regarding service needs. Patients, their families, and disability advocates can be influential allies in such education.

Cost and Value
Awareness of the critical importance of cost has spread throughout health care, including rehabilitation. Even in the midst of pressures, however, it would be naïve to believe that cost is the only issue: the issue is the balance of cost to benefits or, in more general terms, value (31).

Interventions that are likely to produce small improvements in functioning of patients with severe chronic conditions are commonly available, but are these improvements—and hence, the interventions—worth the money? Systematic research is needed to provide evidence regarding the value of various rehabilitative interventions, bearing in mind that this value may vary depending on context. CARF accreditation regulations have touched on the issue. They have addressed the need for rehabilitation programs to formulate feasible objectives, involving the patient in this process, and to communicate appropriate information regarding associated results to stakeholders (18,68). The agreement provides prima facie evidence that attaining the objective is probably of some worth to individuals involved, but it is difficult to summarize the value of attainment of such diverse individualized results to payers and policy makers. In any case, honest communication regarding patient benefits associated with rehabilitation, including objective data on goal attainment and evidence regarding treatment effectiveness, to major stakeholders remains essential to promoting a reputation for quality care and for nurturing the resources needed to maintain that quality.

Rehabilitation managers today are charged to develop systems that place each individual in a care setting that optimizes both outcomes and cost containment. The ability to predict outcome and to tier rehabilitative care to the level needed for the individual’s improvement is becoming a necessity. To assure that each patient is provided the most effective rehabilitative care, objective indicators need to be developed that specify the level, type, and dosages of rehabilitation that should be provided.

SUMMARY AND CONCLUSIONS
The IOM has documented the “quality chasm” between care realistically possible and that actually provided (1). To cross the quality chasm, the IOM has proposed six objectives for 21st century health care systems: health care should be safe, effective, patient-centered, timely, efficient, and equitable. The current nonsystem of care for persons with disabilities in the United States does not often meet these objectives. Persons with disability frequently are unable to obtain needed items and services from their health insurance plans and “experience more problems than others with follow-up care, availability of specialists, getting to doctors, and obtaining help during off hours” (199). Few would contest the assertion that there is a chasm between the quality of rehabilitation possible and that commonly provided to patients with disabling conditions.

Strategies for Performance Improvement
Health care quality monitoring and use of QA data have changed greatly since QA programs were originally established. The focus has shifted from identifying poor care toward identification of feasible improvements that will benefit the majority of patients with frequently encountered problems and the implementation of these improvements. Quality of care is now understood to be complex and multifaceted, and no aspect of the total system of care can be excluded from consideration if quality is truly to be improved.

More than a decade after their initial vogue, TQM and CQI remain theoretically attractive approaches (155). The emphasis on quality care, not as performance of a defined action or even sequence of defined steps but as processes organized around patient problems, remains persuasive, as does the emphasis on improving systems, improving cultures, and developing smooth transitions between systems. Rehabilitation will do well to learn from CQI and TQM efforts in health care more generally, including acute medical care, psychiatric care, geriatric medicine, home health care, and long-term care (99,144,203,204). Cooperation within specially trained multidisciplinary QI teams leads to improved quality in medical care (205). Multidisciplinary team approaches are also particularly valuable approaches to operational QI in rehabilitation.

At the same time, limitations of TQM/CQI have become evident. As Grol points out, “the current TQM approaches should be better adapted to the realities of health care. Links
to other approaches (e.g., professional development and external assessment) should be established. Physicians should be actively involved and occupy leadership roles, and the QI activities should be related to their needs and to patient-related problems more directly” (160, p. 2583).

Multiple approaches are needed to assure and improve the quality and effectiveness of rehabilitative care. Processes and short-term and long-term outcomes all need to be measured either continuously or in occasional focused studies to identify opportunities for performance improvement as well as to keep score. More broadly, improving the quality of care in practice requires “building bridges among professional pride, payer profit, and patient satisfaction” (160). Major stakeholders to the rehabilitation process—patients, families, referral sources, payers, government, and the public—need to be educated to promote the supportive partnerships necessary for provision of quality rehabilitative care.

To improve rehabilitative care systematically, EBP guidelines are critical. Without such guidelines, there is no standard by which quality of care can be evaluated and neither process nor outcomes monitoring has definite action implications. The development of practice guidelines in rehabilitation is difficult, given the breadth and complexity of rehabilitation, and their implementation and results in practice should be monitored and studied. Much work remains to be done to determine the most valuable rehabilitative care processes and to incorporate the needed tailoring of treatment objectives to patient values. Evidence-based guidelines will not be possible for all rehabilitation problems, but even with limited evidence (e.g., stroke rehabilitation), it is possible to develop useful recommendations.

Rehabilitation professionals have developed systems for monitoring general outcomes and patient satisfaction. In their day, these systems were steps forward, but the limited evidence basis of these systems—including the misidentification of functional gain with true evidence of treatment effectiveness, combined with limited transparency and public disclosure—limited their utility. Work is needed to develop full-featured performance information systems that monitor care processes and outcomes and provide information from controlled research to support the interpretation of data. Such information systems should enable clinical professionals and patients to make better decisions as well as provide reports to administrators, payers, and government. Developing and using the results of such systems to assure and improve the quality and outcomes of rehabilitation programs are a primary challenge to the field and its stakeholders.

REFERENCES

133. Post-Stroke Rehabilitation Guideline Panel.
PART II

The Rehabilitation Team: Systems-Based Practice
Patients undergoing comprehensive rehabilitation require the services of multiple health care providers who possess unique skills, training, and expertise that are employed for the full restoration of these patients’ function and their optimal reintegration into all aspects of life. The competent physiatrist must be able to communicate in an optimal fashion with all these providers to meet the many needs of the patient. Prescriptions, referrals, and orders are basic tools by which the physiatrist may communicate the desired involvement of other rehabilitation or medical specialties in assessment, treatment planning, treatment delivery, provision of equipment, and fitting of adaptive devices. Medical specialties that are commonly involved with the rehabilitation patient include neurosurgery, neurology, geriatrics, primary care (including family practice, internal medicine, and pediatrics), psychiatry, urology, and orthopedics. Many other medical and surgical specialties are consulted as needed. Assessment, treatment planning, and therapy are often provided by rehabilitation clinicians specializing in occupational therapy, physical therapy, kinesiotherapy, prosthetics and orthotics, psychology and neuropsychology, recreational therapy, speech and language pathology, rehabilitation nursing, social work, dietary science, case management, and others (Fig. 13-1, Table 13-1) (1,2). Which professions are involved with a particular patient and the extent of those involvements are largely determined by the nature of the patient’s deficits and the structure of the setting in which rehabilitation is being conducted. As indicated by an initial comprehensive physiatric assessment, the physiatrist requests the participation of other rehabilitation specialists for their assistance in determining the appropriate rehabilitation services and level of care, as well as for comprehensive rehabilitation planning, conduct, and monitoring of treatment, discharge planning, and patient and family education.

The health care team is a group of health care professionals from different disciplines who share common values and objectives (3). Halstead performed a literature review, covering the years 1950 to 1975, on team care in chronic illness and concluded that a coordinated team care approach appears to be more effective than fragmented care for patients with long-term illness (3). More recently, the efficacy and efficiency of team care has continued to be lauded (1,4–9).

Writing physical medicine and rehabilitation (PM&R) therapy referrals, equipment prescriptions, and coordinating care requires the skills of a well-rounded clinician who is adept in both therapy and patient interactions to form an effective health care team individualized to the needs of that particular patient. Deficits in knowledge base or team and patient interaction skills lead to suboptimal treatment plans and care. The well-trained rehabilitation medicine specialist is able to develop comprehensive PM&R treatment plans of substantial detail when warranted. The degree of documentation and specification required depends on the mode of team interaction and treatment adopted by the professionals involved. Effective participation in treatment planning, nevertheless, requires the ability both to generate and to support the rationale behind multiple interventions as well as a thorough knowledge of the methods and systems required to achieve a particular rehabilitation goal, within each of the disciplines involved. These interventions must be appreciated in terms of their impact on function as well as on each patient’s pathophysiology processes.

Treatment plans are generated from goals that arise from the problem list developed during evaluation. The evaluation (see Chapters 1–3, 9 and 18) results in a set of identified problems that can be classified in various ways but typically are organized as medical, rehabilitation, and social problems. A set of goals or desired treatment outcomes is generated, along with an initial estimate of the duration of therapy necessary to accomplish each. Such goals assist the treating professionals in establishing therapeutic, discipline-specific goals that serve to support the overall medical rehabilitation plan, assist in identifying target skills that will be required to reach these goals, and serve as the foundation of a comprehensive treatment plan. This plan is a tool that patients, families, therapists, and other treating professionals examine for prognosis and expectations. It forms the basis from which all team members may suggest additions, deletions, methods of achievement, or modifications. The treatment plan is not a static document but rather remains dynamic as goals are accomplished, new goals are
identified and added, or some goals, which become irrelevant or unachievable, are eliminated.

Treatment strategies are developed to accomplish the identified goals. The specific strategies can be physician-directed, therapist-directed, or, ideally, mutually derived by the patient and team through the interdisciplinary process. The rehabilitation medicine specialist should be knowledgeable about all pertinent therapies, their methods and modalities, and the potential benefits and risks of each, in order to optimally apply the specific interventions desired from each therapy specialty that will help to accomplish the desired patient goals. The availability, benefits, and risks of adaptive equipment and their use to facilitate

### TABLE 13.1 Rehabilitation Team Members Associations, Organizations and Journals as of April 2008

<table>
<thead>
<tr>
<th>Discipline</th>
<th>Assn</th>
<th>Address/Phone/Fax/Web site/e-mail</th>
<th>Journals Published</th>
<th>Certif. Required?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Occupational therapist</td>
<td>American Occupational Therapy Association</td>
<td>4720 Montgomery Lane P.O. Box 31220 Bethesda, MD 20824-1220 Tel: (301) 652-2682 Fax: (301) 652-7711 <a href="http://www.AOTA.org">www.AOTA.org</a></td>
<td>American Journal of OT (monthly); OT Practice (semi-monthly)</td>
<td>Yes</td>
</tr>
<tr>
<td>Physical therapist</td>
<td>American Physical Therapy Association</td>
<td>1111 North FairFax Street Alexandria, VA 22314 Tel: (800) 999-2782 Fax: (703) 684-7343 <a href="http://www.APTA.org">www.APTA.org</a></td>
<td>PT Journal (monthly)</td>
<td>Yes</td>
</tr>
<tr>
<td>Prosthetist/orthotist</td>
<td>American Orthotic and Prosthetic Association</td>
<td>330 John Carlyle Street, Suite 200 Alexandria, VA 22314 Tel: (571) 431-0876 Fax: (571) 431-0899 <a href="http://www.AOPAnet.org">www.AOPAnet.org</a></td>
<td>The O&amp;P Almanac (annually)</td>
<td>Yes</td>
</tr>
<tr>
<td>Rehabilitation nurse</td>
<td>Association of Rehabilitation Nurses</td>
<td>4700 West Lake Avenue Glenview, IL 60025 Tel: (847) 375-6481 or (847) 375-4700 Fax: (847) 375-4700 <a href="http://www.rehabnurse.org">www.rehabnurse.org</a></td>
<td>Rehabilitation Nurse (bimonthly)</td>
<td>Yes</td>
</tr>
</tbody>
</table>
### TABLE 13.1 Rehabilitation Team Members Associations, Organizations and Journals as of April 2008 (Continued)

<table>
<thead>
<tr>
<th>Discipline</th>
<th>Assn</th>
<th>Address/Phone/Fax/Web site/e-mail</th>
<th>Journals Published</th>
<th>Certif. Required?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speech-language pathologist</td>
<td>American Speech Language Hearing Association</td>
<td>2200 Research Boulevard Rockville, MD 20850</td>
<td>Journal of Speech and Hearing Research (bimonthly); American Journal of Audiology (three issues per year)</td>
<td>Yes</td>
</tr>
<tr>
<td>Social worker</td>
<td>National Association of Social Workers</td>
<td>750 First St. NE, Suite 700 Washington, DC 20002-4241</td>
<td>Social Worker (quarterly); Health and SW (quarterly); SW Research (quarterly); SW Abstracts (quarterly)</td>
<td>Yes</td>
</tr>
<tr>
<td>Vocational counselor</td>
<td>American Counseling Association</td>
<td>5999 Stevenson Avenue Alexandria, VA 22304</td>
<td>Journal of Counseling and Development (quarterly); Counseling Today (monthly)</td>
<td>Yes</td>
</tr>
<tr>
<td>Child life specialist</td>
<td>Child Life Council</td>
<td>11820 Parklawn Drive, Suite 240 Rockville, MD 20852-2529</td>
<td>One publication, for members only: The Bulletin (quarterly)</td>
<td>Certification not required, but strongly recommended</td>
</tr>
<tr>
<td>Kinesiotherapist (corrective therapist)</td>
<td>The American Kinesiotherapy Association</td>
<td>P.O. Box 1390 Hines, IL 60141-1390</td>
<td>Clinical Kinesiotherapy (quarterly)</td>
<td>Certification not required, but strongly recommended</td>
</tr>
<tr>
<td>Horticultural therapist</td>
<td>The American Horticultural Therapy Association</td>
<td>201 East Main Street, # 1405 Lexington, KY 40507</td>
<td>Journal of Therapeutic Horticulture (annually)</td>
<td>Yes</td>
</tr>
<tr>
<td>Music therapist</td>
<td>American Music Therapy Association</td>
<td>8455 Colesville Road, Suite 1000 Silver Spring, MO 20910</td>
<td>Journal of Music Therapy (quarterly); Music Therapy Perspectives (2/y)</td>
<td>Yes</td>
</tr>
<tr>
<td>Recreation therapist</td>
<td>National Recreation and Park Association</td>
<td>22377 Belmont Ridge Road Ashburn, VA 20148-4501</td>
<td>Therapeutic Recreation Journal (quarterly)</td>
<td>Yes</td>
</tr>
<tr>
<td>American Therapeutic Recreation Association</td>
<td></td>
<td>207 Third Avenue Hattieburg, MS 39401</td>
<td>Annual of Therapeutic Recreation</td>
<td>Yes</td>
</tr>
<tr>
<td>Dance therapist</td>
<td>American Dance Therapy Association</td>
<td>2000 Century Plaza-Suite 108 10832 Little Patuxent Parkway Columbia, MD 21044</td>
<td>American Journal of Dance Therapy (semi-annually)</td>
<td>Yes</td>
</tr>
</tbody>
</table>
McGregor’s Characteristics of an Effective Work Team

<table>
<thead>
<tr>
<th>Table 13.2 McGregor’s Characteristics of an Effective Work Team</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. The atmosphere tends to be informal, comfortable, and relaxed. There are no obvious tensions. It is a working atmosphere in which people are involved and interested. There are no signs of boredom.</td>
</tr>
<tr>
<td>2. There is a lot of discussion in which virtually everyone participates, but it remains pertinent to the task of the group. If the discussion gets off the subject, someone will bring it back in short order.</td>
</tr>
<tr>
<td>3. The task or the objective of the group is well understood and accepted by the members. There will have been free discussion of the objective at some point, until it was formulated in such a way that the members of the group could commit themselves to it.</td>
</tr>
<tr>
<td>4. The members listen to each other! The discussion does not have the quality of jumping from one idea to another unrelated one. Every idea is given a hearing. People do not appear to be afraid of being foolish by putting forth a creative thought even if it seems fairly extreme.</td>
</tr>
<tr>
<td>5. There is some disagreement. The group is comfortable with this and shows no signs of having to avoid conflict or to keep everything on a plane of sweetness and light. Disagreements are not suppressed or overridden by premature group action. The reasons are carefully examined, and the group seeks to resolve them rather than to dominate the dissenter. On the other hand, there is no “tyranny of the minority.” Members who disagree do not appear to be trying to dominate the group or to express hostility. Their disagreement is an expression of a genuine difference of opinion, and they expect a hearing so that a solution may be found. Sometimes there are basic disagreements that cannot be resolved. The group finds it possible to live with them, accepting them but not permitting them to block its efforts. Under some conditions, action will be deferred to permit further study of an issue between the members. On other occasions, when the disagreement cannot be resolved and action is necessary, it will be taken but with open caution and recognition that the action may be subject to later reconsideration.</td>
</tr>
<tr>
<td>6. Most decisions are reached by a consensus, in which it is clear that everybody is in general agreement and willing to go along. However, there is little tendency for members who oppose the action to keep their opposition private and thus let an apparent consensus mask real disagreement. Formal voting is at a minimum; the group does not accept a simple majority as a proper basis for action.</td>
</tr>
<tr>
<td>7. Criticism is frequent, frank, and relatively comfortable. There is little evidence of personal attack, either openly or in a hidden fashion. The criticism has a constructive flavor in that it is oriented toward removing an obstacle that faces the group and prevents it from getting the job done.</td>
</tr>
<tr>
<td>8. Team members are free in expressing their feelings as well as their ideas both on the problem and on the group’s operation. There is little pussyfooting, there are few hidden agendas. Everybody appears to know quite well how everybody else feels about any matter under discussion.</td>
</tr>
<tr>
<td>9. When action is taken, clear assignments are made and accepted.</td>
</tr>
<tr>
<td>10. The chairman of the group does not dominate it, nor does the group defer unduly to him or her. In fact as one observes the activity, it is clear that the leadership shifts from time to time, depending on the circumstances. Different members, because of their knowledge or experience, are in a position at various times to act as resources for the group. The members use them in this fashion and they occupy leadership roles while they are thus being used. There is little evidence of a power struggle as the group operates. The issue is not who controls but how to get the job done.</td>
</tr>
<tr>
<td>11. The group is self-conscious about its own operations. Frequently, it will stop to examine how well it is doing or what may be interfering with its operation. The problem may be a matter of procedure, or it may be a member whose behavior is interfering with the accomplishment of the group’s objectives. Whatever it is, it gets open discussion until a solution is found.</td>
</tr>
</tbody>
</table>

built-in feedback mechanism through which it constantly monitors itself and maintains its effectiveness. When a team is not functioning well, effective function can be developed or restored through the process of team building (6). Team building requires commitments of time and energy, but the rewards of improved patient outcomes and satisfaction of the team members are worth the effort (6,10,11).

A newly formed team, or a team with several new members, faces several major tasks if the team is to function effectively (6,11). The members must build a working relationship and establish a facilitative climate. This is particularly challenging in training atmospheres, since new members are frequently being added or removed for new rotations, and new trainees must learn and adapt to the culture of the permanent team in which negotiated roles have already been established. New teams must work out methods for setting goals, solving problems, making decisions, ensuring follow-through on task assignments, developing collaboration of effort, establishing lines of open communication, and ensuring an appropriate support system that will let team members feel accepted yet allow open discussion and disagreement. In a newly formed team, it is advisable to designate meetings in which members can share personal expectations and develop working policies.

**CONFLICT AND DISAGREEMENT**

Conflicts is a normal, necessary, and not necessarily destructive part of team development (7,10). The potential for conflict is high in health services organizations (12). How it is handled will determine its effect on team objectives and the group process. A good rehabilitation team creates an atmosphere in which members can agree to disagree without making personal accusations or faulting each other’s personalities. In this atmosphere, conflict can be used as a vehicle for growth and innovation.

The interactionist perspective is one current view toward conflict. According to this view, a certain level of conflict is healthy and leads to a group that is viable, self-critical, and innovative. A group can have too little conflict. Without conflict, it may be viewed as harmonious, cooperative, and tranquil, but the team may become apathetic, noninnovative, and nonresponsive to needs for change and may show low productivity. Team members may leave the apathetic team because they are bored. If this occurs, then it becomes the responsibility of team leaders to stir up enough conflict or tension to promote creativity, innovation, and productivity among the team members. The manager who creates conflict must use great skill to see that the conflict does not accelerate to the point where it becomes disruptive, divisive, or chaotic. If conflict is not controlled, then communication suffers, cooperation ceases, and the quality of patient care decreases (7). When conflict repeatedly occurs with no resolution, action must be taken to restore the team’s effectiveness. An appropriate setting for conflict resolution is a team-building session.

**TEAM BUILDING AND DEVELOPMENT**

A group of professionals brought together for the purpose of helping a particular patient or set of patients will not automatically form the most efficient and effective force to accomplish that purpose. Understanding the factors that lead to the development of a team in which members are synergistic in their care of patients is of paramount importance to the physiatrist. To make interdisciplinary rehabilitation teams effective, Rothberg believes the following functions must be performed (13):

- Show/teach team members how to work together and provide sufficient practice time in teamwork.
- Ensure that all members learn, understand, and respect the knowledge and skills of others.
- Develop clear definitions of the roles and behaviors expected of team participants and lessen ambiguities regarding expectations of others.
- Encourage use of the full potential of each member.
- Direct attention to initiation and maintenance of communication and to the breaking down of barriers to interdisciplinary communications.
- Attend to the maintenance of the teams in the same way that other organizations engage in activities that strengthen their cohesion and offer satisfaction to their personnel.
- Acknowledge that leadership should shift as necessary in terms of the patients’ needs.
- Ensure that the person in the leadership role respects the other members, as evidenced by consultation, active listening, and their inclusion in planning.
- Develop an internal system for demonstrating the accountability of each team member to the group, as well as to the institution in which the team practices.
- Develop a process to acknowledge conflict as it arises and to address it in a manner that strengthens the group and its members.

Table 13-3 lists individual characteristics that help one integrate into an interdisciplinary health care team. A professional who is unwilling to accept such roles cannot participate in a significant way in the interdisciplinary health care process.

**TABLE 13.3 Personal Characteristics of Successful Interdisciplinary Team Participants**

1. Accept differences and perspectives of others
2. Function interdependently
3. Negotiate role with other team members
4. Form new values, attitudes, and perceptions
5. Tolerate constant review and challenge of ideas
6. Take risks
7. Possess personal identity and integrity
8. Accept team philosophy of care

New Team Development

Initiating an effective team is a particular challenge. No matter what type of team is being developed, whether formal or informal, multidisciplinary, interdisciplinary, or transdisciplinary (as defined later), or a business group or committee, five basic stages of group development are encountered (Fig. 13-2). These are (a) forming, (b) storming, (c) norming, (d) performing, and (e) adjourning (14).

- During the forming stage, initial entry and identification with the group are the primary concerns. Group members are interested in what the group can offer them and what they can offer the group. During this stage, individuals are usually on their best behavior and may temporarily overlook conflicts for the good of the group.

- Storming is the most difficult stage and is characterized by high emotional tension. The level of trust becomes low during this phase. Team members tend to pressure the rest of the group to accept their preferences. Status and control in the group may become an issue during this phase. Cliques and coalitions may form here, and “hostility and infighting” (14) are common. During this phase, members begin to understand one another’s interpersonal styles and learn to interact within those parameters (15). Team members also attempt to find ways to work toward the team’s goals while they seek concurrently to meet their individual needs.

- The norming phase is a transition to more comfortable and stable interaction and is referred to as initial integration. Balance begins to emerge during this phase, and the team begins to function more as a unit. This initial balance is not completely stable and can give way at any time, but balance and focus are usually reestablished fairly quickly. The newfound harmony usually comes as a great relief after the storming and may become the primary objective of the team for a period of time. Trust improves; however, the group has not yet matured, and the balance between group needs and individual needs is precarious.

- Performing, also referred to as complete integration, is characterized by maturity and a high level of functional efficiency. Complex tasks and disagreements no longer suspend or preoccupy the group. They are quickly resolved, often creatively, and the group moves on toward goal accomplishment. Trust is a key component of the successful team and becomes very high during this phase.

- The adjourning phase occurs when the team disbands. The ability to do this and reconvene in the future as needed is the true test of a team’s integration, maturity, and ultimate success.

The physiatrist, as a team leader, must appreciate that these phases of team development are normal, realizing that to some extent they are inevitable, are acceptable, and represent progress toward the desired goals of an effective and efficient team (14). Leading the team through these tumultuous times takes calm, steady leadership, and the leader must have the ability to remind the members of the group of these normal phases as they pass through them, with the goal of something better resulting eventually. The team must be reminded that complete integration is the goal, but this may not necessarily occur without first going through these other, less effective and efficient phases of negotiation. The leadership qualities defined by Lundberg should be cultivated by rehabilitation team leaders approaching this task (Table 13-4) (16). During the storming and early norming

![Figure 13-2](image-url)
phases, extra care needs to be taken to avoid the appearance of selling out for personal gain during this time of naturally high distrust. Emphasizing the value and importance of each member will help to establish trust and facilitate progress through these tumultuous phases of team development.

**Established Team Complacency**

Another factor that may be detrimental to the team’s effectiveness is complacency among established teams (2,6,7). Whereas much transitional energy is present on initial team development and negotiation of roles, a mature team may lose its edge by accepting routine patterns of behaviors even when change is indicated. A complacent team may be recognized by one or more of the following characteristics: the same members seem to be doing the same thing the same way year after year despite advances in the field; products prescribed are predictable; new members transfer out of the team because of the lack of challenge; there is a fear of, or resistance to, risk taking; and the rewards go to team members with average performance. These characteristics are especially detrimental to the rehabilitation team because external conditions that define the team’s direction and individual patient’s needs are always changing.

Despite similar diagnoses, each patient presents a unique picture; thus, treatment goals and procedures should always vary in some customized way. Treatment techniques should change in response to new research finding, and creativity and problem solving should be important to the operation of a rehabilitation team. Steiner has identified the following characteristics of a creative team: unusual types of people, open channels of communication, interaction with outside sources, openness to new ideas, freedom (i.e., not run as a “tight ship”), an atmosphere in which members have fun, rewards go to people with ideas, and risk taking occurs (2,6).

**Barriers to Communication**

Communication networks associated with rehabilitation are complex, and there are many potential barriers to effective communication (2). Understanding flows of communication, natural barriers to effective communication, and strategies to overcome communication barriers can improve internal communication within the rehabilitation team and health care organization, and thus improve patient care. Communicating well in a rapidly changing health care market, especially external communication with stakeholders outside the rehabilitation facility, can benefit the health care organization in ways that ensure the health, or even survival, of the organization. For example, the rehabilitation organization that communicates well may benefit in terms of being selected as the first-choice provider of rehabilitation services, obtaining contracts at favorable reimbursement levels, or helping to establish favorable regulatory policies (17). Communication skills enhancement is also important for marketing to external stakeholders, as emphasized by CARF (Commission on Accreditation of Rehabilitation Facilities) (18).

An important issue in facilitating rehabilitation team communication is the identification and resolution of barriers to communication. Given and Simmons have identified communication barriers that can interfere with the achievement of treatment goals (11):

- **Autonomy**
- **Individual members’ personal characteristics that may contribute to personality conflicts**
- **Role ambiguity**
- **Incongruent expectations**
- **Differing perceptions of authority**
- **Power and status differentials**
- **Varying educational preparation of the patient care team members**
- **Hidden agendas**

These barriers stem from interpersonal, interprofessional, and practice issues, and these are not intrinsic defects of the team concept (11).

A special barrier to effective communication on rehabilitation teams is the presence of many professional disciplines in rehabilitation, particularly the differing perspectives of professionals with a physical background (e.g., physiatrists and physical therapists) and a psychosocial background (e.g., psychologists and social workers) (19). This adds strength to the holistic assessment and consideration of all aspects of the individual patient’s life needs but can also produce a frustrating set of varying backgrounds, priorities, and initial perspectives that may not be well understood by other team members of a differing discipline. A portion of this barrier can be varying definitions and understanding of rehabilitation-related terminology by different members of the rehabilitation team. A recent review of objective evidence that members of rehabilitation teams have “a disturbing lack of common understanding for some basic rehabilitation terminology” and that “only about half of the personnel providing rehabilitation services are currently sensitive to this issue” (20). The authors suggested several courses of action for this problem: alert rehabilitation professionals that it exists, adopt a standardized rehabilitation glossary for the team, avoid the use of vague terms, define terms operationally, and express descriptions of patients and their progress objectively using standardized functional assessment instruments (20). The use of a communication instrument to help keep the information comprehensible, relevant, and compact can help improve discussion between professionals with different backgrounds (19).

Lack of effective communication can be detrimental to the rehabilitation process and uncomfortable for team members. Time must be designated to maintain an effective team process and to help overcome any existing communication barriers. When a team is functioning suboptimally because of conflict, complacency, or poor communication, the problem can be resolved through the team-building process (6). Dyer cites three prerequisites for conflict negotiation:

1. All parties must agree to come together and work on the problems.
2. Members must agree that there are problems that need to be solved and that solving them is everyone’s responsibility.
3. Members accept the position that the end result is that the team will communicate better, thus enhancing the rehabilitation process (6).

Once these prerequisites have been met, the team identifies the conflicts or barriers in need of resolution. It is important that concrete suggestions be made for the resolution of these problems and that the team agrees on the solutions. This creates a problem-solving session rather than a detrimental process in which the members attempt to determine fault or place blame. Once solutions are agreed on, each member has the responsibility to follow through according to his or her role.

An outside consultant may be extremely helpful, since some signs of poor team function are more easily discerned by an outsider (6). Other symptoms are more easily observed by team members, but an outside consultant can help interpret and resolve these symptoms. The consultant can guide the team away from interpretations of problems that are not likely to lead to resolution, such as erroneously labeling incomplete or inadequate conflict resolution as personality conflict, or placing blame rather than finding effective solutions (6). Consultants can guide the team toward constructive ways to resolve problems such as appreciating the expectation theory, which simply states that negative reactions can be predicted whenever the behavior of one person violates the expectations of another (15). A vicious cycle of escalating conflict can result when the negative reaction itself violates the expectations of the first person. However, because this theory focuses on behavior rather than personality, it allows a greater possibility for conflict resolution. If the parties involved, or even one of the parties, can identify the behaviors that violate expectations, then behaviors can be changed or agreements can be reached. Team members can then reward one another’s behaviors rather than negatively reinforce them (6,15). Appreciating our differences and anticipating how others desire to be treated, including how they prefer to communicate, has been called the Platinum Rule (15). A consultant can help the team learn to sustain healthy communication by developing its own internal mechanisms for problem identification and diagnosis, planning remediation, implementing changes, and evaluating its own results in a healthy feedback loop. The beneficiaries of healthy communication on the rehabilitation team are both the patients and the team members.

It is especially important that health care teams and organizations be able to manage a particular type of conflict—the conflict that arises when something goes wrong. Even in the best-managed organization, things will go wrong. In a health care organization, the result of mistakes can be injury, pain, suffering, or even death. In such cases, the rehabilitation team and the organization also experience distress. There are always ripple effects that can affect multiple stakeholders inside and outside the organization. Excellent communication skills in this situation can contain the damage and may help to redress the consequences, the most difficult step. Healthy communication can help to build trust and even strengthen future relationships with affected stakeholders, and demonstrate a proactive approach toward helping to prevent recurrences of similar mishaps (2,17).

**REHABILITATION TEAM COMMUNICATION METHODOLOGY**

Comprehensive medical rehabilitation requires the interactions of multiple caregivers to provide the breadth of services needed by people with physical and cognitive impairments (3,4,21,22). Patient needs range from acute and chronic medical problems to physical impairments, their complex interactions, and the impact each has on the patient’s psychological, vocational, and social integration. The primary goal of interactions between care providers is communication of the patient’s needs and coordination of his or her efforts in a synergistic manner (23). Physician-initiated prescriptions, referrals, or orders are written communications that are intended to provide for patient needs by initiating the services to be provided by multiple caregivers. The form such written communications take depends in part on the style of interaction adopted by involved professionals. Redundant, poorly coordinated, or incomplete care can occur when a patient’s desires and needs are addressed from multiple vantage points without effective communication and coordination among the different caregiving professionals. Despite the widespread perception that a coordinated team effort enhances the effectiveness of such complex patient care, definitive studies are not available to prove this point. The results of the available studies have varied outcomes related to different measured variables (4,5,7).

Accrediting agencies such as CARF, and, more recently, The Joint Commission (TJC, formerly known as The Joint Commission on the Accreditation of Health Care Organizations/JCAHO), and federal regulations in certain instances, require “interdisciplinary teams” (5,9,24,25), yet many styles of interaction exist that are influenced in part by the practice environment (26). Four general styles of interaction between physicians and other professional caregivers will be discussed: the traditional medical model without a formal team; the multidisciplinary team, which some call the traditional medical model of team interaction; the interdisciplinary model; and the transdisciplinary model. Each model’s advantages and disadvantages are outlined, and its impact on prescriptions, orders, referrals, and treatment plan writing is discussed. These four models of interaction are described in pure form, though features of each are often combined to take the greatest advantage of the benefits each model’s features may offer for a particular practice setting. Effective team dynamics and communication discussed earlier are always important, but they are especially necessary for successful implementation of the interdisciplinary and transdisciplinary models.

**STYLES OF INTERACTIONS**

**Medical Model**

Traditional medical care results in a model in which a physician attends to the patient’s needs. If services of another discipline are desired, that professional is consulted and given either specific or general requests for assistance to meet the needs of the
patient as determined by the attending physician. The quality of the service rendered by the consultant, and thus future consultations, depends on meeting the needs of the patient and the attending physician. The consultant identifying additional needs would usually discuss them with the attending physician before proceeding with the additional treatment, in recognition of the fact that the attending physician may have additional information and insight not available to the consultant. This traditional system results in a clear chain of responsibility that continues to be well respected and is reinforced medicolegally. This traditional autocratic model of leadership, in which the physician assumes an authoritarian role and other team members obey, is not effective in the rehabilitation setting (10). Multiple consultations may result in many professionals doing multiple tasks. Coordination of these efforts by the attending physician or among the involved professionals can often be difficult or incomplete, resulting in less efficient and sometimes redundant patient care. This is one of the major disadvantages to the medical model of patient care (21,23,27).

Rehabilitation professionals have recently favored the concept of “client-centered therapy.” This is not meant to trivialize the patient’s needs, as physicians may suppose, but rather to emphasize the patient as the director and arbiter of the interventions according to the patient’s own desires (28). The term client is used in the place of patient in order to indicate that the role is an active one. The client, his or her caregivers, and the service providers enter into a collaborative relationship with the assumption that the client is the most knowledgeable about his own functional needs. The professionals advise and educate and assist in creating an optimal environment for the client to achieve independence in those areas that the client has identified as being important. Some advantages of this approach include empowerment and decreased dependency for the client, and a truly individualized treatment program, since each patient identifies the issues that he or she wishes to master. However, client-centered care is challenging to provide within the structure of current health care systems, which emphasize professional assessment and medical necessity over patient desires in establishing allowed treatment interventions. It also assumes a fluid interaction that can be problematic within the bounds of the medical model in which a particular attending-defined problem was the cause for the initial referral and for which authorization for treatment was received.

Medical ethics in recent decades has prioritized patient autonomy over attending beneficence (paternalistic actions deemed by the practitioner to be in the patient’s best interest), which is also consistent with patient-centered care (2,29,30). The medical model, as compared with more team interactive models in which the patient is part of the team, is not particularly well suited for patient-directed care because of the additional effort required of the physician in this system. This is because all therapies and consultant plans are coordinated by the attending, and not by the patient. For patient autonomy to have priority, the full weight of patient education, including advice regarding all possible interventions and their respective risks and potential benefits, and recommendations is borne by the attending physician. This has become more difficult in an era of time-limiting managed care. Indeed, much of the decision making over what is best for the patient’s health care is defined by what the patient’s health insurance is willing to cover and is frequently removed from both patient and attending preferences by the coverage certification mechanisms of managed-care systems. These decisions are based more on economic considerations than on considerations of optimal health benefit.

**Multidisciplinary Team Model**

The multidisciplinary team model provides a means for multiple professionals who require frequent interactions to meet and coordinate efforts on a consistent basis. The multidisciplinary model is analogous to the classic pyramid-shaped model of management, which features vertical communication between supervisor and subordinates. It typically remains an attending physician–controlled team in which most interactions are between consultants and the primary attending. Discussion between consulting professionals is held to a minimum or, when necessary, directed by the attending physician. This emphasis on vertical communication (Fig. 13-3) is
evolved from the medical model attending physician’s role and relationship with consultants (26).

Team conferences can be conducted efficiently with such clear lines of authority and control, but lateral communication may suffer (27,31) (see Fig. 13-3). This tendency to impede the free, horizontal flow of communication between the team members is recognized as an obstacle to the optimal use of each participant’s specific expertise and problem-solving skills. This may negate the possible group synergy that can create a product greater than the sum of its parts; or, in clinical terms, a care plan better than any one participant could have developed alone (26,32). The interdisciplinary team model does attempt to improve this communication and enhance group synergism, thus fostering a sense of mutual authority and responsibility (22,31,32).

**Interdisciplinary Team Model**

Interdisciplinary teams benefit from lateral communication flow that occurs as easily as vertical communication in the multidisciplinary team. Because the interdisciplinary model is designed to facilitate such lateral communication, it is theoretically better suited for rehabilitation teams (34–36). The expected norm is group decision-making and group responsibility for developing optimal care planning (31). The problem orientation and ease of flow of lateral communication in the interdisciplinary team processes are similar in function to the project orientation and communication patterns of matrix organization (33,34). The patient is considered part of this planning group and has a central role in the team’s considerations (see Fig. 13-1) (23,37,38). With the emphasis on mutual communication and responsibility, the patient care coordinating conferences may be led by any team member (Fig. 13-4) (27,39). One objective of this model is to allow a freer exchange of ideas and thereby benefit from the group synergy concept (23,40). The interdisciplinary model has been described as a compromise between the benefits of specialization and the need for continuity and comprehensiveness of care (13). Its disadvantages can include considerably less time-effectiveness in completing patient care conferences. In theory, this inefficiency is offset by improved communication and better problem solving. Such teams also require considerable training in the team process, which generally is not received during the years of formal training in the individual disciplines (31). Team communication, development, and conflict resolution have been discussed earlier. This non-patient care training is expensive and does not ensure success (22,31,41). The commitments and personality traits found in the members of a successful interdisciplinary team are similar to those that engender good referral patterns between physicians (see Tables 13-3 and 13-5) (27,41). The physician may be uncomfortable with the team decision-making process because the physician is the one who must usually assume the greatest medicolegal responsibility for the team’s actions and plans. There may be difficulty in having the physician complete the appropriate prescription for such team-generated plans, especially if the plans seem to be different from what the physician recalled or desired. Such conflicts are ideally resolved in team meetings, but delays in completion of the paperwork can jeopardize the optimal patient care.

**Transdisciplinary Team Model**

Transdisciplinary teams, a more recent development, encourage not only communication but cross-treatment among disciplines. This atypical strategy has developed with the central focus on improving patient care through a team approach in which responsibilities are shared as in the interdisciplinary team and also where the normal boundaries of the various health care professions are blurred (42–45,47). These teams have mostly developed out of educational models (21,26,43,44) and have been justified on the basis of economic market forces, and in part by shortages in adequate numbers of therapy professionals (22,47,48). Cross-training, or multiskilling, of available teachers and aides is reported to be very helpful in providing the needed educational services. Such programs, when self-rated against no integration of the disciplines, are rated very favorably (43,45,47). The transdisciplinary team has also found favor with traumatic brain injury teams (26,47). Consistency of information exchange, with the patient intrinsic to co-treatment, is cited as an advantage (43,44,47). Furthermore, the exchange of information among disciplines is highly valued, with therapists or teachers noting expansion of their own professional expertise. A future corps of rehabilitation generalists as the main therapy providers has been predicted and advocated by some (47,48). Whether such informally shared professional knowledge and co-treatment lead to competent therapists in one another’s fields is doubtful. The issues of technically competent care, professional society opposition, state licensure, and qualifications may limit the development of truly transdisciplinary rehabilitation on a widespread scale.

---

**FIGURE 13-4.** In an interdisciplinary team, communication and decision-making mutualism are encouraged. One team member usually acts as the patient care coordinator, but lack of role dominance allows any team member to be eligible for this role.
To further differentiate transdisciplinary from interdisciplinary or multidisciplinary team approaches, attributes from the literature have been defined by Walker and Avant (49). They define five premises that need to be met before the transdisciplinary approach is possible. They are role extension, role enrichment, role expansion, role release, and role support (42).

Role extension is a process whereby one’s own discipline-specific knowledge is continually increasing (50). “Feeling secure in one’s role and knowing that individual contributions are facilitating positive patient outcomes are essential components in creating an environment where each discipline is comfortable collaborating with the other” (51). Role extension involves understanding one’s own discipline well and understanding how it compares to the aspects of other disciplines that can benefit the patient. Further, “having the security in one’s own role leads to the resolution of role, turf, and status issues” (52).

Role enrichment is gaining awareness and knowledge of the other disciplines present on the team (50). Although health care personnel recognize and appreciate the various other professionals, the level of role enrichment to which this defining attribute refers is only achieved through a high degree of collaboration. Team members are encouraged to communicate, collectively plan and implement assessments, discuss results, and develop integrated treatment goals during team meetings (53).

Through team meetings, role expansion is expressed. Each team member from a particular discipline educates the others regarding his or her own expertise (52). Knowledge is shared as team members teach each other to make specific judgments and decisions about interventions that transcend the boundaries of traditional roles (52).

The fourth critical characteristic, role release, is frequently lauded in the literature. Incorporating the skills acquired from other disciplines can help in problematic settings and enhance an individual’s skill set, but such actions tend to blur the traditional discipline boundaries (54). A simple example would be that of a speech-language pathologist (SLP) helping a patient to sit more comfortably for speech therapy by applying proper body mechanics principles learned from a physical therapy colleague. This health care provider assumed responsibility for a needed task by applying techniques learned from another discipline.

Role support, the fifth defining attribute, would be best captured if the physical therapist should walk by at just the moment the SLP was helping the patient to a more comfortable position and gave the SLP feedback on how he or she was instructing the patient about body mechanics. Support of others and feedback about the implementation of a particular skill are the hallmarks of role support (50).

One application of the transdisciplinary team approach is the “arena” or group assessment, in which the patient and all therapists gather at one setting. One primary team facilitator conducts the initial assessment, with all the other disciplines observing, adding, or questioning as needed. This approach limits the number of times the patient is required to answer the same question or demonstrate the same activity or skill, and is considered the best for detecting difficulties because of its thoroughness by several specialties observing the same patient at the same time but from differing perspectives (55).
Though collegiality is enhanced by the transdisciplinary approach, the blurring of roles is discouraged by many specialties and regulatory agencies. The question is whether adequate competence can be developed by the informal training of interdisciplinary teams as compared with the years of formal training during which individual specialists have acquired their competence. Billing rules and regulations can also offer practical problems for appropriate billing for transdisciplinary treatment sessions. As can regulatory requirements such as Medicare’s 3-hour rule for Inpatient Rehabilitation Facilities (IRF), which credits only the actual time spent with the patient, not the time billed by each therapy professional.

**Which Team Approach to Use**

Research is lacking on which of the previous models is most effective. The usefulness of such studies almost certainly will depend on which parameters are examined (i.e., team and patient satisfaction versus outcome). It may be that different models are more effective in different practice environments. The medical care, multidisciplinary team, interdisciplinary team, and transdisciplinary team models can be found in various settings in rehabilitation. The medical care model often is used in a freestanding office practice or in inpatient consultations in an acute-care hospital. This becomes especially true when referrals are made to therapists who are geographically distant or with whom frequent interaction may be difficult. Standing hospital programs that often include nonphysiatric physicians, such as cardiac rehabilitation, pulmonary rehabilitation, geriatrics, prosthetics clinic, myelomeningocele clinic, and the like, may use the multidisciplinary model with one physician in charge. Interdisciplinary teams generally consist of a stable population of health care providers who often can be found in association with specialized units in a comprehensive rehabilitation hospital, unit, or service. Transdisciplinary teams are more common when a stable population of professionals is to provide long-term care for a patient, and cognitive-educational needs are more prominent than intense physical needs.

These models of interaction are meant to enhance communication and thereby coordination of care. The practicing physiatrist may prefer one style over the others but often finds it necessary to communicate with patients and multiple care providers in all these models, or some combination, depending on practice setting. Specificity of orders and the methods in which treatment plans are developed will vary with the treatment and communication models that are adopted.

**Regulatory Organizations**

The two primary regulatory organizations for rehabilitation programs, TJC and CARF, have standards (as noted above) addressing the rehabilitation treatment team and its composition as it relates to the individual needs of the person served. Both use the term *interdisciplinary team*, but they generalize the application of this term in ways that suggest that multidisciplinary teams as defined here would also qualify. CARF uses specific language about the interdisciplinary team and its role and composition. Specifically, CARF indicates in the standard for comprehensive integrated inpatient rehabilitation programs, that the interdisciplinary team includes the following (18):

1. An occupational therapist
2. A physical therapist
3. A psychologist
4. A rehabilitation nurse
5. A rehabilitation physician
6. A social worker
7. A Speech Language Pathologist
8. A therapeutic recreational specialist.

The intent statement for this standard indicates that this dynamic, changing team always includes the person served, the rehabilitation physician, and the rehabilitation nurse. Other team members are determined by the assessment and individual planning processes (18).

TJC, although less specific, implies the importance of the interdisciplinary team approach. They indicate that, “An interdisciplinary team implements and coordinates planned treatment and services” (24). Its intent statement indicates that a “collaborative interdisciplinary approach” helps to achieve optimal outcomes. This plan is developed by the interdisciplinary team based on the needs of, and in conjunction with, the person served.

These regulatory standards do not address the various types of team formats specifically, but they do address the need for the person served to have a team approach that defines the various disciplines that are likely to be needed for the assessment, planning, and implementation of a comprehensive treatment plan for rehabilitation. This developed plan should be goal oriented and should produce positive, measurable outcomes for the person served (18,24). In addition, the regulatory requirements covering admissions to IRF should be considered. Currently, Medicare has specific requirements which must be satisfied prior to a patient’s admission to an IRF. Per Medicare Benefit Policy Manual, 110.1—General (Rev. 1, 10-01-03), A3031001.11.A, HO-211.A, there are two basic requirements that must be met for inpatient hospital stays for rehabilitation care to be covered:

1. The services must be reasonable and necessary (in terms of efficacy, duration, frequency, and amount) for the treatment of the patient’s condition.
2. It must be reasonable and necessary to furnish the care on an inpatient hospital basis, rather than in a less intensive facility such as an SNF, or on an outpatient basis.

In order to meet these criteria, the entire rehabilitation team, as supervised by the physiatrist, must document exactly how the patient has met the established admission criteria, including the above; why that patient must remain under the direct care of a physiatrist; how the patient’s home and community life is practically and significantly benefiting from the established treatment; and why that patient cannot be safely and effectively treated in any, less intense, medical setting. Ideally,
this documentation should begin when the patient is initially consulted for rehabilitation, must be clear upon the patient’s referral to the IRF, and should continue throughout the IRF stay and also help to justify any outpatient therapy needs upon IRF discharge.

Managing a Team Meeting

Although the rehabilitation health care team makes effective use of the combined expertise of all participants, this increased communication comes with a cost, which is time (2). Because much of the communication occurs during scheduled weekly staffings, or patient care conferences, these meetings must be run in an efficient and effective manner. A good team meeting is productive, stimulating, and goal oriented, and it involves creativity and problem solving through facilitated interdisciplinary interaction. The team leader, frequently the physiatrist, has the responsibility of facilitating the meeting, including maintaining the team’s investment, productivity, and efficiency, while not inhibiting effective problem-solving contributions by all participants.

The structure of a team meeting can facilitate an effective team process. The simplest and most popular structure for team meetings is for a member from each discipline to give a progress report for each patient with opportunity for comment by other disciplines, especially when problems arise. This structure can work well because each member tends to focus on the problems particular to his or her discipline. Where cross interactions occur (Fig. 13-5), the observations of other disciplines can be added during the discussion of progress in the common area or task.

Another approach is the problem-oriented agenda, where the problem list for each patient is reviewed in sequence. As each problem on the agenda is discussed, any team member may address his or her role in managing that particular problem. This format tends to promote team ownership to solve problems and to promote the synergistic interdisciplinary approach. It may be easier to keep the meeting goal-oriented because the problem list inherently defines the objectives for the meeting. However, this model requires more skill on the part of the team leader to keep the meetings efficient. For unusually complex problems, it may be prudent to delegate a subgroup to explore resources and possible solutions outside the larger group, and then have the subgroup report on their solutions, to be discussed by all, at the following large group meeting.

Regardless of the model used, it is the team leader’s responsibility to keep the group focused and on task. This involves facilitating discussion, ensuring that ideas are understood, negotiating compromises, and clarifying responsibilities as well as minimizing non–problem-solving discussions (56). The documentation for the meeting should include an action summary of the agreed upon responsibilities with assignments and deadlines (Fig. 13-6).

Another important factor to facilitate communication within a team meeting is the physical setting. A specific meeting time must be designated, and team members must be committed to respecting this time. The meeting room’s size, lighting, and temperature may help or hinder effective group process. Seating should allow face-to-face communication among all members. Adequate physical space and time can help minimize some easily discernible barriers to effective communication (2).

MULTIPICITY OF CARE PROVIDERS

Comprehensive rehabilitation of people with physical and cognitive impairments can be an enormously complex task. The treatment goals are not disorder-specific but rather are patient individualized. The patient’s psychological, religious, vocational, social, and personal needs, desires, and priorities are used to establish and prioritize rehabilitation goals. As an integrated member of the treatment team, the patient

<table>
<thead>
<tr>
<th>Decision</th>
<th>Who is to do it</th>
<th>Date of completion</th>
<th>Date to report progress</th>
</tr>
</thead>
<tbody>
<tr>
<td>Train to do safe independent car transfers with a sliding board</td>
<td>Jan Hoover, PT. Patient’s mother is to bring her car in for use in the training sessions</td>
<td>May 23</td>
<td>May 27 (next scheduled team meeting)</td>
</tr>
</tbody>
</table>

FIGURE 13-5. The appropriate therapy specialties must be chosen for the patient’s specific deficiencies. Examples from physical therapy, occupational therapy, and speech pathology are shown. Team coordination is required to prevent duplication of services and avoid gaps in services needed.

FIGURE 13-6. Example of an action summary.
is expected to make a transition from the passive observer role, common during the acute treatment phase of an injury or debilitating disease, to an active participatory role. This shift in roles requires some patient orientation and education about the team process of evaluating and establishing goals. Patient autonomy should be supported and encouraged. Not only are the patient’s medical needs addressed, but the psychological, social, religious, and vocational impacts of his or her disabling disorder require attention. Planning and facilitating all desired interventions can be accomplished best by each of the disciplines evaluating patient issues from their unique point of expertise. This knowledge from each professional must then be shared and formulated into a cohesive team plan of treatment.

Professional health care providers in the many disciplines common to rehabilitation (see Fig. 13-1, Table 13-1) spend many years acquiring specific skills necessary to effectively assess patient problems that fall within their professional expertise. This often includes the use of test instruments standardized for specific disorders. They learn to formulate and communicate their discipline-specific treatment plans and goals; educate patients, family, and other professionals; apply discipline-specific individualized interventions; and monitor patient progress. From their unique vantage points, they often uncover problems or issues not apparent to others. Although physiatrists have the broadest and most specific training among physician specialties in the issues involved in physical impairments, disabilities, and handicaps, their perspective generally will not be as specific in any one area as that of a therapist who focuses exclusively in that treatment area. Because the interventions required are more than any one provider can reasonably give, the expertise of many professionals is used to divide the needs according to areas of treatment or intervention expertise. To avoid fracturing or neglecting needs and goals that cross disciplines, team communication is used to formulate comprehensive treatment plans (22). The specific capabilities and general roles of the various health care providers commonly found on the rehabilitation team (see Fig. 13-1, Table 13-1) must be learned, well understood, and appreciated by the competent physiatrist for optimal application of these disciplines’ capabilities toward enhancing the function of any particular patient.

Through the medical model, the patient often first encounters a physiatrist, a specialist in PM&R, by referral from another physician. With rehabilitation as the first goal, the physiatrist also will address associated medical problems unique to disabling disorders such as dysfunctional spasticity or optimal pharmacologic bladder management in the patient with spinal cord injury. The physiatrist often initiates the referrals to the necessary rehabilitation professionals. In the inpatient comprehensive rehabilitation setting, the physiatrist may be the sole physician involved.

A physiatrist is knowledgeable in the medical care issues of physically debilitating diseases and trauma, and has the broadest knowledge of the expertise available from each of the other team professionals (see Fig. 13-1). Identifying the areas of patient need addressed by each of these professionals will ensure that appropriate resources are used (examples are shown in Fig. 13-5). The physician who specializes in medical rehabilitation must be aware of the techniques and therapeutic interventions available from each discipline that could have positive impacts on the care of patients and the interventions that are specifically contraindicated.

The appropriately consulted professionals, the patient, and the family form the rehabilitation team. The multitude of potential needs (e.g., medical, physical, psychological, vocational, educational, social, or spiritual) require multiple health care professionals to whom referrals, orders, or prescriptions may be sent. The combined input of the team members should form the basis for a coordinated, comprehensive treatment plan, including methods, goals, and estimates of length of time for completion of each. The treatment plan is dynamic and will require frequent modifications, updates, and revisions as the patient progresses.

**THERAPY REFERRALS AND ORDER WRITING**

Therapy referrals and order writing are based on the initial evaluation (see Chapters 1 and 2). This may or may not include team evaluation input or consensus toward the treatment plan. In the medical model and multidisciplinary team model, orders and treatment plans usually are developed initially by the physiatrist, although they may be modified later as input is received from consultants. In the interdisciplinary team model, a period of evaluation by appropriate disciplines occurs before group development and consensus on the comprehensive treatment plan. Depending on the frequency of team meetings, this may introduce a delay before coordinated team interventions begin. In the transdisciplinary model, group evaluations are the rule, frequently allowing team treatment plans to be developed during the same evaluation and treatment session. Although time efficient, the transdisciplinary team may have less time available for deliberation or complex problem solving, since the concurrent patient treatment is occurring.

Once the problems and treatment goals have been delineated, the process of referrals and order writing can proceed. Often this can be facilitated by organizing problems into functional areas of concern. One organizational scheme is to list problems that are primarily medical in nature first, followed by functional limitations or rehabilitation problems, and then associated social–environmental problems. This allows orders to be broken down into medical, therapeutic, and psychosocial issues, although overlap of problems between these categories is common. Problem-based management of medical issues is now commonplace and integrates well into this scheme.

Interactions with the other professionals providing rehabilitation of complex problems require the physiatrist to possess both a diverse professional knowledge and a highly developed communication skills. The resources available should be applied optimally to obtain the best results for the patient. Through correct identification of the suitable providers, appropriate
referrals or orders communicate in as complete a fashion as possible without limiting creative problem solving or the reciprocal feedback that helps take full advantage of the available expertise. The format of these orders and referrals depends on the practice setting and the model of communication customary in that setting.

Medical Model Referrals and Orders
In the outpatient setting, the practice may involve a well-integrated cohesive team, but more often the rehabilitation medicine specialist is a sole practitioner using community-wide resources. The former type is discussed later under the appropriate team model section. The latter outpatient practice is similar to inpatient physiatric consultations in an acute-care hospital where individual therapy departments may exist without organized teams. In such settings, referrals and orders need to be more specific because frequent verbal feedback and clarification are not as readily available. Recording notes in the hospital chart helps with inpatient coordination and communication, but the chart is less available in the outpatient setting. Written prescriptions help avoid ambiguity and ensure that the patient is being treated as desired (see the section entitled “Written Protocols, Prescriptions, Orders, and Referrals”).

Although treatment recommendations ideally are based on clear physiologic rationale and clinically proven efficacy, such a literature base often is lacking or incomplete. Practitioners tend to be strongly influenced by their own successes and failures, applying lessons learned from past patients to future patients. If the physiatrist does not know which approach to treatment a consulted discipline is taking, then it is unlikely any specific learning will occur from that interaction to benefit the rehabilitation physician’s future management of similar patients. Thus, knowing the particular interventions to be used will help enhance the clinical acumen of the referring physiatrist. Indeed, knowledge of how to prescribe therapeutic interventions in as much detail as is necessary is one measure the American Board of Physical Medicine and Rehabilitation examiners use to determine certification.

Another unique advantage the physiatrist holds is the understanding of how therapeutic interventions affect the pathophysiologic process of disease states. This knowledge may serve as a safety check for his or her patients. The physical medicine aspect of physiatry demands that the physics, biophysics, physiology, and pathophysiology of all prescribed physical modalities be well appreciated. This allows rational prescription of intensity, application methods, sites, duration, frequency, and precautions as warranted for such treatments. The physiatrist must both prescribe appropriate interventions and proscribe inappropriate interventions. It is from such patient safety concerns that legal requirements for physician prescriptions were mandated. Without specific understanding of and concurrence in the treatment strategies used, this safety net of supervision is lost.

The major disadvantage in an extremely precise prescription format is that it may be taken as a signal by the consultant not to think, question, or be creative in addressing the patient’s problems but, rather, merely to perform the services as a technician. This perception may occur even though an order to evaluate the patient has been included, which often is legally required by state rules, whether prescribed or not. To minimize this potential negative impact on professional creativity and problem-solving expertise, requests for feedback should be specifically included. It often is helpful to request phone consultation with therapists, after their evaluation but before they begin treatment, to explore additional options or to convey significant yet sensitive information. If a phone consultation is requested, priority must be given to receiving such calls. Otherwise, this form of feedback and collaboration will not be reinforced enough to be maintained. Phone consultations may allow a better treatment approach to be pursued through modifications of orders by phone while providing the attending rehabilitation physician with the knowledge to adequately coordinate the specific interventions being applied. This will also help provide the order specifics often necessary for reimbursement of therapist-provided services.

Occasionally, team members are found who are unwilling to follow specific treatment orders and who proceed on a treatment plan based on their impression of what is in the patient’s best interests, without consulting the prescribing physician. This violates the trust placed in the consultant and the rules by which one should engender referrals between health care providers (see Table 13-5) (57). Such practices also expose the therapist and patient to medically unsupervised care. If this situation cannot be corrected, the patient, for his or her safety and optimal care, should be redirected to more cooperative and collegial therapy professionals. General orders requesting “evaluate and treat,” sometimes because of lack of better knowledge, tend to promote such cross-purpose practices. Although this takes advantage of the therapist’s creativity and expertise, it may restrict the physician’s ability to supervise or coordinate patient care and may reduce the advantage of multiple professionals’ synergism. Habitual poor physician support has, in part, encouraged some therapy groups to seek independent practices (also called direct access), available in 30 of the 50 states of United States, wherein no medical supervision is required (58). The relationship between a physiatrist and consulted professionals should be collegial and mutually supportive because a domineering, rigid posture serves only to dampen creativity and problem solving among professionals and thus may diminish the quality of patient care (45). Managed care may restrict access to only certain providers. This adds weight to the value of being able to generate rapport, collegiality, and a sense of teamwork with many different rehabilitation professionals in many different settings.

A physiatrist may evaluate patients in the outpatient setting, in which no other professional consultations are required. In this situation, instructions to the patient about medications, side effects, therapeutic exercise, home programs, or simple modalities (e.g., heating pads, ice packs, home traction) are the important communications. Informational brochures and pictographic flyers frequently are available from national advocacy groups (Table 13-6) or can be devised to help reinforce patient
comprehension and therefore compliance with the prescribed home program. Without the benefit of a therapist who interacts frequently with the patient and reports problems regularly, more frequent reevaluations may be necessary to ensure both compliance and progress. Increasingly, case management nurses are involved and may serve as valuable coordination resources and advocates for the patient with third-party payers.

When formal therapy is ordered, treatment referrals should specify any patient education or instruction desired. This includes requesting home programs and follow-up to verify compliance as necessary. Home health care services often terminate treatment because of funding constraints before all goals have been accomplished. Using therapy time before such terminations to provide patient and family training in home programs may significantly extend gains.

**Multidisciplinary Team Referrals and Orders**

In the multidisciplinary team setting, the physiatrist may be a team member, a consultant, or more often one who acts as the primary attending physician. In such a group, the same specificity of orders often is required to initiate therapy but is modified more readily after input from team members at regularly scheduled patient care conferences. Priorities of goals and treatments also are more easily discussed verbally than in the written form. This allows some of the subtleties of comprehensive management to be more effectively conveyed and coordinated. Some degree of coordination between consultants also occurs at multidisciplinary conferences, but the flow of problem-solving creativity is not as free as seen with interdisciplinary team conferences. The format usually consists of consultants giving their reports (i.e., initial evaluations or progress since last conference) and recommendations. Other members ideally monitor the input, but the primary consultant determines the solution to any perceived problems and organizes all the input into a modified problem list and treatment plan. Many treatment modifications are made by verbal orders, with feedback guaranteed by regular meetings. In this setting, it is not necessary to include the time until next physician follow-up or desired frequency and mechanism of follow-up therapy reports based on the original orders.

**Interdisciplinary Team Referrals and Orders**

The format of initial orders to consultants who comprise an interdisciplinary team often is based on requesting a general evaluation, with specific evaluation instruments and the comprehensive treatment plans to be discussed and mutually derived. Occasionally, to avoid delays in initiating therapy, broad categories of intervention also are requested (e.g., "ADL training"). The specifics, however, should be discussed and integrated by the team into a comprehensive individualized patient treatment plan. The shortcomings of a setting with no dynamic, creative problem-solving interactions may persist when the patient’s treatment plan is not specified and discussed but consists only of general orders for initial evaluations and general treatment, or treatment ordered according to a protocol (e.g., "tetraplegic protocol"). Such a generalized order format implies little attention to the patient’s specific and unique needs. It may be countered that therapists adapt the program to this patient’s unique needs, but each team member’s professionalism still functions in isolation, which defeats the advantage of the interdisciplinary team process. Although the mutualism of the interdisciplinary team implies no dominant specialty, it does not exclude any member from the responsibility to be interdependent with the creative input from other members in establishing his or her own specific treatment interventions. This means the physician should consider input from the physical therapist and therapeutic recreation specialist as well as from a consulting psychiatrist before starting antidepressant medications. Territory is both relinquished by all and embraced by all, although, in the end, specific needs and interventions are assigned by the group to those individual team members who have the greatest expertise in that area.

Because the comprehensive treatment plan is not developed solely by the admitting physician, and specific interventions are decided by mutual consensus among all team members, the actual specifics of treatment can be difficult to grasp in the training environment for the PM&R resident physician. This is especially true if the medicolegally required orders remain generalized or if the specific treatment plans are signed much later by the attending physician without the resident necessarily being in the loop. This may occur because only the attending signature is required to meet hospital and third-party payer rules. Much of resident training is funded by inpatient rehabilitation hospitals or units in which the interdisciplinary team process is most often used. It is necessary that not only the generalized order formats be appreciated but also the specifics of therapy interventions and efficacy be prescribed for other, less-integrated settings. If the specifics are not discussed in team meetings, then the full benefit of the interdisciplinary team process is not being realized. Many times, multidisciplinary teams with good mutual interactive skills will be labeled interdisciplinary, but each professional maintains full control of his or her specialty area, with little cross-disciplinary discussion of methods and approaches. Such teams remain multidisciplinary despite labels to the contrary. In this setting, general orders may become accepted but may be counterproductive to the educational process of the physician, the medical supervision

**TABLE 13.6 Patient Education Resources**

<table>
<thead>
<tr>
<th>Name</th>
<th>Address</th>
<th>Phone</th>
<th>Fax</th>
</tr>
</thead>
<tbody>
<tr>
<td>Channing Bete Company</td>
<td>1 Community Place</td>
<td>(800) 628-7733</td>
<td></td>
</tr>
<tr>
<td>South Deerfield, MA 01373-7328</td>
<td>Fax: (800) 499-6464</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Krames/Staywell</td>
<td>780 Township Line Road</td>
<td>(800) 333-3032</td>
<td></td>
</tr>
<tr>
<td>Yardley, PA 19067</td>
<td>Tel: (800) 772-4377</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fax: (800) 722-4377</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

This table lists resources for patient education, including telephone, fax, and address information for each provider.
Transdisciplinary Team Referrals and Orders
All members are involved collaboratively in treatments in the transdisciplinary team approach. Collective hands-on treatment is an excellent method for learning, especially when the information shared among the treating professionals is pertinent and applicable to the moment of care needs. Having another professional depend on your input as you are co-treating is both rewarding and self-affirming. The importance of each member’s beliefs can be emphasized and appreciated in a very practical hands-on experience. The team member does not have to wonder whether a communication about a belief’s importance was received adequately when it becomes essential to the treatment approach integrated between professionals during a co-treatment. Many reports on the transdisciplinary approach emphasize the high ratings such approaches receive by the treating disciplines (43,44,47). Collaboration and coordination of effort certainly are optimized because the disciplines have the opportunity to communicate throughout both the evaluation and the treatment of patients. There may not be a need for a formal team meeting apart from ongoing patient care, except as necessary to provide regulatory or third-party payer required documentation. If all the disciplines in comprehensive rehabilitation could be integrated sufficiently that each felt comfortable treating any patient’s problem, regardless of usual discipline specificity, then a rehabilitation therapy generalist could be envisioned (47). Such a corps of professionals would certainly appreciate problems from a broader perspective and, in an era of shortage, allow for a certain ease of cross-coverage. The greatest impediment to such a development is the necessity by certification laws and ethical considerations of providing skilled, competent professional care (5,22). Billing also can be an ethical dilemma. Should a single patient treated for 1 hour by three co-treating professionals be billed for 1 hour of therapy or three? Should the therapy time be billed on the schedule of the best-paying specialty present or equally divided among the therapy disciplines treating? It remains an open question as to whether co-treatment results in each participating professional becoming more globally competent or merely exposes the patient to less than fully professional care in those areas in which the co-treating caregiver lacks certified expertise and competence.

The advantage of dynamic, fluid, and constantly changing treatment plans that adapt to the patient’s changing status can be a disadvantage when such plans must be developed on the spot. This allows little time for deliberation or consideration of alternatives because treatment must be given promptly. Written plans may not keep up with the current flow of treatment, causing difficulties when a patient must make the transition to other care providers. Indeed, written treatment plans often are generated as a retrospective report of the patient’s past treatments and progress.

If the physician team member is a part of the treatment team, then any concerns about medical safety and medical treatment coordination can be addressed as treatment progresses. This, however, is unusual, with the rehabilitation physician often referring patients to such teams in which the physician will not act as a co-therapist. In such a setting, the more generalized order format may not allow adequate communication of the physician’s concerns and treatment goals, especially because formal team conferences may not be frequent. Because the treating professionals may be addressing areas outside their specific expertise, the comprehensiveness with which all specific therapy issues are addressed may be of concern. In such a setting, it may be to the patient’s advantage for the physician to write more specific and detailed orders to ensure that the breadth of patient issues identified by the physician will be addressed. Some mechanism to allow flexibility in approaches while maintaining direction toward the desired goals is important. Thus, treatment orders or referrals are written in a very goal-directed way, giving guidelines for treatment methods or intervention models. A mechanism for feedback on the approaches taken also is important to enhance the prescribing physician’s supervision and learning experience. Without such interaction, prolonged ineffectiveness or perhaps even contraindicated approaches may result without the benefit of a physiatrist’s professional expertise. Because of the possible professional “dilution” in the transdisciplinary approach, even closer re-evaluation of care by the prescribing professional may be indicated.

WRITTEN PROTOCOLS, PRESCRIPTIONS, ORDERS, AND REFERRALS
Communication
The purpose of physician-generated protocols, prescriptions, orders, and referrals is to communicate patient’s needs adequately and to request services from another professional. In the case of medications, this applies to the prescription sent to the pharmacist. In rehabilitation, it applies to the services requested from the various professionals described earlier. The rehabilitation medicine specialist must use his or her expertise first to decide what the patient’s needs are. The physiatrist’s broad knowledge of the capabilities of various rehabilitation professionals allows selection of the appropriate consulting professionals (see Figs. 13-1 and 13-5). Each professional is then sent a referral or order, depending on the setting. The content depends on the team process in effect in that setting. Referrals in the medical model or multidisciplinary team model should include all elements listed in the first part of Table 13-7 to provide adequate communication (22). The referral should include a mechanism for feedback and possibly an invitation for a formal team meeting apart from ongoing patient care, especially because formal team conferences may not be frequent.
for pretreatment discussion as to the most efficacious plan of treatment. If such an approach is taken, priority must be given to responding to therapist-initiated phone consultations, similar to the courtesy that should be offered to referring physicians (see Table 13-5) (57). Referrals to interdisciplinary team members often are requests for evaluation, with the specifics of treatment to be discussed and agreed on at the next team conference, at which time a written treatment plan and orders will be developed. Referrals that are to be addressed by a transdisciplinary team can be performed in a way similar to the interdisciplinary team, if a post-evaluation conference can be planned; otherwise, they are best left in the detailed format of the medical model. Protocols may be established by the collective consensus of a treatment team, especially for commonly seen disorders that require little variation in approach. These should be agreed on by all treating professionals before their implementation and often require significant development time. A protocol must not become an excuse not to think or customize the treatment approach according to the patient’s unique needs and circumstances. All formats for orders should provide a mechanism for feedback, and the treating team should discuss any changes in treatment required as the patient progresses.

### TABLE 13.7 Seven Requirements for Therapy Referrals

<table>
<thead>
<tr>
<th>Required of All Referrals</th>
<th>Specifics Possibly Needed in Therapy Referrals</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Discipline of therapist to whom referral is directed: may include referral to a specific team</td>
<td>1. (a) if a specific therapist is desired, it may be listed as “Disciplinary Attention: Therapist’s Name”</td>
</tr>
<tr>
<td>2. Diagnosis for which treatment is being requested</td>
<td>(b) If to a specific team, may include each therapy discipline desired or left to be defined, implying referrals will be generated to all disciplines for an initial evaluation; specific therapy orders would then be determined at team conference</td>
</tr>
<tr>
<td>3. Request for evaluation</td>
<td>2. (a) Onset of diagnosis or associated problem</td>
</tr>
<tr>
<td>4. Goals of treatment with expected duration</td>
<td>(b) Include both physical problem and relevant medical diagnosis and onset</td>
</tr>
<tr>
<td>5. Intensity, frequency, and initial duration of treatment desired: may be modified after consultation with therapy professional according to patient’s rate of progress</td>
<td>(c) May include multiple relevant problems and respective underlying diagnoses</td>
</tr>
<tr>
<td>6. Precautions: include other diagnoses or problems that could impede or contraindicate certain interventions, and necessary patient monitoring during therapy with recommended limitations to maintain patient’s safety</td>
<td>(d) Associated psychosocial problems that may affect goals or outcome</td>
</tr>
<tr>
<td>7. Mechanism for feedback, date, and signature: date when physician is to reevaluate the patient, request for phone consultation or progress reports, or implied team staffing if referred to an established team</td>
<td>3. (a) Specify desired testing and reporting mechanisms</td>
</tr>
<tr>
<td>Specifics Possibly Needed in Therapy Referrals</td>
<td>(b) Specify intervals between any retesting or reevaluation desired.</td>
</tr>
<tr>
<td>1. (a) if a specific therapist is desired, it may be listed as “Disciplinary Attention: Therapist’s Name”</td>
<td>4. (a) Detailed short- or long-term goals usually based on problems listed above, or</td>
</tr>
<tr>
<td>(b) If to a specific team, may include each therapy discipline desired or left to be defined, implying referrals will be generated to all disciplines for an initial evaluation; specific therapy orders would then be determined at team conference</td>
<td>(b) Detailed component tasks to be accomplished and sequence desired</td>
</tr>
<tr>
<td>2. (a) Onset of diagnosis or associated problem</td>
<td>(c) Estimated length of time expected to accomplish each of the above goals</td>
</tr>
<tr>
<td>(b) Include both physical problem and relevant medical diagnosis and onset</td>
<td>5. (a) Location of therapy desired (e.g., bedside, department or gym, inpatient, outpatient)</td>
</tr>
<tr>
<td>(c) May include multiple relevant problems and respective underlying diagnoses</td>
<td>(b) Desired duration of each treatment session</td>
</tr>
<tr>
<td>(d) Associated psychosocial problems that may affect goals or outcome</td>
<td>(c) Specific therapeutic modalities desired, with intensity, duration, frequency, and timing with other therapeutic interventions described (see Chapters 11, 13, 17–20)</td>
</tr>
<tr>
<td>3. Specify testing and reporting mechanisms</td>
<td>6. (a) Specifics of monitoring: desired type, frequency, timing during therapeutic interventions, and criteria to discontinue or specifically modify intervention</td>
</tr>
<tr>
<td>(b) Specify intervals between any retesting or reevaluation desired.</td>
<td>(b) Criteria for immediate physician notification</td>
</tr>
<tr>
<td>4. (a) Detailed short- or long-term goals usually based on problems listed above, or</td>
<td>(c) Specific precautions to ensure therapist safety (e.g., infectious, patient behavior, or violence risks)</td>
</tr>
<tr>
<td>(b) Detailed component tasks to be accomplished and sequence desired</td>
<td>(d) Specific modality precautions, given the patient’s diagnoses</td>
</tr>
<tr>
<td>(c) Estimated length of time expected to accomplish each of the above goals</td>
<td>(e) Complete list of patient problems or complete diagnosis list</td>
</tr>
<tr>
<td>5. (a) Location of therapy desired (e.g., bedside, department or gym, inpatient, outpatient)</td>
<td>7. (a) Next physician follow-up date</td>
</tr>
<tr>
<td>(b) Desired duration of each treatment session</td>
<td>(b) Anticipated physician follow-up frequency</td>
</tr>
<tr>
<td>(c) Specific therapeutic modalities desired, with intensity, duration, frequency, and timing with other therapeutic interventions described (see Chapters 11, 13, 17–20)</td>
<td>(c) Possibly desired phone consultation before initiating therapy</td>
</tr>
<tr>
<td>6. Endpoints or decision points and criteria for increasing or decreasing therapy in general, or a specific intervention’s frequency</td>
<td>(d) Desired frequency of follow-up reports and mechanism (written or phone)</td>
</tr>
<tr>
<td>7. Specific education for patient and mechanism to evaluate effectiveness of this teaching</td>
<td>(e) Details desired in follow-up reports</td>
</tr>
<tr>
<td>8. Home program training desired, including timing or criteria for such transition</td>
<td>(f) Third-party reporting required or desired</td>
</tr>
<tr>
<td>9. Nature of home program to be taught: frequency, duration, and intensity of modalities, therapeutic exercise, or other interventions</td>
<td>(g) Criteria to discontinue or duration to continue therapy should physician follow-up not be obtained</td>
</tr>
<tr>
<td>10. Handout materials specifically desired</td>
<td>(h) Date or week desired first to discuss this patient at team conference</td>
</tr>
<tr>
<td>11. Anticipated or desired home equipment training or trials</td>
<td>(i) Frequency of team conferencing desired, especially if different from team’s norm</td>
</tr>
<tr>
<td>12. Duration until therapist follow-up, if any desired, to reverify or enhance compliances with home program and maintenance of gains</td>
<td>(j) Desired emergency health system to be activated should patient decompensate</td>
</tr>
<tr>
<td>13. Complete list of patient problems or complete diagnosis list</td>
<td>(k) Provision of phone, address, and paging numbers to contact the referring physician; mechanism for emergency contact provided</td>
</tr>
</tbody>
</table>
Medical Necessity
A broad understanding of reasonable and medically necessary standards defined by the Centers for Medicare and Medicaid Services (CMS) is useful in identifying rehabilitation therapy needs of a patient. A prescription generated based on a good understanding of these guidelines allows for better medical justification of rehabilitation therapy needs. These standards are set around the definition of “skilled therapy” services as defined by CMS. These are used by rehabilitation therapy professionals to assist in determination of plan of care. CMS guidelines are used by many payer sources besides Medicare and Medicaid. The judgment of the physician and the needs of the patient, however, should always supersede a standard that may not be appropriate in a specific setting. Critical points in these guidelines include:

1. Skilled therapy: CMS states (59), “The services shall be of such a level of complexity and sophistication or the condition of the patient shall be such that the services required can be safely and effectively performed only by a therapist, or in the case of physical therapy and occupational therapy, by or under the supervision of a therapist. Services that do not require the performance or supervision of a therapist are not skilled and are not considered reasonable or necessary therapy services, even if they are performed or supervised by a qualified professional.” Only services that are skilled in nature are valued added to provide benefit to a patient, per this definition. Chronic diagnoses, repetitive services, community/recreational services even performed by a therapist can be viewed as a non-skilled service.

2. Reasonable and necessary: CMS states (60), “While a beneficiary’s particular medical condition is a valid factor in deciding if skilled therapy services are needed, a beneficiary’s diagnosis or prognosis should never be the sole factor in deciding that a service is or is not skilled. The key issue is whether the skills of a therapist are needed to treat the illness or injury, or whether the services can be carried out by non-skilled personnel. There must be an expectation that the patient’s condition will improve significantly in a reasonable (and generally predictable) period of time, or the services must be necessary for the establishment of a safe and effective maintenance program required in connection with a specific disease state.”

3. Maintenance program: CMS validates the necessity of skilled services for setting up a maintenance program for chronic diagnoses and states (61), “The specialized skill, knowledge and judgment of a therapist would be required, and services are covered, to design or establish the plan, assure patient safety, train the patient, family members and/or unskilled personnel and make infrequent but periodic reevaluations of the plan. The services of a qualified professional are not necessary to carry out a maintenance program, and are not covered under ordinary circumstances. The patient may perform such a program independently or with the assistance of unskilled personnel or family members.” Although this allows more frequent interventions for chronic diagnoses, the intervention should be of very short duration with a home exercise program as an ultimate goal.

Reviewing the following prescriptions will allow a better understanding of these guidelines:

EXAMPLE 1
PROBLEM: 74-year-old male s/p CABG, now with generalized weakness and unable to return to golf.
GOALS: Return to golf.
INTERVENTIONS: Physical therapy referral:
Dated: __________
Dx: Deconditioning S/P CABG 6 weeks ago
Date of Onset: __________
Frequency: 2 to 3 times a week × 4 weeks
Goals: Improve strength to be able to return to golf activity
Precautions: Cardiac precautions: HR<_____ ; Systolic BP<_____ ; other:_____
Instructions: Please evaluate and assist with strengthening and endurance. Reduced strength in lower extremities, perform training in standing and walking endurance. Patient to return for follow-up in PMR clinic in 4 weeks.
CRITICAL REVIEW: Per CMS guideline, community/recreational/leisure goals are optimal goals and not functional goals. These are not medically justified goals for therapy intervention.
• Golf activity is a recreational goal: Cannot justify “skilled” physical therapy services
• Endurance goal: Skilled services cannot be justified. Research validates that endurance can improve with time and skilled services are not indicated.
• Skilled therapy can be justified if specific balance and gait deficits that affect daily function are identified in the referral.
• Patient might be suited for cardiac rehabilitation programs that are now generally run by nursing or are private pay programs.

EXAMPLE 2
PROBLEM: A 67-year-old female with acute exacerbation of Parkinson’s. Unable to perform home exercise program (HEP) independently
GOAL: Independent with HEP
INTERVENTIONS: Occupational therapy (OT) and physical therapy (PT) referrals
Dated: __________
Dx: Acute Exacerbation of Parkinson’s
Date of Onset: __________
Frequency: 3 to 4 visits each for OT and PT
Goals: Independent with home exercise program
Precautions: Fall risk
Instructions: Patient with acute exacerbation of Parkinson’s disease. Now with increased difficulty performing existing HEP. Please evaluate for adaptive equipment at home for dressing and address challenges with independent HEP. Please add
new HEP as appropriate. Please address transfer training with adaptive equipment. Please send report to FAX:_____

CRITICAL REVIEW: CMS guidelines approve setting up of maintenance programs for chronic diagnoses.

- Goals are functional, are related to safety and address household function
- Exacerbation refers to acute change in a chronic condition
- Setting up home exercise program covered per "skilled" needs identified by CMS guidelines
- Goals and frequency relate to short duration of intervention

**Therapy Quality Control**

Without appropriate follow-up, the efficacy of any intervention cannot be evaluated or documented. Referrals for therapy, and any prescribed equipment or medication, require follow-up. Often what was desired and presumed to be well communicated by written prescription or referral by the physician is not what occurred with the patient and the therapist. Feedback enhances the accuracy of conveying the correct messages and should be encouraged (32). Receiving feedback helps the rehabilitation physician obtain a broader and more complete perspective on the patient’s required needs and provides information necessary to evaluate progress. A certain degree of feedback is built into the multidisciplinary and interdisciplinary models as well as the transdisciplinary model if the physician is a co-therapist. Otherwise, mechanisms to ensure feedback are essential to good written referrals and orders. Quality control requires not only feedback and follow-up but appropriate corrective actions (Fig. 13-7).

**FIGURE 13-7.** Treatment-planning algorithm.
In the medical model, the attending physician alone bears the responsibility for ensuring feedback, providing appropriate follow-up, and initiating additional contacts as problems arise. The physician’s judgment determines whether outcomes are adequate and whether treatment is brought to closure. Table 13-7 shows examples of referrals and orders that arise from this model.

In the multidisciplinary team, the members give input as to whether goals are being achieved and, if not, why. Problems with progress can also be discussed in team conferences, with solutions derived between the consultants and the primary team-leading physician. Identification of problems and necessary corrective actions becomes a more joint effort as the consultants collectively discuss a single patient’s care.

In the interdisciplinary team, a sense of mutual responsibility demands rapid identification and problem solving by the team for any problem perceived by a member. Solutions are achieved by interactive discussions, brainstorming, and finally mutual consensus over the optimum course to be taken. Problem identification and corrective actions are a team process. It is this process that forms the strength of the interdisciplinary team and occupies the greatest time in patient care team conferences.

In the transdisciplinary team, problems are agreed on so rapidly during treatment that larger-scale or longer-term problems may be ignored or not formally considered. If such a problem currently is affecting treatment, it should be solved by mutual discussion and problem solving. Feedback, discussion, and problem solving occur so fluidly and rapidly that the physician member may be left out until after a solution has been decided and enacted. This diminishes the value of the physician’s expertise to the team and can create conflicts should the physician subsequently believe alternative approaches are indicated. Even if by mutual consent the treating therapist and patient concur, the patient may come to doubt the treatment team’s or physician’s expertise. More likely, the physician’s opinion will be viewed as obtrusive, given the relative paucity of time he or she has been involved in the patient’s team care. Thus, corrective action can be a disjointed process in the transdisciplinary model unless the physician is an integral co-therapist.

Diagnosis-related groups (DRGs), implemented in 1983 to cap Medicare expense growth, created incentives for improved efficiency in terms of hospital costs. This prompted the development of care paths in an attempt to optimize

**TABLE 13.8** Examples of Detailed Prescription for Durable Medical Equipment (DME) and Therapy Generated from a Problem List, Leading to Goals, which Lead to Specific Interventions

<table>
<thead>
<tr>
<th>Problems</th>
<th>Goals</th>
<th>Interventions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peroneal palsy, likely recovery within 6 mo</td>
<td>Eliminate toe catching with ambulation.</td>
<td>DME prescription: Please fit with off-the-shelf polypropylene ankle-foot orthosis set at 90 degrees,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>provide single cane adjusted to patient. Diagnosis: Foot drop secondary to peroneal palsy;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>needed for more than 6 mo; medically necessary to provide safe gait. Date:_____</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Therapy referral: Date:_____</td>
</tr>
<tr>
<td></td>
<td></td>
<td>P.T. X 1 visit</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Dx: Peroneal palsy with footdrop. Goals: Improved gait safety/efficiency and eliminate toe</td>
</tr>
<tr>
<td></td>
<td></td>
<td>catching during gait; prevent heel cord contracture.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Precautions: Mildly decreased sensation on foot dorsum, gait instability, chronic intermittent</td>
</tr>
<tr>
<td></td>
<td></td>
<td>hepatitis B fluid precautions and train in the use of AFO.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Please evaluate and train in the use of AFO; including donning, doffing, and skin checks.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Teach home program for gastro: stretching. Also train in proper gait, including stairs and</td>
</tr>
<tr>
<td></td>
<td></td>
<td>rough terrain, with cane to be used until patient becomes comfortable with balance. I will</td>
</tr>
<tr>
<td></td>
<td></td>
<td>follow up pt. in 2 wk; please send report. Thank you.</td>
</tr>
<tr>
<td>Left shoulder adhesive capsulitis.</td>
<td>Improve shoulder external rotation and then</td>
<td>DME prescription: Please issue yellow, red, and green theraband exercise elastic cords. Diagnosis:</td>
</tr>
<tr>
<td></td>
<td>abduction, while avoiding impingement, to</td>
<td>Deficient rotator cuff on left; medically needed over 6 mo for proper shoulder function. Date:____</td>
</tr>
<tr>
<td></td>
<td>functional and, if possible, normal range.</td>
<td>Therapy referral: Date:_____</td>
</tr>
<tr>
<td></td>
<td></td>
<td>P.T. TIW (3 times per week) for 4 wk.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Dx: Left shoulder adhesive capsulitis. Goals: Improve external rotation by 20 degrees and</td>
</tr>
<tr>
<td></td>
<td></td>
<td>shoulder abduction by 30 degrees in 4 wk.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Precautions: Patient S/P rotator cuff repair 12 wk ago with weak rotator cuff muscles and</td>
</tr>
<tr>
<td></td>
<td></td>
<td>increased impingement risk; pulleys are contraindicated.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Please evaluate and manually stretch left shoulder, including scapular mobilization, avoiding</td>
</tr>
<tr>
<td></td>
<td></td>
<td>shoulder impingement, while applying ultrasound to anterior, then posterior joint to start at</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.5 W/cm² and adjust to highest level comfortable for 10 min at each site. Teach home program</td>
</tr>
<tr>
<td></td>
<td></td>
<td>for self-stretching of external rotation only (abduction will be added later), and rotator cuff</td>
</tr>
<tr>
<td></td>
<td></td>
<td>muscle strengthening with progressive therabands. Pt. will be reevaluated in 4 wk. Please send</td>
</tr>
<tr>
<td></td>
<td></td>
<td>progress report, including shoulder active and passive range of motion measurements every 2 wk.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Please call 123-4567 for any questions. Thank you.</td>
</tr>
</tbody>
</table>

*This level of specificity may not be required when working with a known therapist with whom treatment protocols have been previously discussed, however, the seven elements required from Table 13-7 should be included. This is for both patient safety, billing needs, and to optimize care.*
hospital resource utilization. The CMS implemented a prospective payment system (PPS) in the year 2000 for rehabilitation hospitals, also called Inpatient Rehabilitation Facilities (IRF), which is in part based on the functional independence measure (FIM) plus diagnostic criteria defining categories of patients, called Case Mix Groups (CMGs), which serve to define reimbursement, similar to the DRG system for acute care (48,62,63). Besides increasing the efficiency of cost containment, managed care desires to document maintenance of quality. The ability to measure quality requires some degree of uniformity of approach across institutions for comparisons. Care paths, or protocols for patient treatment, help establish a greater degree of uniformity of patient care. Care paths are ideally targeted for populations for whom 75% are expected to follow a typical course. More customization is typically required in rehabilitation than for other medical care. Care paths are most easily developed for postsurgical patients in whom few complications are expected, such as elective orthopedic cases. In the acute-care setting, 20% to 60% shorter lengths of stay have been obtained by implementing such protocols. Few studies have been done in rehabilitation settings, but one study showed no impact on cost or length of stay when a care path was used versus the usual interdisciplinary team approach. The high level of coordinated care that already exists in comprehensive rehabilitation settings may explain this result (64).

Managed care dominates many markets, as had been predicted (65). Competition for these contracts and the quality improvement management techniques fostered by the TJC and CARF promote the development and implementation of care paths even in rehabilitation, where their value has yet to be firmly established (9,24,64). Care paths or treatment protocols have many labels, including critical paths, practice guidelines or parameters, clinical guidelines, clinical pathways, care maps, flow charts, anticipated recovery plans, or case management (64,66). Care paths for inpatient comprehensive rehabilitation units will become more common as managed care increases. The development of such paths always needs to be customized to each particular setting, ideally by an interdisciplinary team. It is hoped that such protocols will improve efficiency without sacrificing the individualization of care required by most complex rehabilitation inpatients.

The complexity and nonhomogeneity of comprehensive rehabilitation patients makes their study difficult (67,68). Despite poorer outcomes in terms of morbidity and return to community for skilled nursing facility patients compared to IRF patients as a whole, for the populations they serve, few randomized controlled trials of matched patients exist (68). CMS has tried to use evidence-based medicine as its standard, with the absence of evidence often being inappropriately attributed by many of CMS’s fiscal intermediaries as evidence of lack of efficacy (69–71). CMS’s pending fiscal shortfalls are well known, are severe and unresolved, and will continue to pressure denial of care that is not of specifically proven benefit over other perceived less expensive options including no care (67,68,71). This has and will continue to constitute a dynamically changing scene from what is now considered the community standard of care to what the fiscal intermediary chooses to restrict from reimbursement.

**EQUIPMENT PRESCRIPTIONS**

Prescriptions can involve not only needed therapies but also necessary adaptive equipment. The use of adaptive equipment or tools is an important component of the rehabilitation of people with disabilities (72,73). Frequently, appropriate equipment may represent the difference between being functionally independent and requiring the assistance of a caregiver to perform necessary tasks for self-care and ADL. The ultimate goal in rehabilitation is for the patient to achieve the highest level of independence possible. This requires the effective integration of adaptive equipment into the patient’s individualized treatment plan as necessary (74).

The importance of adaptive equipment and devices is illustrated by the diversity of items available to enhance mobility or perform ADL (75). Adaptive devices and equipment are categorized according to the functional skills they are devised to facilitate. These include ADL (e.g., dressing sticks, sock-aides, rocker knife, reachers, etc.), mobility (e.g., gait aids, positioning and transfer equipment, wheelchairs, etc.), communication (e.g., picture boards, Passy-Muir valves, Electrolarynx, computer augmentation, etc.), environmental management (e.g., adaptations, environmental controls, etc.), and leisure and recreation (e.g., adaptive sports and leisure equipment, specialized prostheses, etc.). Concomitant with the availability of an increasing array of adaptive aids is the inconsistent evaluation of the efficacy of these devices (76). Information regarding development and research and printed information resources are variable in quality and content.

Changes in a patient’s functional status through the rehabilitative process introduce another aspect of consideration in adaptive equipment use and length of need. Improved function through time could mean decreased need for adaptive equipment. However, patients with progressive neurologic conditions may require more sophisticated assistive technology if they are to be as functional as they wish to be. The goal would be the attainment of the highest level of independence with the least adaptive equipment. All these factors are part of the process for the selection of adaptive equipment.

Devices and equipment can be very costly. Costs can be minimized by using rentals or 30-day trial usage when utility is uncertain. When items are more unique and customized, costs increase, whereas more commonly used devices can be mass-produced and become more economical. Some costs may be decreasing because some adaptive devices are being increasingly used by a normative aging population to facilitate their self-care and functional performance (e.g., a reacher which is used rather than standing on a step stool to reach items in the closet or pantry, or a long-handled shoe horn that makes it unnecessary to bend over when putting on shoes, etc.).
Frequently, commercial devices must be adapted to the patient to provide for appropriate, individualized fit. Training also is required for the patient to achieve proper and optimum use of the equipment. In addition, the items of adaptive equipment on the market do not remain static; there are ongoing additions to the available repertoire. These changes can represent refinements on current items, new adaptations, or equipment using new technology.

Types of Adaptive Equipment and Devices

**Self-Care and ADL Equipment**

Equipment that aids in self-care and facilitates the ADL constitutes a category that encompasses a wide variety of items. These adaptive devices generally contribute to the patient's independence in performing these activities, with the ADL dysfunction dictating the type of adaptive equipment needed. For example, a stroke patient may be able to perform only one-handed activities, owing to upper-extremity hemiparesis. Therefore, to facilitate eating, a plate guard and adaptive eating utensils such as a rocker knife may be appropriate. The use of the equipment should serve the two purposes of providing the patient with independence in eating a meal and releasing the caretaker from the supervisory or assistive task during the eating process. Other examples for this patient might include a suction-based hand brush to assist in hygiene and a dressing stick to help in performing dressing activities independently with one hand. (See Chapter 9 for additional details.)

The diagnosis, prognosis, and residual function of the patient indicate the types and extent of equipment (77,78). A general rule of thumb is that the more restricted the patient is in performing ADL, the more adaptive the equipment must be and the more extensive will be the collaboration among occupational therapists, physical therapists, and SLPs to provide adequately for the patient. Rehabilitation engineers also can be helpful team members, particularly when the custom adaptations are extensive or complex. Common vendor sources are listed in Table 13-9; databases in Table 13-10; and consumer sources in Table 13-11. Many consumer advocacy groups and their patient information resource materials can now be accessed through Internet searches. Table 13-11 lists some of the more germane Internet addresses.

**Mobility**

There are several types of mobility equipment. There are many items that help to facilitate ambulation. Canes and walkers are frequently used, with each offering a method of facilitating ambulation as well as enhancing ambulation safety. Prostheses or orthoses also are designed to assist in ambulation and mobility. (See Chapters 74–76, and 78 for additional details.)

Positioning adaptations and seating systems can range from the very simple to the intricate. Proper positioning provides the preliminary and necessary basis of posture for the teaching and development of other life skills. Good body support and head control are essential for many activities. Positioning can be simple, such as the placement of a wedge or bolster. It also can involve elaborate seating configurations that require training for proper measurements and construction. The maintenance of skin integrity is another aspect of seating and positioning equipment that must be considered. As with adaptive devices, the more involved the seating requirements are, the more important it is to have representatives from occupational therapy and physical therapy with the physiatrist to formulate the seating requirements. (See Chapter 78 for additional information.)

Transfer equipment to facilitate a patient's movement from one place to another will depend on the amount of assistance required by the patient for the transfer. A transfer board is the simplest item of equipment for transfer. This is used to facilitate safe movement of the patient from a wheelchair to a bed, a chair, or an automobile. The less able the patient is to assist in the transfer process, the more elaborate or complex is the equipment...
Communication

Augmentative and alternative communication systems are a major area of equipment provision. These devices are becoming increasingly more involved and more readily available. They range from low-tech communication boards and pointing devices through complex computer-based communication systems. The physiatrist should collaborate with an SLP and an occupational therapist in the selection of the optimal augmentative communication system for the patient. The SLP would recommend the most appropriate device for communication. For the motorically challenged patient, the occupational therapist would recommend the most appropriate switch- and control-operated devices for the patient. The provision of adapted computer inputs is now an important area of consideration in the lives of people of all ages (79). (See Chapters 15 and 73 for additional information.)

Environmental Management

Various adaptive devices are available to provide assistance in the home (80). Most of the devices are engineered for use in the kitchen. These can include one-handed cutting boards, one-handed sandwich holders, stove overhead mirrors for wheelchair-mobile people to see the top of the stove, and other items. Adaptations also are available for washer and dryer operation. Numerous devices are available for the bathroom to assist in independence and safety. Environmental controls for

<table>
<thead>
<tr>
<th>TABLE 13.10 Databases and Resources for Rehabilitation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ABLEDATA</strong></td>
</tr>
<tr>
<td>Listings of assistive technology products and devices for people with disabilities and seniors. The database includes adaptations as follows: environmental, blind and low vision, deaf and hard-of-hearing, deaf blind, communication, computers, controls, education, housekeeping, orthotics, aids for daily living, prosthetics, recreational, safety and security, seating, therapeutic aids, transportation, workplace, walking, and wheeled mobility. Sponsored by the National Institute on Disability and Rehabilitation Research, U.S. Department of Labor.</td>
</tr>
<tr>
<td>ABLEDATA</td>
</tr>
<tr>
<td>8630 Fenton Street, Suite 930</td>
</tr>
<tr>
<td>Silver Spring, MD 20910</td>
</tr>
<tr>
<td>Tel: (800) 227-0216</td>
</tr>
<tr>
<td>Fax: (301) 608-8968</td>
</tr>
<tr>
<td><a href="http://www.abledata.com">http://www.abledata.com</a></td>
</tr>
<tr>
<td>The Boulevard</td>
</tr>
<tr>
<td>A disability resource directory of products and services for the physically challenged, elderly, caregivers, and health-care professionals</td>
</tr>
<tr>
<td>1205 Savoy Street, Suite 101</td>
</tr>
<tr>
<td>San Diego, CA 92107</td>
</tr>
<tr>
<td>Tel: (619) 222-8735</td>
</tr>
<tr>
<td>Fax: (619) 226-2675</td>
</tr>
<tr>
<td><a href="http://www.blvd.com">http://www.blvd.com</a></td>
</tr>
<tr>
<td>Closing the Gap Solutions (and Annual Conference)</td>
</tr>
<tr>
<td>Assistive Technology in Special Education and Rehabilitation</td>
</tr>
<tr>
<td>Closing the Gap</td>
</tr>
<tr>
<td>P.O. Box 68</td>
</tr>
<tr>
<td>Henderson, MN 56044</td>
</tr>
<tr>
<td>Tel: (507) 248-3294</td>
</tr>
<tr>
<td>Fax: (507) 248-3810</td>
</tr>
<tr>
<td><a href="http://www.closingthegap.com">http://www.closingthegap.com</a></td>
</tr>
<tr>
<td>National Health Information Center</td>
</tr>
<tr>
<td>This health information referral service was established by the Office of Disease Prevention and Health Promotion (ODPHP) within the Public Health Service. The objectives are to identify health information resources, channel requests for information to these resources, and to develop publications in print and electronic form on health-related topics of interest to health professionals, health-related media, and the public.</td>
</tr>
<tr>
<td>National Health Information Center</td>
</tr>
<tr>
<td>P.O. Box 1133</td>
</tr>
<tr>
<td>Washington, DC 20013-1133</td>
</tr>
<tr>
<td>Tel: (800) 336-4797</td>
</tr>
<tr>
<td>Fax: (301) 984-4256</td>
</tr>
<tr>
<td><a href="http://www.healthfinder.gov">http://www.healthfinder.gov</a></td>
</tr>
<tr>
<td>National Rehabilitation Information Center (NARIC)</td>
</tr>
<tr>
<td>Maintains a research library of more than 65,000 documents and responds to a wide range of information requests, providing facts and referral, database searches, and document delivery. Through telephone information referral and the Internet. NARIC disseminates information gathered from NIDRR-funded projects, other federal programs, and from journals, periodicals, newsletters, films, and videotapes. NARIC maintains REHABDATA, a bibliographic database on rehabilitation and disability issues, both in-house and on the internet. NARIC also prepares and publishes the annual NIDRR Program Directory, available in database format from NARIC’s website. Services: Information and referral available by phone, email, fax, letter, website, and walk-in (free); document delivery is 5 cents per page ($5 minimum), and 50 abstracts for $5.</td>
</tr>
<tr>
<td>REHABDATA</td>
</tr>
<tr>
<td>8201 Corporate Drive, Suite 600</td>
</tr>
<tr>
<td>Landover, MD 20785</td>
</tr>
<tr>
<td>Tel: (800) 346-2742 or (301) 459-5900</td>
</tr>
<tr>
<td>Fax: (301) 459-4263</td>
</tr>
<tr>
<td><a href="http://www.naric.com">http://www.naric.com</a></td>
</tr>
</tbody>
</table>
of equipment required for an activity may be indicated. Just as with other previously described devices, coordination would be indicated for equipment provision. The therapeutic recreation specialist may assist in the identification of the leisure or recreational activity that a patient wants to pursue. The occupational therapist or orthotist may assist in the provision and fitting of the appropriate adaptation or splint needed to perform tasks involved in the activity of interest. (See Chapter 9 for additional information.)

**Resources**

Commercial vendors sell equipment and many adaptive devices (81). Most items can be used as purchased; others will need adaptations that customize the devices for the person. At other times, equipment will have to be individually designed and constructed to meet patient needs. These special items are done on an individual basis by bioengineering or orthotics with professional input from occupational or physical therapists or others.

There is wide variation in the cost of equipment, the requirements for documentation for procurement of equipment, and the availability of equipment for patient evaluation trials for efficacy. Some items, such as wheelchairs and EADLS, can be very costly to procure. Ideally, there would be a range of items available within each category (e.g., ADL, mobility, communication) that can be used for patient assessment or training, but often this is not economically feasible. Vendors sometimes can provide equipment for patient assessment and use. Equipment pools are another resource where equipment no longer needed by people or equipment shared among several facilities in an area can be used to assist in the evaluative and training process.

**Prescriptions for Devices**

Any device, be it a simple plate guard for eating or the most elaborately configured electric wheelchair, typically requires a physician prescription for insurance coverage, including Medicare. Even without insurance coverage, a physician’s prescription often allows medically necessary devices to be purchased free of sales tax in many states. Those items that are considered durable medical equipment (DME) require specific information for prescription coverage. Not all items are DME. To avoid confusion and minimize patient expense, it is useful to provide a comprehensive equipment prescription for all medically necessary devices.

The patient’s name and diagnosis are included in the information part of the prescription. Then the initial part of the prescription is the name of the item, the stock number or other identifiers, and, when appropriate, the source of the item. All parts, sizes, adaptations, colors, and so on, are included, as applicable. The justification and rationale for the item are included, as well as the estimated duration of use. A permanent need is documented as “greater than 12 months” for Medicare prescriptions. When expensive devices are being requested, it usually is necessary to receive approval from the third-party payer before ordering such equipment.

---

**TABLE 13.11** Sources for Direct Consumer Adaptive Aids

<table>
<thead>
<tr>
<th>Vendor</th>
<th>Address</th>
<th>Phone</th>
<th>Fax</th>
<th>Website</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adapt-Ability, Inc.</td>
<td>9355 Dielman Industrial Dr. St. Louis, MO 63132</td>
<td>(314) 432-1101</td>
<td>(314) 432-0780</td>
<td><a href="http://www.adapt-ability.org">www.adapt-ability.org</a></td>
</tr>
<tr>
<td>Bruce Medical Supply</td>
<td>411 Waverley Oaks Road P.O. Box 9166 Waltham, MA 02454-9166</td>
<td>(800) 225-8446</td>
<td>(781) 894-9519</td>
<td><a href="http://www.BruceMedical.com">www.BruceMedical.com</a></td>
</tr>
<tr>
<td>Enrichments Catalog</td>
<td>c/o Sammons Preston. 1000 Remington Boulevard, Suite 210 Bolingbrook, IL 60440</td>
<td>(800) 323-5547</td>
<td>(800) 547-4333</td>
<td><a href="http://www.EnrichmentsCatalog.com">www.EnrichmentsCatalog.com</a></td>
</tr>
<tr>
<td>Independent Living Aids, Inc.</td>
<td>P.O. Box 9022 Hicksville, NY 11802</td>
<td>(800) 537-2118</td>
<td>(516) 937-3906</td>
<td><a href="http://www.IndependentLiving.com">www.IndependentLiving.com</a></td>
</tr>
<tr>
<td>North Coast Medical, Inc.</td>
<td>18305 Sutter Blvd. Morgan Hill, CA 95037-2845</td>
<td>(800) 621-9319</td>
<td>(877) 213-9300</td>
<td><a href="http://www.NCMedical.com">www.NCMedical.com</a></td>
</tr>
</tbody>
</table>

Adaptations can also be made in the work environment to facilitate use by those requiring accommodations. Frequently, this is done on an individualized basis. If necessary, site visits are made to determine the needs for either adaptations or equipment. Considerations such as space needed for a wheelchair to turn or to pass through a door, and alterations of table position or height to a comfortable work level are examples of work adaptations.

**Leisure and Recreation**

People may want to pursue old hobbies or develop new leisure and recreational activities. Because of dysfunction, adaptations of home management are another example of assistive devices. Electronic Aids to Daily Living (EADLS), formerly known as environmental control units, can consist of a few simple devices (e.g., for turning on and off a light or television) or more elaborate systems that manage many of the electric functions in a home. These can be introduced in the rehabilitation unit environment (hospital-based EADLS) (79). (See Chapter 73 for additional details.)

Adaptations can also be made in the work environment to facilitate use by those requiring accommodations. Frequently, this is done on an individualized basis. If necessary, site visits are made to determine the needs for either adaptations or equipment. Considerations such as space needed for a wheelchair to turn or to pass through a door, and alterations of table position or height to a comfortable work level are examples of work adaptations.
DME Quality Control

Regardless of what adaptive equipment or device is ordered for a patient, there is a responsibility to ensure that the patient receives training in its use. Also, some items, such as wheelchairs, require fitting. Both the physician and the therapist requesting the equipment are responsible for ensuring that the equipment fits the patient and that the equipment received is operational for its intended purpose. This is especially true in managed-care settings, in which negotiated DME contracts may exist and determine what equipment is available. Timely follow-up is indicated for reassessment of the patient’s use of the equipment and to ensure that the items serve the purpose for which they were initially ordered.

CONCLUSION

The ability to comprehensively define the rehabilitation needs of the patient and to request specific, individualized, and appropriate therapeutic interventions distinguishes the physiatrist from all other medical specialties. To successfully identify and accomplish the goals of rehabilitation, the physical medicine and rehabilitation physician works closely with the allied health rehabilitation disciplines and other medical specialties. Referrals, orders, and equipment prescriptions are basic mechanisms by which the physiatrist requests the participation of the other professions in assessment, planning, and delivery of patient care. The necessary elements and specificity of detail included in the referrals or orders are largely determined by the mode of professional interaction and style of communication developed among the members of the rehabilitation team. A cohesive team with well-developed mechanisms for clear communication among its members can result in an approach to rehabilitation that exceeds the sum of its parts, a concept endorsed by the U.S. Department of Health and Human Services Agency for Health care Research and Quality (82). The rehabilitation medicine physician must be knowledgeable about the treatment strategies used and their potential interactions with the patient’s medical problems. Providing appropriate therapy precautions is a particular responsibility of physiatrists. Effectively written referrals, orders, and equipment prescriptions will fully communicate patient needs, desired inter-ventions, appropriate precautions, and expectations, and provide adequate mechanisms for feedback and quality control.

ACKNOWLEDGMENT

Significant contributions were taken and modified from Chapter 1 by Joel A. DeLisa, Donald M. Currie, and Gordon M. Martin of the 3rd edition of this text as referenced, which also included substantial historical perspective not included here. The prior contributions, now revised, of coauthors to Chapter 13 of the 3rd edition of this text, Mary E. Heye and Mary Nelle Titus and the prior contributions, now revised, of coauthors to Chapter 46 of the 4th edition of this text, T. Russell Nelson, Thomas C. Turturro, and Alison J. Beck, are also acknowledged and appreciated.

REFERENCES

59. Covered Medical and Other Health Services, Chapter 15, Section 220.2, Reasonable and Necessary Outpatient Rehabilitation Therapy Services, Medicare Benefit Policy Manual, Rev. 86, 04-18-08, 165–166.
60. Covered Medical and Other Health Services, Chapter 15, Section 220.2, Reasonable and Necessary Outpatient Rehabilitation Therapy Services. Medicare Benefit Policy Manual, Rev. 86, 04-18-08, 166–168.
This chapter begins by reviewing the history and current status of rehabilitation psychology. This is followed by a description of the direct and indirect services typically provided by rehabilitation psychologists. Frequently encountered psychological measures are described, and their importance for rehabilitation planning is stressed. The final section examines theories of adjustment to disability.

REHABILITATION PSYCHOLOGY: HISTORY AND CURRENT STATUS

History
Rehabilitation psychology is a specialty area within psychology that focuses on the study and application of psychological knowledge and skills on behalf of individuals with disabilities and chronic health conditions in order to maximize health and welfare, independence and choice, functional abilities, and social role participation (1). The field of rehabilitation psychology received initial impetus from veterans returning from the two world wars in the first half of the last century. Physical and occupational therapy trace their roots to World War I. After World War II, the Veterans Administration focused on the psychological needs of the physically disabled, which led to acceptance of psychologists as providers of mental health services. During this same time period, Howard A. Rusk developed the first comprehensive rehabilitation center, which led, with the leadership of others, to the development of physical medicine and rehabilitation as a medical specialty. Thus, the birth and maturation of the disciplines comprising the rehabilitation team have overlapping histories (2,3).

As the number of psychologists working in rehabilitation settings grew, the need for a professional forum arose. In 1949, a special-interest group within the American Psychological Association (APA) was created and in 1958 was granted division status. The Division of Rehabilitation Psychology of the APA provides leadership in formulating federal legislation and in various professional and lay organizations. The Rehabilitation Act of 1973 and the Education for All Handicapped Children Act of 1975 provided mandates for the participation of rehabilitation psychologists in services to the disabled (4). Rehabilitation psychologists promoted passage of the Americans with Disabilities Act of 1990 and its ongoing implementation (5). The injured veterans returning from the wars in Iraq and Afghanistan coupled with the aging population of the United States represent populations with needs that are uniquely addressed by rehabilitation psychologists.

Current Status
Rehabilitation psychologists have struggled with their identity since the field’s inception. Shontz and Wright argued for the distinctiveness of rehabilitation psychology (6). More recently, Glueckauf argued that rehabilitation psychology is subspecialty within the larger domain of health psychology. In this model, rehabilitation psychology is grouped with the subspecialties of clinical neuropsychology, geropsychology, and pediatric psychology (7). To promote this model, an Interdivisional Health Care Committee, with representatives from several divisions of the APA, was formed and continues to actively pursue a common agenda (8).

The identity problem of rehabilitation psychologists arises from the fact that they typically have doctoral degrees in clinical or counseling psychology and enter the field through internship training. A general degree program initially molds their professional identity rather than training in rehabilitation theories, principles, literature, and research; training that is available in only a select number of doctoral programs. Depending on the quality of internship and postdoctoral training, some practitioners may lack training in theoretical models of rehabilitation. There has been ongoing debate within the field of rehabilitation psychology about the “best” training model. The key area of disagreement relates to timing of specialized training (predoctoral vs. postdoctoral), with advocates for both points of view (9–13).

The training and practice of rehabilitation psychologists are changing for two reasons. First, the APA has developed a new model of education in which the graduate-level curriculum is generic, with specialization occurring through postdoctoral training. In 1995, the Division of Rehabilitation Psychology published guidelines for postdoctoral training in rehabilitation psychology (14). These guidelines define what constitutes comprehensive postdoctoral training of rehabilitation psychologists. A comprehensive revision of these guidelines based on the APA Board of Educational Affairs criteria is nearing completion and will be published in Rehabilitation Psychology (15,16). These guidelines will be posted to the Division of Rehabilitation Psychology Web site (www.Div22.org) when complete.

Stiers and Stucky (17) recently surveyed clinical training programs in rehabilitation psychology. Of the 117 sites identified, 94 responded to their survey. They found that the majority of sites did not have a complete rehabilitation focus to their clinical services and did not have faculty with specialty certification in rehabilitation. They found that there...
were only 11 core fellowship sites that had a primary focus on rehabilitation psychology. They concluded that many patients and rehabilitation teams receiving services from psychologists in rehabilitation settings were receiving only partial benefit. They described the unmet needs in rehabilitation psychology training and argued for the development of a “tool kit” that provides standardized curricular materials and program evaluation tools. Second, the American Board of Rehabilitation Psychology (ABRP), www.abrp.org, was established in 1995, with the first diplomate in rehabilitation psychology awarded in 1996. The ABRP provided a comprehensive rationale for specialty definition and competency-based practice standards (15,18). Those awarded the diplomate have typically completed postdoctoral training and are required to have 3 years of experience in rehabilitation psychology, two of which must be supervised. The ABRP is part of the American Board of Professional Psychology (ABPP), an organization of psychologists that accredits subspecialties, similar to the American Board of Medical Specialties for medical specialties.

The goal of the Division of Rehabilitation Psychology is to expand knowledge and seek solutions to problems related to disability and the rehabilitation process. The mission of the organization states:

the Division of Rehabilitation Psychology will: (1) be APA’s voice for the science and practice of psychology as it relates to changes in abilities and social roles arising from illness and disability, (2) be recognized by constituents internal and external to APA for the redefinition of “rehabilitation psychology” as an expanded spectrum of community and clinical services including prevention, rehabilitation, and postacute health care within the context of quality of life across the life-span, and (3) influence the health care marketplace such that rehabilitation psychology services are widely available and accessible and that the quality of life perspective counterbalances both the economic and the medical/curative approaches (19).

The delivery of rehabilitation psychology services is adapting to the corporatization of health care and the reduction of expenditures in the Medicare and Medicaid programs (20–23). The impact of the Prospective Payment System (PPS) on rehabilitation psychologists is uncertain (24). Doctoral-trained rehabilitation psychologists can bill independently under Medicare Part B and, hence, are able to generate revenue in addition to the flat fee provided by the PPS. As a result, rehabilitation psychologists may assume expanded roles in patient evaluation and treatment planning. The shift to health maintenance and treatment, symptom management, health-promoting behaviors, health-related risk-taking behaviors, and overall adjustment to physical illness. Federal reimbursement for these new codes will come from medical rather than from mental health funding and, hence, will not draw from limited mental health dollars. Recent passage of the Mental Health Parity Bill will enhance coverage of mental health services by eliminating the differential in co-payments between mental versus physical health services.

The paradigm shift by the World Health Organization and the National Institutes on Disability and Rehabilitation Research from a biomedical to a social model of disability bodes well for the continued relevance of rehabilitation psychology services and research (27–32). The International Classification of Functioning, Disability and Health (ICF) has been adopted by 191 countries as the international standard for classifying health and disability. As noted by Bruyere and Peterson (33), “the ICF represents a new way for the world to talk about health… The ICF is a classification system that uses a universal, culturally sensitive, integrative, and interactive model of health and functioning that is sensitive to social and environmental aspects of health and disability and covers the entire human life span.” The authors highlight that the concepts and assumptions that comprise the ICF reflect long-standing core values of rehabilitation psychologists including inclusion of and advocacy for people with disabilities in society.

The direct and indirect services described in the following section reflect services typically found within an inpatient medical rehabilitation setting. Similar services are provided in a suitably adapted manner on an outpatient basis. As service delivery shifts to outpatient and community settings, adaptations in these services continue to evolve. Despite the consolidation of health care payer and delivery systems, consumers continue to demand quality and value (34). The value of rehabilitation psychology services is acknowledged by the Commission on the Accreditation of Rehabilitation Facilities through their mandated availability of the rehabilitation psychologist as part of the
rehabilitation team in both acute and subacute rehabilitation facilities. Rehabilitation psychology’s focus on enhancing the quality of life for those with chronic illness and disability remains the central goal despite changes in health care delivery paradigms. Expected changes in the health care delivery system will include universal coverage for all citizens with a focus on prevention, multidisciplinary health care team communication, and measuring outcomes (35,36). Rehabilitation psychology has a long-standing focus on recognizing and maximizing an individual’s strengths and helping them to create a meaningful life despite being confronted with the challenges of living with a disability. Interestingly, a contemporary focus in the field of psychology is on “Positive Psychology.” Positive psychology explores factors that make life worth living, enable people to successfully confront challenges, and facilitate extracting meaning from daily life (37–39). Dunn and Dougherty (40) argue for actively connecting rehabilitation psychology’s foundations and existing strengths to the emerging field of positive psychology, thus enriching both fields of inquiry.

DIRECT SERVICES

The Clinical Interview

The psychologist’s first contact with a patient is pivotal in the development of a therapeutic relationship and may occur before transfer to a rehabilitation unit. The psychologist may visit the patient before the initial interview to explain his or her role. The patient’s expectations of meeting with the psychologist are determined by previous exposure to mental health professionals, communications from other team members, including the physician, and preliminary explanations from the psychologist. The patient’s willingness to interact meaningfully with the psychologist can be strongly influenced by the physician. At the introduction, the psychologist will explain that comprehensive rehabilitation includes help with problematic thoughts and feelings associated with chronic illness or the onset of disability. Frequently, patients are relieved to discover that contact with the psychologist is a routine part of comprehensive rehabilitation.

The initial interview may last an hour or more. Patients with cognitive impairment may be seen only long enough for a general determination of their information-processing capacity and emotional state. Further assessment will await improvement in their cognitive status or contact with an informed family member. The length of the initial interview with non-cognitively impaired patients depends on the complexity of the medical or social issues. There are two major goals for the initial interview. First, a comprehensive history of the patient’s social background is obtained. Table 14-1 lists frequently asked biographical questions. These data provide insight into previous learning experiences that may affect rehabilitation-related attitudes and behaviors. Second, the psychologist attempts to understand the disability as the patient sees it, with the most critical question being, what is the meaning of the disability for the patient and his or her life? The foundation for a meaningful therapeutic relationship is laid, in part, by taking sufficient time to elicit the patient’s perspective. The patient often faces a medical situation that he or she does not fully comprehend. Anxiety and fear often block the reception and communication of information between the patient and rehabilitation team members, especially physicians. The opportunity to have one’s perspective, including cognitive and emotional aspects, aired in a supportive and clarifying manner is often therapeutic in

<table>
<thead>
<tr>
<th>TABLE 14.1 Psychosocial Information Sought During Initial Interview</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data on Family of Origin</td>
</tr>
<tr>
<td>Names, ages, occupations, marital status, and residence of</td>
</tr>
<tr>
<td>parents and siblings</td>
</tr>
<tr>
<td>Religious training</td>
</tr>
<tr>
<td>Stability of family during early development</td>
</tr>
<tr>
<td>History of major mental disorder in immediate and extended</td>
</tr>
<tr>
<td>family, including any history of sexual abuse, chemical</td>
</tr>
<tr>
<td>dependency, suicide, or psychiatric hospitalization</td>
</tr>
<tr>
<td>Relevant Patient Information</td>
</tr>
<tr>
<td>Educational background and school achievement</td>
</tr>
<tr>
<td>Occupation and vocational history</td>
</tr>
<tr>
<td>Avocational activities and hobbies</td>
</tr>
<tr>
<td>History of adjustment to structured environments, such as</td>
</tr>
<tr>
<td>school, work and military service</td>
</tr>
<tr>
<td>Social adjustment, including any previous arrests, chemical</td>
</tr>
<tr>
<td>dependency treatment, or psychiatric diagnosis</td>
</tr>
<tr>
<td>Prior association with hospitals and health care</td>
</tr>
<tr>
<td>Preinjury stresses at the time of injury</td>
</tr>
<tr>
<td>Most difficult loss the patient has had to adjust to previously; success in that task</td>
</tr>
<tr>
<td>Prior associations with people who have a disability</td>
</tr>
<tr>
<td>Family Structure</td>
</tr>
<tr>
<td>Names, ages, and quality of relationship with spouse and children</td>
</tr>
<tr>
<td>Background of dating and sexual relationship with current spouse</td>
</tr>
<tr>
<td>Marital adjustment</td>
</tr>
<tr>
<td>Understanding the Patient’s Perspective</td>
</tr>
<tr>
<td>The patient’s understanding of the cause and probable course of the disability</td>
</tr>
<tr>
<td>The patient’s initial thoughts at the onset of the disability (if traumatic)</td>
</tr>
<tr>
<td>The patient’s most pressing immediate concern</td>
</tr>
<tr>
<td>How well the patient thinks he or she is coping with the situation</td>
</tr>
<tr>
<td>The patient’s perception of how the disability will change lifestyle, including relationships, vocational future, and self-concept</td>
</tr>
<tr>
<td>The patient’s understanding of the behavioral expectations in the rehabilitation unit compared with those in the acute-care unit of the hospital</td>
</tr>
<tr>
<td>The degree to which the patient’s sense of self-esteem or employment is related to physique or physical skills</td>
</tr>
<tr>
<td>The patient’s comfort in meeting with a psychologist</td>
</tr>
<tr>
<td>Techniques used to cope with stressful events in the past</td>
</tr>
<tr>
<td>Techniques used to get and maintain a sense of control over the environment</td>
</tr>
</tbody>
</table>
A personality test conventionally refers to a measure of personal characteristics such as emotional status, interpersonal relations, motivation, interests, and attitudes. Personality inventory development has generally relied on one or more methods including content validation, empirical criterion keying, factor analysis, and personality theory. Personality measurement has generated controversy over two issues. The first concerns the stability of personality traits across situations as opposed to the situational specificity of behavior (45). The second issue involves the degree to which a given personality characteristic reflects a merely transitory state rather than a stable underlying trait. Anastasi and Urbina (46) provided a thorough overview of these issues and of other psychological measurement concepts including norms, item analysis, reliability, and validity. Elliott and Umlauf (47) caution that the insensitive use of personality measures with individuals who have medical symptoms or limited physical abilities can produce misleading results. Johnson-Greene and Touradji (48) provide a contemporary review of the role of personality in rehabilitation outcomes and adjustment to disability. The most frequently used personality inventory designed to measure psychopathology is the Minnesota Multiphasic Personality Inventory-2 (MMPI-2) that is now available in a new form entitled the Minnesota Multiphasic Personality Inventory-2 Restructured Form (MMPI-2-RF). Two personality measures of nonpathologic or "normal" personality relevant to rehabilitation are the Revised NEO Personality Inventory (NEO-PI-R) and the Strong Interest Inventory (SII).

**Minnesota Multiphasic Personality Inventory-2**

The MMPI-2 is the revised version of the MMPI, the most widely used and thoroughly researched objective measure of personality (49–53). The MMPI-2 is composed of statements describing thoughts, feelings, ideas, attitudes, physical and emotional symptoms, and previous life experiences. In general, the material included on the MMPI-2 is usually covered in a clinical interview. However, factors of privacy, time-savings, and the clinical relevance of the items have ensured its acceptance in health-care settings.

The MMPI-2 was originally designed to yield information about personality factors related to the major psychiatric syndromes. The 567 true-false questions are grouped into ten clinical scales (Table 14-2) that continue to reflect important aspects of personality despite their obsolete psychiatric titles. The items composing each scale were determined statistically. An item was included only if a carefully diagnosed group of patients (e.g., those hospitalized for depression) answered that question in a manner statistically different from that of other carefully diagnosed groups of patients (e.g., schizophrenics) and from the normal standardization sample. The MMPI-2 standardization sample consisted of 2,600 persons from several states chosen to reflect several national census parameters, including minority group status. This system of item selection (i.e., empirical criterion keying) fostered the inclusion of subtle items, items that make the MMPI-2 less easily faked when compared with other personality measures.

The ten clinical scales are interpreted with the aid of four validity scales (see Table 14-2). These scales provide information on the client’s response style such as literacy, cooperation, malingering, comprehension, and defensiveness. The empirical nature of the inventory has permitted construction of special scales. For example, there are scales to help predict rehabilitation motivation, headache proneness, and tendencies toward the development of alcoholism. Additionally, there are extensive MMPI norms on persons with specific diagnoses such as multiple sclerosis and spinal cord injury (SCI) (54). Norms are reported as standard scores with a mean of 50 and a standard deviation of 10. A score of 65 or greater is the point at which the normal and the pathologic groups are most reliably discriminated. However, depending on the scores obtained on the validity scales, this "cut-score" may be adjusted by the trained interpreter.

The MMPI-2 requires an eighth-grade reading level and is intended for adults 18 years of age and older. A version of the MMPI entitled the MMPI-A is intended for use with adolescents (55). The MMPI-2 requires about 90 minutes to complete. Although many computerized scoring services are available, this does not obviate the need for interpretation by an experienced psychologist. A variety of factors—including...
race, socioeconomic status, unique family circumstances, ethnic background, and physical disability—may distort the MMPI-2 profile (56).

An important goal in the development of the MMPI-2 was to preserve sufficient item continuity with the original MMPI to allow for the generalizability of the voluminous MMPI research literature to the MMPI-2. Unfortunately, Humphrey and Dahlstrom reported that profiles generated by the MMPI and the MMPI-2 on the same subjects are too frequently at variance to be able to consider the two instruments interchangeable (57). Hence, it would be an error to assume that the MMPI clinical research literature can be uncritically generalized to the MMPI-2 for all patients. Moreover, many of the original criticisms of the MMPI remain problematic for the MMPI-2 (58). In particular, the MMPI was originally written to aid physicians in the medical management of patients who were believed to have psychological factors intertwined with their presenting complaints. MMPI and MMPI-2 normative studies on medical patients (59) consistently reveal elevations of three to six points on clinical scales 1, 2, and 3 when compared to the MMPI normative samples. Appropriate interpretation of the MMPI-2 with medical and rehabilitation patients requires knowledge of these normative biases (60).

Minnesota Multiphasic Personality Inventory-2 Restructured Form

The MMPI-2-RF represents the continuing evolution of the MMPI (61). The original ten MMPI clinical scales were problematic because of their intercorrelations, item overlap, and heterogeneous item content. To address these problems, a set of nine Restructured Clinical Scales (RCS) was derived (Table 14-3) by identifying the major distinctive “core” components of each clinical scale. Tellegen et al. (62) provided

<table>
<thead>
<tr>
<th>Scale</th>
<th>Number</th>
<th>Name</th>
<th>Number of Items</th>
<th>Elevated Scores Suggest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q</td>
<td>Cannot say</td>
<td>567</td>
<td>A large number of items have not been answered, a possible indication that the patient is resentful or is uncomfortable with ambiguity</td>
<td></td>
</tr>
<tr>
<td>L</td>
<td>L scale</td>
<td>15</td>
<td>An effort to create the impression of being a person with high moral, social, and ethical values</td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>F scale</td>
<td>60</td>
<td>The questionnaire has been invalidated by some factor, including lack of comprehension, poor reading ability, mental confusion, a deliberate desire to fake psychiatric difficulty, random marking of responses, or scoring errors</td>
<td></td>
</tr>
<tr>
<td>K</td>
<td>K scale</td>
<td>30</td>
<td>A self-view of being well adjusted, capable, and confident, which, at higher scale elevations, is likely to represent a denial of the true state of affairs</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Hypochondriasis</td>
<td>32</td>
<td>Undue concern with bodily states and preoccupation with possible symptoms of physical illness</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Depression</td>
<td>57</td>
<td>Depression, sadness, pessimism, guilt, passivity, and tendency to give up hope easily</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Hysteria</td>
<td>60</td>
<td>Psychological immaturity, self-centeredness, superficial relationships, and frequent use of denial in everyday life, and a tendency to develop physical symptoms under stress</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Psychopathic deviate</td>
<td>50</td>
<td>Assertiveness and nonconformity at moderate elevations; angry rebelliousness and noncompliance with social mores at extreme levels</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Masculinity-femininity</td>
<td>56</td>
<td>The degree of identification with roles and interests traditionally assigned to the sex opposite that of the respondent</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Paranoia</td>
<td>40</td>
<td>Interpersonal oversensitivity and irritability about motives or behavior of others, and at extreme elevations, suspicious thinking similar to that of people with paranoid personality traits</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Psychasthenia</td>
<td>48</td>
<td>General feelings of anxiety, with excessive rumination about personal inadequacies</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Schizophrenia</td>
<td>78</td>
<td>Feelings of detachment from the social realm, extending to frank mental confusion and interpersonal aversiveness</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Hypomania</td>
<td>46</td>
<td>Talkativeness, distractibility, physical restlessness, and, at times, impatience, irritability, or rapid mood swings</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>Social introversion</td>
<td>69</td>
<td>Social introversion and a lack of desire to be with others</td>
<td></td>
</tr>
</tbody>
</table>
data documenting the improved psychometric properties of the RCS that included improved reliability, reduced scale intercorrelations, and improved convergent and discriminant validity. Once the RCS were devised, 23 additional specific problem scales were constructed that reflected distinctive components of the old clinical scales that were not captured in the new scales with titles such as malaise, suicidal/death ideation, anger proneness, substance abuse, shyness, and ideas of persecution. In addition, the MMPI-2-RF has a total of eight validity scales and three higher order scales, entitled emotional/internalizing dysfunction, thought dysfunction, and behavioral/externalizing dysfunction. The MMPI-2-RF has a total of 338 items and 50 empirically revised scales. The MMPI-2-RF requires 35 to 50 minutes to complete and is written at a sixth-grade reading level.

**NEO Personality Inventory Revised**

The NEO Personality Inventory Revised (NEO-PI-R) reflects the culmination of decades of personality research that concludes that personality traits can be summarized in terms of the so-called five-factor model (63,64). The NEO-PI-R was designed to measure the five major dimensions or domains thought to be central to normal adult personality. These dimensions are entitled Neuroticism (N), Extraversion (E), Openness (O), Agreeableness (A), and Conscientiousness (C). Each domain scale has six facet scales resulting in a total of 35 scales on the inventory. Neuroticism refers to a general tendency to experience negative affect, self-consciousness, poor coping, irrational ideas, feelings of vulnerability, and difficulties controlling cravings and urges. Extroversion relates to interpersonal warmth, gregariousness, assertiveness, activity, excitement seeking, and the tendency to experience positive emotions. Openness pertains to depth of imagination, aesthetic sensitivity, intensity of feelings, preference for variety, intellectual curiosity, and independence of judgment. Agreeableness includes the characteristics of trust, straightforwardness, altruism, methods of handling interpersonal conflict, humbleness, and sympathy for others. Finally, conscientiousness encompasses the characteristics of competence, organization, reliability, achievement striving, self-discipline, and deliberation before acting.

NEO-PI-R item construction was based on rational-theoretical methods. Item selection was determined by internal consistency and factor analytic data. The scale’s 240 items are rated on a five-point continuum from “strongly disagree” to “strongly agree.” The inventory is designed for adults, 17 years of age and older. The inventory requires a sixth-grade reading level and about 45 minutes for completion. There are separate adolescent norms for those less than 21 years old. The NEO-PI-R has a self-report (Form S) and an observer rating form (Form R). This dual-form feature is unique among personality measures and is especially relevant to rehabilitation research. Also noteworthy, the NEO-PI-R items do not contain references to physical abilities or sensations that might distort a physically disabled subject’s responses.
The NEO-PI-R is a reliable and valid measure. Validity has been established through numerous studies that correlate the NEO-PI-R with other measures of personality. All these correlations have been in accord with theory and expectation (65). There are two limitations to the NEO-PI-R. The NEO-PI-R assumes an honest respondent; no subtle items or validity scales are provided. In addition, it remains unclear to what degree the subject’s current mood (state) may impact the response to test items that describe long-standing personality characteristics (trait).

The NEO-PI-R has been used in a number of rehabilitation studies. Rohe and Krause administered the initial version of the NEO-PI-R to males with traumatic SCI 16 years after injury (66). The subjects scored lower on the scales of conscientiousness, assertiveness, and activity; they scored higher on the scales of excitement seeking and fantasy when compared with the adult male normative sample. Scales reflective of negative affect were not elevated. The subjects’ reduced conscientiousness and non-elevated neuroticism scale scores have negative implications for adherence to rehabilitation regimens but positive implications for long-term coping abilities. A subsequent study on the same sample correlated personality and self-reported life adjustment. Rohe and Krause discovered that elevated scores on the depression scale were associated with poorer outcomes, whereas elevated scores on the scales dealing with warmth, positive emotions, actions, and values were associated with superior outcomes. The authors suggested that personality assessment may be an invaluable aid in predicting long-term outcomes and can help delineate those individuals most vulnerable to negative outcomes, perhaps indicating a need for more careful follow-up and supportive services for this subgroup (67). In a seminal study using the NEO-PI-R, the authors assessed whether SCI is associated with personality change by comparing the personality test scores of identical twins, one of whom sustained an SCI. The authors found no significant differences between the NEO-PI-R scores of the twins with SCI and their non-SCI twins (68).

The Strong Interest Inventory

The SII is traditionally considered a measure of vocational interests; however, research has supported its use as a valid non-pathology-oriented measure of personality (69). First published in 1927, the SII is one of the most thoroughly researched, highly respected, and frequently used psychological tests. The SII was revised in 2004. Improvements include more focus on business and technology occupations, number of items reduced from 317 to 291, representative sampling of ethnic, racial, and demographic workforce diversity, and an expanded number of scales. The SII asks the respondent to indicate their level of interest in a five-point Likert scale ranging from strongly like to strongly dislike for occupations, subject areas, activities, leisure activities, and people. The last section asks the respondent to rate their possession of nine personal characteristics. The test requires 35 to 40 minutes to complete and is written at an eighth- to ninth-grade reading level (70).

The General Occupational Themes, one of the four types of scales on the SII, are based on trait theory as derived by Holland (71). Holland drew on factor-analytic studies of personality and interests to produce a typology of six basic personality types. These types are titled realistic, investigative, artistic, social, enterprising, and conventional. Rohe and Athelstan administered the SII to a national sample of persons with SCI (72). Contrary to previous research, they found unique personality characteristics associated with persons having SCI of traumatic onset. These characteristics included an interest in activities requiring physical interaction with things, such as machinery, and a disinterest in activities that require intense or complex interaction with either data or people. Malec used the Eysenck Personality Inventory with people having SCI of traumatic onset and discovered a pattern of personality characteristics congruent with that found in Rohe and Athelstan’s study (73).

Rohe’s review of the literature suggested that when a disability is of traumatic onset and secondary to the individual’s behavior, statements in the literature about the lack of a relationship between disability and personality characteristics appear to be inaccurate (74). He noted that the previous literature either used pathology-oriented measures (e.g., MMPI) or studied individuals whose disability was not the result of trauma associated with their behavior. An additional study sought to determine if those personality characteristics associated with people having SCI would change after years of living with the disability. The data indicated that personality characteristics remained constant over an average of ten years (75). Rohe and Krause conducted a follow-up study to the aforementioned personality stability study. They found that males with traumatic SCI displayed marked consistency in personality characteristics over an 11-year follow-up period (76).

The MMPI-2, MMPI-2-RF, NEO-PI-R, and SII represent four measures of personality relevant to clinical rehabilitation settings. These measures can help answer diagnostic and management questions. For example, a patient’s unwillingness to comply with requested medical interventions or the structure imposed by the hospital environment may be discerned with the use of personality measures. Knowledge of such personality characteristics can help prevent ill-advised interventions and can create a treatment environment designed to maximize patient compliance. While clinical experience attests to the usefulness of such assessment, systematic empirical research on this topic is a pressing need.

Mood

Problematic mood, especially depression, in patients is one of the most common concerns among rehabilitation team members and a frequent reason for psychological assessment. Depression is an imprecise term used to describe an affective state that ranges from “being down” to major depressive disorder (MDD). The incidence and prevalence of MDD in rehabilitation populations has been the focus of significant research and debate. Assessing depression immediately after disability onset is complicated by such medical and environmental factors as sleep disruption, pain, and decreased appetite. Two brief, psychometrically robust and helpful instruments are the Patient Health Questionnaire-9 (PHQ-9) (77) and the Beck Depression Inventory-Fast Screen (BDI-FS).
reliability and validity data. The correlation between the 7-item inventory requires 5 minutes to complete and has acceptable range while a score of 8 is average for persons with MDD. The suicidal thoughts. A score of 4 falls in the mildly depressed past failure, loss of pleasure, self-dislike, self-criticalness, and statements. The item groupings focus on sadness, pessimism, focused items. The BDI-FS consists of seven groups of four components of depression, systematically excluding the somatically based symptoms or normal bereavement. Bombardier found accurate diagnosis, the clinician reviews the responses to ensure that any positive somatic responses are not due to medically based symptoms or normal bereavement. Bombardier et al. (79) studied the PHQ-9 with an SCI sample 1 year post injury. The average score was 5.48, with 11.4% of the sample meeting the criteria for probable MDD with a mean score of 18.1. The PHQ-9 appears to be a promising instrument for detecting MDD in rehabilitation populations.

**Beck Depression Inventory-Fast Screen for Medical Patients**

The BDI-FS for medical patients (80) is a self-report inventory that screens for depression in adults and adolescents. The items were extracted from the original 21-item Beck Depression Inventory-II (81) and focus on the cognitive and affective components of depression, systematically excluding the somatically focused items. The BDI-FS consists of seven groups of four statements. The item groupings focus on sadness, pessimism, past failure, loss of pleasure, self-dislike, self-criticalness, and suicidal thoughts. A score of 4 falls in the mildly depressed range while a score of 8 is average for persons with MDD. The inventory requires 5 minutes to complete and has acceptable reliability and validity data. The correlation between the 7-item and the 21-item inventories is 0.91. A recent study using the BDI-FS with persons with multiple sclerosis supported its concurrent and discriminative validity (82).

**Intellectual Ability**

Intellectual ability tests provide a summary score that serves as a global index of a person’s general problem-solving ability, frequently referred to as an IQ or intellectual quotient score. This summary score is validated against a broad criterion such as scholastic achievement or occupational success. Although such tests are constructed of a number of subtests that sample facets of intellectual functioning, they are usually weighted toward tasks requiring verbal ability. The most frequently encountered measure of intellectual ability is the Wechsler Adult Intelligence Scale, III (WAIS-III) (83).

**Wechsler Adult Intelligence Scale, III**

The WAIS-III is the second revision of the WAIS originally published in 1955. The WAIS-III was standardized on a normative sample of 2,450 ‘normal adults’ from ages 16 to 89, divided into 13 age groups. They were chosen to be representative of the US population as determined by the 1995 census update. The sample was stratified by gender, educational level, ethnicity, and region of the country. On the WAIS-III, items from earlier versions of the test that had become obsolete because of societal changes or bias against ethnic groups were altered or eliminated and the artwork was updated. Other improvements include extending the range of possible scores both downward and upward (full-scale IQ range = 45 to 155), decreased reliance on timed tests, and addition of a new subtest that requires conceptual reasoning using abstract symbols (matrix reasoning). The WAIS-III must be administered by a trained examiner and requires 75 to 90 minutes for completion.

The WAIS-III consists of 14 subtests, one of which is optional. Table 14-5 lists in their order of administration these subtests and what they measure. On the WAIS-III, six subtests are used in the computation of the verbal IQ, and five subtests determine the performance IQ. In general, the WAIS-III has deemphasized speed of responding by reducing the number of items with time bonus points, eliminating the picture arrangement subtest, and replacing the timed object-assembly subtest with the nontimed matrix reasoning subtest. Object assembly remains available when time and motoric response capabilities are not an issue. All WAIS-III subtest scores are corrected for age and standardized with a mean of 10 and a standard deviation of 3. The full-scale IQ is calculated by averaging scores obtained from the verbal and performance IQs.

The WAIS-III incorporates earlier research of Cohen, who discovered that three underlying factors accounted for most of the WAIS-R test variance (84). The Cohen factors are now codified as three of the four WAIS-III “index scores.” The names of the index scores and the subtests used to compute them are verbal comprehension (vocabulary, similarities, information), perceptual organization (picture completion, block design, matrix reasoning), working memory (arithmetic, digit span, letter-number sequencing), and processing speed (digit symbol-coding, symbol search). As with the traditional IQ

---

**TABLE 14.4 PHQ-9 Items**

| Over the last 2 weeks, how often have you been bothered by any of the following problems? |
| Response options: (0) not at all, (1) several days, (2) more than half the days, (3) nearly every day |
| 1. Little interest or pleasure in doing things. |
| 2. Feeling down, depressed, or hopeless. |
| 3. Trouble falling or staying asleep, or sleeping too much. |
| 4. Feeling tired or having little energy. |
| 5. Poor appetite or overeating. |
| 6. Feeling bad about yourself—or that you are a failure or have let yourself or your family down. |
| 7. Trouble concentrating on things, such as reading the newspaper or watching television. |
| 8. Moving or speaking so slowly that other people could have noticed? Or the opposite—being so fidgety or restless that you have been moving around a lot more than usual. |
| 9. Thoughts that you would be better off dead or of hurting yourself in some way. |

Reprinted from the PRIME-MD® Patient Health Questionnaire. ©1999, Pfizer Inc.
TABLE 14.5 The 14 Subtests of the Wechsler Adult Intelligence Scale-III

<table>
<thead>
<tr>
<th>Test Number</th>
<th>Number of Items</th>
<th>Task Description</th>
<th>Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Verbal Scale</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Vocabulary</td>
<td>35</td>
<td>Define the meaning of words presented both orally and visually</td>
</tr>
<tr>
<td>4</td>
<td>Similarities</td>
<td>14</td>
<td>Explain the way in which two things are alike</td>
</tr>
<tr>
<td>6</td>
<td>Arithmetic</td>
<td>14</td>
<td>Solve arithmetic problems presented in a story format without using pencil or paper</td>
</tr>
<tr>
<td>8</td>
<td>Digit span</td>
<td>14</td>
<td>Listen to and orally repeat increasingly long lists of numbers, with separate lists presented in forward and reverse directions</td>
</tr>
<tr>
<td>9</td>
<td>Information</td>
<td>29</td>
<td>Answer oral questions about diverse information acquired through living in the United States</td>
</tr>
<tr>
<td>11</td>
<td>Comprehension</td>
<td>16</td>
<td>Explain what should be done under certain circumstances and why certain social conventions are followed; interpret proverbs</td>
</tr>
<tr>
<td>13</td>
<td>Letter-number sequencing*</td>
<td>7</td>
<td>Order sequentially a series of numbers and letters initially presented in a specified, random order</td>
</tr>
<tr>
<td>Performance Scale</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Picture completion</td>
<td>20</td>
<td>Determine which part is missing from a picture of an object or scene</td>
</tr>
<tr>
<td>3</td>
<td>Digit symbol-coding</td>
<td></td>
<td>In a timed code substitution task, pair nine symbols with nine digits</td>
</tr>
<tr>
<td>5</td>
<td>Block design</td>
<td>9</td>
<td>Reproduce a two-dimensional design on a card by using 1-inch block whose sides are red, white, or red and white</td>
</tr>
<tr>
<td>7</td>
<td>Matrix reasoning</td>
<td>26</td>
<td>Continuous and discrete pattern completion, classification, reasoning by analogy, serial reasoning</td>
</tr>
<tr>
<td>10</td>
<td>Picture arrangement</td>
<td>10</td>
<td>Arrange sets of cards containing cartoon-like drawings so that they tell a story</td>
</tr>
<tr>
<td>12</td>
<td>Symbol search*</td>
<td>60</td>
<td>Visually scan a target group and a search group of symbols, then indicate whether the target symbols appear in the search group</td>
</tr>
<tr>
<td>14</td>
<td>Object assembly*</td>
<td>4</td>
<td>Properly arrange four cut-up cardboard figures of familiar objects</td>
</tr>
</tbody>
</table>

*aSupplementary subtest.
*bOptional subtest.

scores, index scores have a mean of 100 and a standard deviation of 15. They are sometimes reported in lieu of the traditional IQ scores. The subtests of digit symbol coding and symbol search were added to the WAIS-III to measure the proposed fourth factor, speed of information processing, as suggested by Malec et al. (85). The test exceeds all standards of reliability and validity. Reviewers have been uniformly impressed with the quality of the resulting instrument (86,87). The WAIS-IV was published in late 2008 (88). This new version has reduced emphasis on motor demands and timed performance, enlarged visual stimuli, reduced testing time, updated norms, and improved psychometric properties. Many of these improvements reflect specific requests by rehabilitation psychologists working with populations who have compromised sensory and motor functions.

Given the emotional significance of IQ scores, psychologists usually convert both IQ scores and discussions about them into either percentiles or classifications (Table 14-6). When the physician is confronted with questions about test results from patients, the use of either percentiles or classifications
is recommended. Measures of intellectual ability help the physiatrist set appropriate expectations about the rate and complexity of learning legitimately expected from the patient. They also serve as the cornerstone for determining the presence of organic brain dysfunction and provide guidance for post-dismissal vocational planning.

**Academic Achievement**

A frequently overlooked but nonetheless important factor within rehabilitation settings is academic achievement. Reading and mathematics achievement are of particular concern not only during inpatient rehabilitation but also for longer-range educational and vocational planning. The patient’s reading level is a potential limiting factor in tasks ranging from filling out hospital menus to incorporating ideas presented in patient education materials. The average reading level in the United States is at the sixth grade, the level required to read a newspaper. Patient education materials, however, often reflect the reading levels of the professionals who devise them. As the patients’ reading level falls below the national average, progressively greater reliance on oral instruction and audiovisual materials becomes necessary. Patients are often expected to use mathematics when recording fluid intake and taking correct dosages of medications. Two frequently used measures of reading and mathematical achievement are the Wide Range Achievement Test-4 (WRAT-4) and the Woodcock-Johnson Psycho-Educational Battery-III.

**Wide Range Achievement Test-4**

The WRAT-4 is the current edition of the WRAT (89). The test provides assessment of reading, spelling, and mathematics achievement. Reading achievement is based on the subtests Word Reading and Sentence Comprehension. The Word Reading subtest requires correct pronunciation of single words. The Sentence Comprehension subtest requires that the subject demonstrate understanding of the sentence by supplying one or two words that complete the meaning of the sentence. These two subtests yield a reading composite score. The addition of the Sentence Completion subtest effectively addressed the criticism that the WRAT-4 was not truly measuring reading achievement. The Spelling test requires correct spelling of words presented by the examiner in the context of a sentence. Mathematics achievement has two subtests entitled Oral Math and Mathematics Computation. The Math Computation subtest is administered to children aged 8 and older and ranges from simple addition to algebra. The Math Computation subtest has a time limit of 15 minutes. The WRAT-4 is for individuals from 5 through 94 years of age. There are alternate forms, entitled “Blue” and “Green.” The entire test can be administered in 30 to 45 minutes, and results are presented in the form of standard scores, percentiles, stanines, grade equivalents and Rasch ability scaled scores. The WRAT-4 is reliable, and the stratified national sample of 3,000 individuals is representative of the U.S. population in terms of age, gender, ethnicity, geographic region, and educational attainment.

**Woodcock-Johnson Psycho-Educational Battery-III**

The Woodcock-Johnson Psycho-Educational Battery-III (WJ-III) contains 42 subtests subdivided into two parts: cognitive ability (20 subtests) and academic achievement (22 subtests). The WJ-III replaces the Woodcock-Johnson-Revised, which received critical acclaim for ease of administration, reliability, validity, and normative sampling procedures. Improvements of the current version include expanding the number of subtests, co-norming the cognitive ability and the academic achievement tests on a large nationally representative sample of 8,818 subjects from 100 geographically diverse U.S. communities, providing age norms from 2 to more than 90 years, and decreasing testing time through focusing testing at the subject’s ability level. Efforts were made to minimize the use of manipulatives and eliminate items that might be culturally insensitive to individuals with disabilities. Both improvements are helpful when assessing those with disabilities (90).

The academic achievement subtests are grouped into five broad curricular areas: reading, oral language, mathematics, writing, and academic knowledge. The achievement tests are further organized into groupings that aid interpretation. Reading achievement is determined by averaging three subtests requiring about 15 minutes to complete. These three subtests assess letter-word identification, reading fluency, and passage comprehension. Thus, the resulting score reflects diverse aspects of reading including reading decoding, reading speed, and reading comprehension. Achievement scores are reported in age- and grade-normed percentiles, and reading ranges from easy to difficult. Mathematics achievement is based on three subtests consisting of calculation, math fluency, and applied problems that require about 20 minutes to complete. The three remaining achievement clusters are of less relevance to inpatient rehabilitation but may be useful in post-dismissal planning. The WJ-III is an exceedingly well-constructed test and has become the standard to which other tests of ability and achievement are compared. The cognitive test portions of the WJ-III can serve as an alternative to the WAIS-III when a broad-based, well-constructed measure of intellectual ability is desired (91,92). The WJ-III Normative Update (WJ-III NU) provides updated norms based on the 2005 United States Census Projections. The WJ-III NU renews and refines the validity of both the cognitive and achievement subtests.

<table>
<thead>
<tr>
<th>IQ Score</th>
<th>Percentile Range</th>
<th>Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>130 and above</td>
<td>98 or greater</td>
<td>Very superior</td>
</tr>
<tr>
<td>120–129</td>
<td>91 to 97</td>
<td>Superior</td>
</tr>
<tr>
<td>110–119</td>
<td>74 to 89</td>
<td>High average</td>
</tr>
<tr>
<td>90–109</td>
<td>25 to 73</td>
<td>Average</td>
</tr>
<tr>
<td>80–89</td>
<td>9 to 23</td>
<td>Low average</td>
</tr>
<tr>
<td>70–79</td>
<td>2 to 8</td>
<td>Borderline</td>
</tr>
<tr>
<td>69 and below</td>
<td>&lt; 2</td>
<td>Extremely low</td>
</tr>
</tbody>
</table>
Neuropsychological Assessment

Individuals with cognitive dysfunction represent one of the largest groups receiving rehabilitation services. For many, the deficits are transient, whereas for some, cognitive deficits are permanent and not only will complicate the learning of independent living skills but also will determine future living arrangements, social interactions, and vocational prospects. In both situations, the rehabilitation psychologist is frequently asked to clarify the nature and type of cognitive deficits. This section describes three helpful screening tests of cognitive status: the Montreal Cognitive Assessment (MoCA), the Galveston Orientation and Amnesia Test (GOAT), and the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS). Next, two approaches to neuropsychological assessment—the quantitative approach, as exemplified by the Halstead-Reitan Neuropsychological Battery (HRNB), and the qualitative approach, as exemplified by the Boston Process Approach—are described.

Screening Measures of Cognitive Status

The physician frequently encounters patients of doubtful potential for a rehabilitation program. A question that often arises is whether the patient shows evidence of organic brain dysfunction. Mental status scales can be divided into three groups: lengthy scales with multiple content, abbreviated scales with one or two items per cognitive area, and short scales of ten or fewer items. Lengthy scales often require 1 hour to administer and may not provide enough information to warrant the time required. Short scales tend to focus on orientation questions, ignoring the diversity of cognitive abilities. Nasreddine et al. devised the MoCA to generate helpful information in a minimum of time (94).

The Montreal Cognitive Assessment

The MoCA is the first brief screening measure specifically designed to detect mild cognitive impairment (MCI). MCI is conceptualized as an intermediary state between normal cognitive aging and dementia (95). Although the Mini-Mental State Examination (MMSE) remains in frequent use (96), there have been many criticisms of the measure, including its lack of sensitivity to MCI (97). The MoCA was derived from observation of cognitive domains frequently impacted in those with MCI. The MoCA is a one-page, 30-point instrument that requires 10 minutes to administer. The MoCA tasks include five-item verbal learning and delay recall, clock drawing, cube copy, alphanumeric sequencing, phonemic fluency, verbal abstraction, sustained attention on a tapping task, serial subtraction, digits forwards and backwards, three-item confrontation naming, sentence repetition, and orientation to time and place. The MoCA was found to have high sensitivity and specificity for detecting MCI. When compared to the MMSE on samples of patients with Parkinson’s disease and memory impairment, the MoCA was found to be a superior instrument (98,99). The MoCA is available for free (www.mocatest.org) and has been translated into 17 languages. Persons scoring 26 or higher are in the non-impaired range, 26 to 21 are in the MCI range, and below 21 are in the Alzheimer’s disease range. The MoCA has much to offer the rehabilitation inpatient setting where rapid and reliable screening for subtle cognitive dysfunction is frequently requested.

The Galveston Orientation and Amnesia Test

The GOAT was developed by Levin et al. (100) and measures amnesia and disorientation after head injury. The scale consists of ten questions that focus on temporal orientation, recall of biographical data, and memory of recent events. The patient can obtain a maximum of 100 points; the final score is computed by subtraction of the number of error points. The GOAT was standardized on a group of 50 young adults (median age 23) who had recovered from mild closed head injury, usually consisting of a momentary loss of consciousness. Scores below those received by all members of the control group (<65) are designated impaired; scores between 66 and 75 are designated borderline; and those above 75 are considered normal. The greatest scoring difficulty occurs where points are assigned for the patient’s accuracy in recalling events before trauma. More recently, Bode et al. (101) rescored the items as dichotomies and identified three strata of posttraumatic amnesia (PTA). Equal-interval measures of PTA were created, and a self-scoring key was developed to assess PTA more efficiently. PTA is defined as the time during which the GOAT score is 75 or less. Validity data were generated by a comparison of the length of the PTA with the variables of initial neurologic impairment and scores on the Glasgow Outcome Scale. In both cases, the GOAT score readily discriminated according to the severity of head injury. Scaling recovery of cognitive function in the noncomatose patient permits meaningful discussion with the family and rehabilitation team members. Most important, attempts at more involved neuropsychological assessment usually prove nonproductive until the patient consistently obtains scores of 70 or greater. Once a score of 70 is achieved, neuropsychological test data usually are reliable for further rehabilitation and postdischarge planning.

The Repeatable Battery for the Assessment of Neuropsychological Status

The RBANS is a recently developed, brief battery that fills an important niche between brief screening measures of cognitive status and comprehensive neuropsychological assessment. The RBANS is an individually administered instrument designed to measure attention, language, visuospatial-constructional abilities, and immediate and delayed memories (102). The test requires about 30 minutes to complete. The RBANS is intended for use with adults between 20 and 89 years. Norms are based on a stratified, nationally representative sample of 540 healthy adults. The RBANS subtests are subdivided into five cognitive domains: immediate memory, visuospatial reasoning, language, attention, and delayed recall. A total scale score is
also generated. The 12 subtests are entitled list learning, story memory, figure copy, line orientation, picture naming, semantic fluency, digit span, coding, list recall, list recognition, story recall, and figural recall. An alternate form is available. The test is a collection of item types from other well-known tests that are used to assess memory and cognitive function. Although there are some limitations to the depth of the test manual and associated data, the RBANS has acceptable reliability and validity coefficients (103). The RBANS is useful for a broad-based but intermediate level of screening for cognitive deficits in acute-care settings and for tracking improvements or declines in cognitive function over time.

Goals of Neuropsychological Assessment

The field of neuropsychological assessment has come full circle since 1935 when Halstead established his laboratory (104). Halstead observed the behavior of individuals with brain damage and then developed psychological tests to measure the characteristics that he observed. Thus, the initial goal of assessment was a better understanding of brain-behavior relationships, particularly in individuals with brain impairments. Halstead’s assessment methods quickly proved their worth when used by experienced clinicians, by reliably and validly diagnosing brain damage, and localizing malfunctioning regions of the brain. Their use as a neurodiagnostic instrument gained prominence and is now routine. Their diagnostic validity has been shown to be equal to that of neurodiagnostic techniques in use before the introduction of computed tomography and magnetic resonance imaging (105). As brain imaging technology has improved, however, the importance of neuropsychological assessment has returned to the original goal of describing brain-behavior relationships.

The field of neuropsychological assessment is increasingly focusing on the development of new methods for assessing rehabilitation potential, functional competence, and valid cognitive remediation procedures for patients with brain damage (106–110). For example, Sherer et al. (111) demonstrated how early brief cognitive assessment combined with severity of head injury indicators can more reliably predict long-term employment outcome. Miller and Donders (112) demonstrated that neuropsychological assessment during acute rehabilitation significantly improved the prediction of educational outcomes in children with traumatic brain injury 2 years later. Unfortunately, most neuropsychological tests were designed with diagnosis, not prediction or remediation, as their major goal. The tests were not constructed to determine which combinations of cognitive abilities are minimally necessary for survival in complex environments. Heaton and Pendleton (113), in an early comprehensive review of the neuropsychological literature, lamented that prediction of everyday functioning is a largely ignored topic of research in a field still dominated by diagnostic issues. Optimal prediction about the satisfactory matching of persons with environments will require new approaches to the measurement of environments and careful attention to the ecological validity of neuropsychological tests (114,115).

Halstead-Reitan Neuropsychological Battery

The HRNB consists of three batteries (i.e., child, intermediate, and adult) (116). The adult version consists of Halstead’s original five tests of seven variables selected for their ability to distinguish adults who have frontal lobe dysfunction. Halstead’s first graduate student, Ralph Reitan, established his own laboratory and added measures of aphasia, sensory-perceptual integrity, grip strength, and sequential visual scanning, collectively entitled “Allied Procedures.” In addition to the HRNB, a complete battery usually includes the WAIS-III and a measure of personality (e.g., the MMPI-2). The results of the HRNB can be presented as the Halstead Impairment Index, which is not a meaningful indicator of brain damage by itself but must be understood in the context of other test data that have been analyzed using inferential methods. Reitan described these inferential methods as follows:

Level of performance. The individual’s score is compared with that of a criterion group and the normality of the score is determined. Statements describing the amount or degree of a specific attribute are provided.

Pattern of performance. Variations of scores within and between tests can be analyzed for specific strengths and weaknesses. Specific behavioral deficits or pathognomonic signs. Behaviors occurring only with brain damage (e.g., anosmia, hemianopsia, and hemiplegia) can be detected.

Comparison of performance of the right and left sides of the body. Measures of motor, sensory, and sensory-perceptual functions from one side of the body are compared with those of the opposite side. Significant discrepancies may reflect brain dysfunction and can help rule out competing explanations of poor performance on more complex neuropsychological tests.

In rehabilitation, as opposed to general medicine, the issue of diagnosing the presence of brain damage is of reduced importance. Brain damage is frequently the criterion for entry into a rehabilitation unit. Rehabilitation team members are more concerned with the degree to which the patient will be able to understand and profit from rehabilitation services or function in their home environments. As noted in the introduction to this section, a lengthy test battery may not be physically possible because of recent onset of impairment or because of shortened lengths of stay in acute rehabilitation. In these cases, the rehabilitation psychologist may administer select portions of the standard HRNB or use tests that place fewer demands on the patient such as the RBANS. A comprehensive battery of neuropsychological tests may be administered on an outpatient basis to more fully assess cognitive function and provide guidance on such issues as supervision, return to driving, and ability to return to work.

The HRNB is not without its limitations. For example, Halstead’s original cutting scores were based on the performance of young patients, with age-graded norms becoming available recently (117). Patients with minimal education may spuriously score in the impaired range. Finally, the battery’s ability to
Although neuropsychological assessment has proved effective in detecting cognitive dysfunction, translating test results into ecologically valid rehabilitation recommendations has proved more elusive.

The Boston Process Approach to Neuropsychological Assessment

The Boston Process Approach to Neuropsychological Assessment (BPA) is the designation given to neuropsychological assessment that focuses on the manner in which the patient produced his or her response rather than concentrating solely on test norms and patterns of scores. The three goals of the BPA are to (a) understand the qualitative nature of the behavior being assessed, (b) reconcile descriptive richness with the reliability and validity of quantitative tests, and (c) relate the behavior assessed to neuropsychological theory.

The BPA emphasizes developmental psychology’s distinction between “process and achievement” to understand cognition in individuals with brain damage. Systematic observation of the problem-solving strategies used by patients during standardized testing is the hallmark of this approach. This method allows both a quantitative assessment and a dynamic description of the information-processing style of each patient.

The BPA uses a core set of tests with supplementary tests to clarify problem areas and confirm clinical hypotheses. The core set consists of frequently used tests that assess functions in six cognitive domains including intellectual and conceptual, memory, language, visuoperceptual, academic achievement, self-control, and motor control. The BPA is concerned with the extent the patient gives priority to processing low-level detail or “featural” information versus higher-level “configural” or “contextual” information. Patients with impairments in contextual processing often have difficulties organizing their behavior.

Both the quantitative and qualitative approaches to neuropsychological assessment have proponents. The process approach is perhaps more compatible with the rehabilitation emphasis on practical and functional improvements. Given its focus on qualitative observation, the process approach provides potential insights regarding the most effective cognitive rehabilitation methods to use for remediation of cognitive deficits. Lezak et al. noted that the use of fixed neuropsychological test batteries is waning in favor of a flexible choice of tests. Increasingly, the norm is for neuropsychologists to flexibly borrow from both the quantitative and qualitative traditions to design briefer batteries that are appropriate to a particular assessment question and setting. Moreover, the current trend toward reduced lengths of initial rehabilitation often precludes administration of a comprehensive neuropsychological test battery. A common alternative is to perform a brief assessment of cognitive function before discharge from inpatient rehabilitation followed by comprehensive assessment weeks to months after discharge. This brief initial assessment provides a “benchmarking” of current levels of cognition and can provide guidance about the amount of supervision needed for the patient. A report by the American Academy of Neurology confirmed the efficacy of neuropsychological assessment.

Chemical Use Assessment

Background

In the mid-1980s, Rohe and DePompolo highlighted the failure of rehabilitation professionals to assess and intervene in the domain of chemical health among those with disabilities. In 2002, Basford, Rohe, and DePompolo found significant improvements in chemical health screening for those with disabilities but a continued failure to provide staff training on the issue of chemical health assessment and intervention. Cardoso et al. report continued inadequacies in the training of rehabilitation psychologists in substance abuse assessment and treatment. Throughout this section, alcohol and drug abuse are considered jointly. The focus of discussion, however, is on alcohol, the more frequently abused substance.

Bombardier and Turner provide a contemporary overview of alcohol and drug use in persons with disabilities.

The importance of alcohol screening is related to both the drug’s impact on bodily functions and the associated behavioral aberrations occasioned by its excessive use. Eckardt et al. reviewed the detrimental effect of alcohol on most organs, especially the liver, pancreas, and heart. Alcohol ingestion may potentiate the action of prescribed medications, most notably central nervous system (CNS) depressants. This potentiation is of particular importance for rehabilitation patients. For example, medications that control blood clot formation and reduce spasticity are frequently used with rehabilitation patients. Alcohol may decrease blood-clotting activity and act in an additive manner with muscle relaxants such as diazepam (Valium) and baclofen (Lioresal). In addition, altered consciousness may result in less vigilance in health-compromising situations. If alcohol is ingested in the form of beer, the large fluid volume could seriously compromise a bladder retraining program.

The cognitive and behavioral aberrations associated with drug intoxication often result in the onset of a disability. Rohe and DePompolo noted that vehicular crashes and falls, especially while the individual is under the influence of alcohol, account for a large proportion of admissions to rehabilitation units. Retrospective chart reviews of persons with CNS trauma often show alcohol present at the time of injury. Heinemann et al. found that 39% of their spinal cord–injured patients admitted to being intoxicated at the time of injury. Corrigan’s review of the literature found that as many as two thirds of head-injured patients have a history of substance abuse that preceded their injury. These studies revealed alcohol intoxication present in one third to one half of hospitalizations. Heinemann et al. found that drinking patterns before and after SCI are strongly related. In a study of long-term spinal cord–injured patients, problems resulting from substance abuse were reported by more than one half of the patients sampled. Rivara et al. found that 47% of general trauma patients had a positive blood alcohol level and
that 36% were intoxicated (136). The preceding data suggest that individuals admitted to rehabilitation units with traumatic CNS injuries are not a random sample of the drinking public. Assessment and intervention with this population represents an opportunity at reducing future medical, social, and personal costs. Corrigan argues that rehabilitation professionals, under scrutiny from third-party payers, cannot afford to have a significant proportion of their patients display poor long-term outcomes secondary to failure to address substance abuse (133).

The screening of chemical health for all rehabilitation patients, especially those who incur their disability as a result of trauma, must become a standard of care. For this to occur, the administrators of rehabilitation facilities will have to require that drug screening be standard policy and that rehabilitation professionals, especially physiatrists, have necessary screening skills. Although one recent survey of rehabilitation unit administrators indicated that chemical health screening has become common; as recently as 1996, Schmidt and Gavin found that only 4% of persons receiving initial traumatic brain injury rehabilitation were screened for substance abuse (124,137). Thus, the data on administrators’ perception of the adequacy of chemical health screening occurring in their facilities versus actual practice appear inconsistent. Unfortunately, staff training on this issue remains sorely lacking, with only 23% of rehabilitation unit administrators reporting that staff were provided education (124). Obtaining accessible substance abuse treatment for those with physical disabilities remains difficult (138). Perhaps the most overlooked chemical health issue is screening and intervention for nicotine dependence in those with disabilities. Basford, Rohe, and DePompolo point out that rates of smoking may be higher in those with disabilities when compared with the general population, and the health consequences in terms of lung function and wound healing may be particularly devastating to this population (124).

Screening for Chemical Dependency

Individual attitudes about alcohol use are diverse, strongly held, and determine the perception of another person’s use. Unless one first examines personally held attitudes and values about alcohol use, perceptions of another’s use may be biased. A training program developed by the Center for Substance Abuse Prevention stresses the importance of attitudes in prevention programs (139). The two most frequent problems encountered during screening are viewing alcohol as a moral problem and judging the deviance of the patient’s drinking through comparison with the interviewer’s personal pattern of use. Weinberg stated that the most important aspects of interviewing about alcohol use are (a) getting a detailed history, (b) demonstrating nonjudgmental acceptance, (c) asking direct, specific, and factual questions, (d) maintaining persistence, (e) never discussing alibis, and (f) titrating hostility (140).

The most promising approach to helping people change their alcohol use is “motivational interviewing,” which is based on Prochaska’s stages of change model (141,142). Two measures useful in assessing alcohol use are the CAGE Questionnaire and the Alcohol Use Disorders Identification Test (AUDIT).

The CAGE Questionnaire. The CAGE Questionnaire was originally developed on 130 randomly selected medical and surgical inpatients at North Carolina Memorial Hospital. The goal was to find the least number of questions that would reliably identify those suffering from alcoholism. The four CAGE questions are

1. Have you ever felt you ought to cut down on your drinking?
2. Have people annoyed you by criticizing your drinking?
3. Have you ever felt bad or guilty about your drinking?
4. Have you ever had a drink first thing in the morning to steady your nerves or to get rid of a hangover (i.e., an eye opener)?

Ewing summarized the scale’s development and data on four normative samples (143). One positive response to any of the questions raises the suspicion that alcohol dependence is present. Two positive responses identified 97% of his alcoholic sample correctly, and only 4% of his nonalcoholic sample incorrectly. Three or more responses are clearly symptomatic of alcohol dependence. The CAGE has sensitivity of from 60% to 95% and specificity ranging from 40% to 95% (144).

The Alcohol Use Disorders Identification Test. The AUDIT is a measure designed to aid in the process of accurate and reliable assessment of alcohol use. The AUDIT was developed by the World Health Organization to detect alcohol problems in primary medical care settings. The AUDIT consists of ten questions: three on alcohol consumption, four on alcohol-related problems, and three on alcohol dependence symptoms. The AUDIT requires 2 minutes to administer and 1 minute to score, with a recommended cutoff score of 8. Sensitivities are typically above 90%, with specificities in the 80% to 90% range (145).

In summary, the field of rehabilitation has yet to fulfill its responsibility in the screening and treatment of chemical abuse and dependency, especially nicotine dependence. Research data suggest that patients with traumatic CNS injuries have a high probability of chemical abuse. Physiatrists have a responsibility for learning the skills needed for systematic screening of their patients. This intervention, early in the rehabilitation process, is a crucial aspect of prevention of future medical complications. For example, Bombardier et al. have studied both those with recent SCI and Traumatic Brain Injury (TBI) in the inpatient setting. For both patient groups they found that the majority of at-risk drinkers were considering changes in their alcohol use. They conclude that intervening shortly after the onset of disability provides a window of opportunity to reduce post-injury alcohol abuse and related impairments (146,147).

Psychotherapeutic Interventions

Individual Psychotherapy

Psychotherapy is a generic term denoting psychological interventions that ameliorate emotional and behavioral difficulties. Psychotherapy can be defined as an interpersonal process whose goal is modification of problematic affect, behavior, or cognition (148). Although there are more than 130 varieties of psychotherapy, psychotherapy research has become
increasingly sophisticated in understanding key variables (149). Current research is focused on what specific types of therapy are best suited for what particular patient groups. For example, meta-analytic studies indicate that cognitive-behavioral therapy is as effective as anti-depressant medication for individuals with MDD. The effectiveness of psychotherapy seems to be related to a therapist's degree of training, and enthusiasm for, the theory and methods espoused by the particular therapy. Data suggest that psychotherapy does produce measurable change in patients. This change, however, can be negative as well as positive. The important principle of “above all, do no harm” is as important in psychotherapy as it is in medicine. Inadequately trained therapists are thus a source of concern in a field lacking firm boundaries.

The three basic assumptions underlying psychotherapy are as follows:

1. The person seeking services desires change.
2. The dysfunctional affect, behavior, or cognition is understood and amenable to change.
3. The process is a collaborative endeavor that assumes active client participation.

Psychotherapeutic intervention is thus contraindicated in patients on whom it must be forced or in those with significant communication or learning impediments. Additionally, if the difficulties are due to factors solely in the patient’s environment (e.g., long hospitalization, unpleasant medical interventions, prejudice, non-understanding staff), the focus of the psychotherapist’s intervention may shift from the patient to the environment (150).

The qualities of effective therapists have been studied and delineated. Therapeutic effectiveness initially depends on good assessment skills. Knowing when and how to intervene and, conversely, when to do nothing is fundamental to the process. Effective therapists are able to instill trust, confidence, and hope in their clients. Regardless of the type of therapy practiced, effective therapists have been shown to communicate the specific attitudinal qualities of genuineness, unconditional positive regard, and empathy. As opposed to mere friendship, the therapist provides an atmosphere of acceptance, respect, understanding, warmth, and help in conjunction with deliberate efforts to avoid criticizing, judging, or reacting emotionally with the patient. The creation of this atmosphere results in a framework unmatched by any other human relationship, one conducive to therapeutic change.

Most patients in rehabilitation units are faced with discovering and coping with permanent physical, cognitive, and social losses. This discovery is frequently accompanied by anger, anxiety, dysphoria, grief, and fear. Clinical experience suggests that the patient population can be divided into thirds according to the severity of their reactions. One third of the patients copes extremely well through use of previously established skills and the support of significant others. Another one third has greater difficulties, but with brief psychotherapeutic intervention is able to successfully manage the crisis. The final one third has significant difficulties in coping. These patients frequently have histories of difficulties in adjustment, such as chemical abuse, major mental disorder, and inability to tolerate structured living environments. This group is of paramount concern to the rehabilitation psychologist and consumes disproportionate amounts of professional time.

Because of the pressing practical problems faced by rehabilitation inpatients and the increasingly short periods of hospitalization, rehabilitation psychologists tend to use time-limited forms of therapy, also known as brief therapy. Brief therapy is a general term denoting therapies with a small number of sessions (i.e., six to ten) and limited, focused, and readily attainable goals. These goals often include amelioration of the most disabling symptoms, reestablishment of previous levels of functioning, and development of enhanced coping skills. The sessions are focused on concrete content and the “here and now.” Rehabilitation psychologists frequently apply techniques termed cognitive-behavioral. This concept is described by Turk et al. (151).

There are several common elements of cognitive-behavioral therapy. Interventions are active, time-limited, and fairly structured, with the underlying assumption that affect and behavior are largely determined by the way in which the individual construes the world. Therapy is designed to help the patient identify, reality-test, and correct maladaptive, distorted conceptualizations and dysfunctional beliefs. The patient is assisted in recognizing the connections among cognition, affect, and behavior, together with their joint consequences and is encouraged to become aware and monitor the role that negative thoughts and images play in the maintenance of maladaptive behavior.

Behavioral Management and Operant Conditioning Techniques

In contrast to many specialty areas of medicine, in rehabilitation medicine there is a strong and systematic interaction between the medical and the behavioral sciences. Because of their relevance to rehabilitation, the principles underlying behavior modification are discussed in detail. Included are the topics of token economies, behavioral contracting, and misconceptions about behavior modification. The following material is drawn from the writings of Martin and Pear (152), Reynolds (153), Kazdin (154), and Brockway and Fordyce (155).

Types of Reinforcers

There are three types of reinforcers. Primary or unconditioned reinforcers are present at birth. They include food, water, sexual stimulation, rest after activity and activity after rest, a band of temperatures, air, and cessation of aversive stimuli. Conditioned reinforcers are stimuli that have been repeatedly paired with primary reinforcers. They are idiosyncratic and are based on the learning history of the person. Generalized reinforcers are stimuli that have been paired with two or more conditioned reinforcers. The prime example of a generalized reinforcer is money; however, verbal responses such as “thank you,” “correct,” and “great” also are in this category. In addition to
the three types of reinforcers, there is an important principle, the Premack principle, which states that any high-frequency behavior can be used to reinforce a low-frequency behavior. For example, a high-frequency behavior such as watching television can be made contingent on performing a low-frequency behavior such as stretching exercises.

**Token Economies**
A token economy refers to a reinforcement system based on tokens. The tokens, frequently poker chips, function as generalized reinforcers and can be exchanged at agreed-on rates for back-up reinforcers, such as food, activities, and privileges. The behaviors to be changed (i.e., target behaviors) are specified along with the number of tokens earned for their performance. The stipulations of the economy are usually written in the form of a contract, a “reinforcement menu,” that indicates exchange rates and back-up reinforcers and is displayed in a prominent place. Token economies have been used extensively in special education and psychiatric settings. They can be useful with troublesome rehabilitation patients for such behaviors as arriving at therapy sessions late, lack of compliance with fluid schedules, and failure to perform activities of daily living. As with behavioral contracts, ethical considerations and the success of the program mandate full involvement of the patient in the initial design of the program.

**Behavioral Contracts**
Behavioral contracts, also known as contingency contracts, are written agreements between people who desire a change in behavior. The contract precisely indicates the relationship between behaviors and their consequences. The contract serves four important functions. First, it ensures that the rehabilitation team and the patient agree on goals and procedures. Second, because the goals are specified behaviorally, evidence is readily available regarding fulfillment of the contract. Third, the patient has a clear picture of what behaviors are expected if he or she is to remain in the rehabilitation program. Fourth, the signing of a document functions as a powerful indicator of commitment and helps ensure compliance with the agreement.

**Common Misconceptions about Behavior Modifications**
Behavior modification arouses concerns usually because of a misunderstanding of its underlying principles. Kazdin presented a succinct overview of common objections, two of which are iterated here (154). A frequent objection is that use of tangible reinforcers is the same as bribery. Bribery can be differentiated from reinforcement, because bribery is used to increase behavior that is considered illegal or immoral and usually involves delivery of the payoff before performance of the behavior, not, as in behavior modification, after. Bribery and reinforcement share the similarity of being ways of influencing behavior, but that is where the similarity ends.

A second objection is that behavior modification is “coercive.” Although behavior modification is inherently controlling and designed to alter behavior, multiple safeguards prevent its misapplication. These safeguards include involving the patient when contingencies are negotiated, constructing programs that rely on positive reinforcement rather than negative reinforcement or punishment, and making response requirements for reinforcement lenient at the beginning of the program (156). The use of behavioral modification in rehabilitation units requires careful training of staff. A limiting factor in many inpatient rehabilitation units is the lack of stability in team membership, especially where nursing personnel change frequently.

**Social Skills Training**

**Changes in Social Interaction after Disability**
Although only a limited number of the recently disabled may profit from psychotherapy, based on the author’s clinical experience many can benefit from social skills training. Research on the social psychology of disability is plentiful and underscores the social disadvantages encountered by the disabled, especially the recently disabled (157–169). Richardson summarized the literature and found consistently negative public attitudes toward the disabled (166). He noted that when first encountering a disabled person, the nondisabled experience heightened emotional arousal, anxiety, and feelings of ambivalence. These learned but somewhat involuntary reactions usually result in distorted social interactions.

Often the nondisabled focus solely on the disability and ignore personal characteristics normally used to evaluate people and establish relationships. The disabled also suffer from the societal norm to be kind to the disabled, which results in a lack of honest feedback and concomitant decreased accuracy in social perception by the disabled person (170). Consequently, the physically disabled often learn to discount praise and pay close attention to criticism. The factors mentioned earlier suggest that social interactions between the disabled and the nondisabled are complex, ambiguous, and unpredictable. Interventions that ameliorate difficulties with social interaction might help reduce emotional distress, speed the slow process of community reintegration, and reduce the risk of future medical problems (171).

**Social Skills: Types and Methods of Assessment**
Social skills is an inexact term used to describe a wide range of behavior thought necessary for effective social functioning. Dunn and Herman listed three types of social skills: general, general disability-related, and specific disability-related (Table 14-7) (172). Patients with the onset of disability before adolescence may require intensive remedial help with the development of general social skills. Patients with the onset of disability after adolescence enter the social arena with various competencies in general social skills. However, those with the onset after adolescence experience social situations for which they have no previous socialization experiences, hence the importance for training to handle these situations. General social skills can be assessed through a variety of means, including paper and pencil tests, behavioral assessment, and observational techniques.
### TABLE 14.7 General, General Disability, and Disability-Specific Social Skills

<table>
<thead>
<tr>
<th>General Social Skills</th>
<th>Disability-Specific Social Skills</th>
</tr>
</thead>
<tbody>
<tr>
<td>Listening</td>
<td>Dealing with staring</td>
</tr>
<tr>
<td>Positive and negative assertion</td>
<td>Refusing undesired help</td>
</tr>
<tr>
<td>Self-disclosure</td>
<td>Managing unwanted questions</td>
</tr>
<tr>
<td>Receiving compliments</td>
<td>Dealing with social advances</td>
</tr>
<tr>
<td>Confrontation</td>
<td>Acknowledgment of unstated attitudes (making the implicit explicit)</td>
</tr>
<tr>
<td>Touching</td>
<td>Asking for help</td>
</tr>
<tr>
<td>Maximizing physical attractiveness</td>
<td>Acknowledgment of the disability</td>
</tr>
<tr>
<td>Meeting new people</td>
<td></td>
</tr>
<tr>
<td>Use of humor</td>
<td>Refusing undesired help</td>
</tr>
<tr>
<td>Heterosocial skills</td>
<td>Managing unwelcome social advances</td>
</tr>
<tr>
<td>General Disability-Related Social Skills</td>
<td>Dealing with staring</td>
</tr>
<tr>
<td>Acknowledgment of the disability</td>
<td>Acknowledgment of unstated attitudes (making the implicit explicit)</td>
</tr>
<tr>
<td>Asking for help</td>
<td>Refusing undesired help</td>
</tr>
<tr>
<td>Acknowledgment of unstated attitudes (making the implicit explicit)</td>
<td></td>
</tr>
<tr>
<td>Refusing undesired help</td>
<td>Managing unwelcome social advances</td>
</tr>
<tr>
<td>Managing unwelcome social advances</td>
<td>Dealing with staring</td>
</tr>
<tr>
<td>Dealing with staring</td>
<td>Acknowledgment of unstated attitudes (making the implicit explicit)</td>
</tr>
<tr>
<td>Handling unwanted questions</td>
<td>Refusing undesired help</td>
</tr>
<tr>
<td>Disability-Specific Social Skills</td>
<td>Managing unwelcome social advances</td>
</tr>
<tr>
<td>Facilitating communication</td>
<td>Dealing with staring</td>
</tr>
<tr>
<td>Overcoming early deficits in socialization</td>
<td>Acknowledgment of the disability</td>
</tr>
<tr>
<td>Managing bowel and bladder problems</td>
<td>Refusing undesired help</td>
</tr>
<tr>
<td>Handling reactions to deformity and disfigurement</td>
<td>Managing unwelcome social advances</td>
</tr>
<tr>
<td>Disclosing nonvisible disabilities</td>
<td>Dealing with starring</td>
</tr>
<tr>
<td>Dealing with reactions to prosthesis</td>
<td>Acknowledgment of the disability</td>
</tr>
</tbody>
</table>

A social skill frequently identified as a problem is assertiveness; hence, it is used here to illustrate three assessment methods. A paper-and-pencil measure of assertiveness is the Gambrill Assertion Inventory (GAI) (173). The GAI presents 40 situations described by a short phrase—for example, “turn off a talkative friend.” The subject then rates his or her degree of discomfort for each situation on a five-point scale ranging from “none” to “very much.” Next, the subject rates his or her response probability for each situation on a five-point scale ranging from “always do it” to “never do it.” Normative data allow comparisons with the general population and with those having assertiveness difficulties. A measure of assertiveness more relevant to disability is the SCI Assertion Questionnaire (174). The format of this questionnaire is similar to that of the GAI, but social situations that are potential problems for wheelchair users are described.

Behavioral measures help clarify the frequently found discrepancy between what people say they do and what they actually do. Behavioral measures offer direct and quantifiable data on both verbal and nonverbal aspects of social interactions. Such measures might include checklists or rating scales that permit counting responses, measuring length of time spent interacting, and so on. For example, studies using behavioral measures have shown that disabled people receive offers of help from strangers less frequently than do the non-disabled. However, if help is offered to the disabled, it tends to be overly solicitous (175,176). Hastorf et al. (170) found that strangers were more willing to work on a cooperative project if the disabled partner assertively acknowledged the handicap at the beginning of the interaction. Finally, behavioral measures have been used during evaluation of the efficacy of assertion training with disabled people (177–180).

Observational techniques can be used by the person, significant others, or staff members. Generally, this type of assessment is less objective than that provided by behavioral measures. Nonetheless, reduced precision is counterbalanced by the opportunity to observe qualitative aspects of the social skill in a natural setting. Several research projects (e.g., Longitudinal Functional Assessment System and Rehabilitation Indicators Project) use observational techniques in the form of diaries, self-reports, and environmental surveys. Thorough descriptions of methods of assessing social skills can be found in a variety of publications (177,181,182). The reader is encouraged to consult relevant social skills training manuals for intervention techniques (183,184). Social skills training programs for wheelchair users are described by Dunn and Herman (172). Brotherton et al. (185) describe a program for the traumatic brain injured.

### INDIRECT SERVICES

The rehabilitation psychologist’s overall aim is to enhance the quality of rehabilitation outcomes for patients. Indirect services such as maximizing team interaction skills, staff development, administration, and research provide avenues for enhancing patient outcomes that are as important as those of direct patient services.

The rehabilitation team is a unique structure in the delivery of health-care resources (186). Nowhere else are so many professionals with diverse backgrounds of training expected to communicate in a clear, timely, and comprehensive manner. This communication may become tenuous because of different professional terminologies, overlap in roles, and the pressures of productivity in a competitive health-care environment (187). The psychologist can enhance patient outcomes by facilitating cohesion of the rehabilitation team (188). This task can be accomplished through a variety of methods, including chairing committees to improve interdisciplinary cooperation and leading staff meetings to clarify overlap in professional roles. Butt and Caplan provide a comprehensive review of the rehabilitation psychologist’s role in the rehabilitation team (189). In the author’s experience, rehabilitation psychologists play a variety of roles in relationship to the rehabilitation team. The rehabilitation psychologist’s knowledge of distinguishing normal from abnormal behavior in adjustment to trauma and disability is frequently utilized for staff education. Although some
in-service topics focus on patient variables, such as practical management suggestions and brain-behavior relationships, other topics include personal concerns of the staff such as job stress and communication skills. The psychologist's strengths in interpersonal communication frequently prompt their selection for administrative positions.

Rehabilitation psychologists trained at the doctoral level typically have extensive expertise in research design and statistical methods. As such, they frequently consult with other team members interested in conducting research. They may coordinate research or direct research committees. The research expertise of psychologists is reflected in their presence on editorial boards of numerous rehabilitation-related publications. They are also found in local, state, and national organizations whose function is to promote quality rehabilitation and social justice for the physically disabled.

**PSYCHOLOGICAL ADJUSTMENT TO DISABILITY**

This section is divided into two parts. The first part reviews theories of adjustment to disability. The second part describes three models of adjustment to disability. The stage model emphasizes internal cognitions; the behavioral model emphasizes external events; and the coping skills model emphasizes both internal cognition and external events in the adjustment to disability.

**Overview of Theories of Adjustment to Disability**

Theories of adjustment to disability can be grouped along an internal to external continuum (190). On one end are theories that emphasize internal cognitive events, termed *mentalistic theories*, and on the other end are theories that emphasize events external to the individual, termed *social theories* or *behavioral theories*. The middle of the continuum contains integrative theories that attempt to meld the internal with the external determinants.

Before formal theorizing about adjustment to disability, most people thought that the primary source of suffering connected with disability was the disability itself. Hence, removal or amelioration of the disability would presumably reduce distress. However, experience demonstrated that after removal of a disability, some people remained incapacitated. The search for understanding the adjustment process shifted to the then contemporary principles of dynamic psychology and focused on internal events such as motivation. Patients' difficulties in adjusting to disability were conceptualized in psychodynamic terms, and their incapacitation was transformed into a mental health problem.

As time progressed, dynamic psychology models, especially the classic psychoanalytic model with its emphasis on psychopathology, provided insufficient explanatory power. Professionals came to recognize that physical and social barriers, barriers external to the patient, produce the major source of adjustment problems. Emphasis on sociologic concepts such as “sick role” (191) and “illness behavior” (192) ensued. These sociologic theories added to the understanding of adjustment to disability on a societal level (193). When an individual's adjustment to disability behavior is the focus, however, learning theory's emphasis on the sensitivity of behavior to its consequences provides significant explanatory power. The behavioral model of adjustment to disability is described more fully in a later section.

Theories that attempt to simultaneously take into account the internal events of the person and the external demands of the environment are called *integrative field theories* (190,194,195). Integrative field theories grew out of Lewin's concept of life space (196). These theories state that behavior is a joint function of the person and his environment \( B = f(P, E) \). Meyerson, Trieschmann, and Wright applied Lewin's basic formulation to problems encountered by the physically disabled (164,197,198). For example, Trieschmann expanded Lewin's model and described the “educational model of rehabilitation” in which behavior is the function of personal, organic, and environmental variables, designated by the formula \( B = f(P \times O \times E) \). Her acknowledgment of organic variables highlights the concept that behavior is fundamentally dependent on and limited by the physical capacities of the person.

**Models of Adjustment to Disability**

**Stage Model**

Stage theory states that people undergoing a life crisis follow a predictable, orderly path of emotional response. Shontz is the major contributor to the application of stage theory to adjustment to disability (199). Stage theory appears both explicitly and implicitly in a wide variety of rehabilitation-related literature, including that on cancer, hemodialysis (200,201), SCI (202–204), and amputation (205). Additional writers who make implicit or explicit reference to a stage model of adjustment to disability are Dembo et al., Siller, and Guenther (159,167,206). Unfortunately, these studies are merely descriptive and are based on interview data or anecdotal reports.

Most stage theories propose a series of three to five steps beginning with shock and ending with some form of adaptation. Three commonly held assumptions appear to underlie stage theory formulations applied to the disabled. First, people respond to the onset of disability in specific and predictable ways. Second, they go through a series of stages over time. Finally, they eventually accept or resolve their emotional crises. The following discussion draws on the work of Silver and Wortman (207–209) and Trieschmann (197). Much of the stage theory literature is derived from the study of bereavement. Although there are important differences between the death of a significant other and loss (death) of a body part or function, it is reasonable to assume that the processes of coping with death and coping with disability are psychologically similar if not equivalent.

**Are There Universal Responses to Disability Onset?**

Both Guenther (206) and Shontz (199) indicated that once the crisis of disability is realized, virtually all people experience shock. Unfortunately, most studies report retrospective accounts of initial feelings and behavior. In one such study,
Parkes interviewed widows and amputees after their losses (205). Initial feelings of shock and numbness were reported by only 50% of the sample. Tyhurst observed disaster victims and described three types of reactions (210). One group reacted with classic signs of shock, another group appeared cool and collected during the acute situation, and a third group responded with reactions of paralyzing anxiety and hysterical crying. Shock was far from a universal reaction. Silver and Wortman’s 1980 literature review concluded that there is little evidence supporting the belief that people react in specific and predictable ways to undesirable life events (207). Although some patterns are evident, individual variation is inevitably present. Their updated (2001) literature review found no research that would compel them to revise their earlier conclusion. The author’s clinical experience supports this conclusion. The initial reaction to the onset of disability is diverse. Moreover, the more accurate the first thought after the onset of disability (termed by the author the sign-on cognition) and the more honest the self-acknowledgment, the better the initial and longer-term adjustment appears to be. Further research on the concept of sign-on cognition may prove useful in understanding the process of adjustment to disability.

Do Emotional Responses Follow a Pattern After Injury?
The belief that people follow a predictable pattern of emotional response after the onset of a disability is widely held. References to stage models of emotional response occur in the professional literature of nurses (211,212), social workers (213), clergy (214), health-care professionals (215), and psychologists (167,199,206).

Silver and Wortman were unable to discover any studies specifically testing stage theory by measurement of affective states over time (207,209). Four related studies, all conducted on patients with SCI, failed to support state theory. Dunn, cited by Trieschmann studied seven psychological variables during three time periods of inpatient rehabilitation (197). He found no pattern of change in mood over time; variability was the norm. McDaniel and Sexton assessed psychological status over four points in time from ratings by rehabilitation team members (213). Ratings of negative mood states remained relatively constant over the length of the study and were independent of staff ratings of the patient’s degree of acceptance of loss. Dinardo, in a cross-sectional study, found that the degree of depression experienced by his subjects was independent of the time that had elapsed since their injury (216). Finally, Lawson, in a longitudinal study, used a variety of methods to assess the presence of depression (217). He found no period of at least a week when any of his patients scored consistently in the depressive range on any of the measures. His results suggest that patients with SCI do not experience a stage of depression during initial rehabilitation.

Although there is considerable popular and professional literature attesting to the veracity of stages of adjustment to disability, the empirical data do not support such a contention. Silver and Wortman summarized the available data by stating, “Perhaps the most striking feature of available research, considered as a whole, is the variability in the nature and sequence of people’s emotional reactions and coping mechanisms as they attempt to resolve their crises” (207). Their follow-up review of the literature highlighted that there are deep-seated assumptions about how people “should” react to loss. These assumptions include that (a) individuals suffering a loss are supposed to go through a period of intense distress; (b) failure to experience intense distress is suggestive of a problem; (c) successful adjustment requires that the individual “work through” his or her feelings; (d) continued attachment to the deceased (in the case of disability, attachment to previous levels of body function) is viewed as pathologic; and (e) within a year or two people will recover from their loss and return to earlier levels of functioning (209). They note that individuals who do not comply with these assumptions may incur negative reactions from their peers and subsequently shift their behavior to be in accord with others’ expectations.

Is a Final Stage of Resolution Reached?
Do people who have suffered a major undesirable life event eventually reach a final stage of resolution or acceptance of their disability? The findings across studies suggest that a large minority of people continue to suffer years after a traumatic life event. The unquestioned expectation of resolution or acceptance appears unwarranted for such traumatic life events as severe burns, SCI, cancer, death of a spouse, and rape (207,218). For example, Shadish et al. studied a cross-sectional sample of patients with SCI (219). They found that those who had been disabled for as long as 38 years continued to think about and miss physically impossible activities.

Wortman et al. discuss a theoretical framework that suggests that the impact of a major undesirable life event is determined by whether the event can be incorporated into an individual’s view of the world (220). The term worldview denotes the system of beliefs, assumptions, or expectations related to oneself, others, and the world that provides a sense of coherence and meaning (221). Losses that are sudden, uncontrollable, and random may readily shatter people’s assumptions about the world. Thus, the extent to which an individual’s worldview is violated will determine the intensity of the disequilibrium and distress that they experience. The degree to which they can reconcile the event with their preexisting worldview or create a new worldview that adequately accounts for the event will determine their long-term adjustment. Hence, disability of sudden onset would have a greater impact on worldview than disability caused by chronic illness.

Given the lack of support for three common assumptions underlying stage theory, one must consider alternative explanations of why such beliefs permeate clinical folklore and descriptive writing in the area. As suggested earlier, stage theory may represent a codification of strongly but implicitly held culturally determined expectations about reaction to loss. Two additional reasons come to mind. First, professionals working with the recently disabled often encounter unpredictable and emotionally charged situations. One way to help neutralize
fears occasioned by such situations is to label the patient as being in a particular stage. This may help transform potentially threatening and seemingly unpredictable behavior into meaningful and predictable categories. A negative outcome of such conceptualizing may be the documented tendency for rehabilitation personnel to overdiagnose psychopathologic conditions in patients (222). This interpretation of behavior may also result in the staff inappropriately distancing themselves from the patient by negating the necessity for careful listening. A second reason may be the enticing belief that all patients eventually resolve the negative effect occasioned by their disability and achieve a final stage of adjustment or resolution. Such a belief has intrinsic appeal to health-care professionals who strive to maximize functional abilities and enhance quality of life.

**Behavioral Model**

The behavioral model of disability adjustment emphasizes the importance of external factors in determining a person's adjustment. In this model there is reduced interest in the patient's cognitions and a primary focus on observable behaviors. The most frequently cited proponent of this model is Fordyce, and much of what follows is culled from his writings (155, 223). Additional applications of the behavioral model to rehabilitation problems can be found in the works of Ince (224) and Berni and Fordyce (225). In the behavioral model of adjustment to disability, the newly disabled face four tasks. The patient must remain in the rehabilitation environment, eliminate disability-incongruent behaviors, acquire disability-congruent behaviors, and maintain the output of disability-congruent behaviors.

The onset of physical disability and entry into the rehabilitation environment represent punishment. In learning theory, punishment is defined as the loss of access to positive reinforcers or the response-contingent onset of aversive stimuli. Thus, the newly disabled find themselves initially operating under a pattern of punishment. Two types of behavior follow the onset of aversive stimuli. The first is escape or avoidance, and the second is aggression. Escape or avoidance behavior is frequently seen in the rehabilitation setting in the form of daydreaming, verbal disclaimers of disability, unauthorized forays off the medical unit, and refusal to participate in scheduled treatments. Aggressive behaviors may consist of either rebellious behavior or verbal and sometimes physical attack. If avoidant or aggressive behaviors are not understood and dealt with therapeutically, rehabilitation may end prematurely.

The intervention strategy for these problems involves the discovery and, if possible, reduction of aversive aspects of the rehabilitation environment. This is accomplished by reinforcement of approximations to active participation in the rehabilitation program. Selecting and systematically graphing a mutually agreed indicator of rehabilitation progress can help the patient focus on tangible improvements. Patient reactions of hostility are common and should be tolerated within limits. These reactions should never be dealt with through counter-hostility, which only increases the probability that the environment, including the treatment staff, will become conditioned aversive stimuli. Systematically ignoring unwanted behavior and establishing therapeutic rapport enhance the probability that the patient will remain in the rehabilitation environment.

The reduction of disability-inappropriate behaviors and the acquisition of disability-congruent behaviors are synonymous with the concept of "adjustment to disability." Disability-inappropriate behaviors are decreased by withdrawal of reinforcers after their occurrence, a process known as extinction. Paradoxically, the laws of behavior demonstrate that withdrawal of reinforcers initially results in a temporary increase in the rate of behavior. This is true for both verbal and performance behaviors.

The patient's verbal behavior is likely to change more slowly than performance behavior. Statements indicating a belief in the eventual return of physical function may require years to extinguish. The staff should neither reinforce nor punish unrealistic verbalizations. Rather, a verbal response suggesting the need to maintain hope tempered with a focus on the present is least likely to offend the patient. These statements of patients are more frequent at the onset of rehabilitation and may reflect the beginning of extinction. Detailed explanations of anticipated recovery of functional abilities help decrease unrealistic patient or family verbalizations and keep everyone focused on achievable functional goals. This is especially important for family members, who may erroneously believe that the proper way to help the disabled family member cope is through agreeing with unrealistic fantasies about eventual recovery of function.

Difficulties in the acquisition of disability-congruent behaviors are usually considered to be problems in motivation. Learning theory rejects this formulation because it relies on an inference about the internal state of the person. Usually, this label is applied to people who have failed to reach expected levels of performance set by the rehabilitation staff. In learning theory, the problem is that of adjusting contingencies to increase the rate of desired behavior or reduce the rate of behaviors competing with the desired behavior. Unfortunately, most disability-congruent behaviors are initially of low frequency, strength, and value. The steps in changing this situation include establishing reinforcing relationships with the treatment staff, enhancing long-term reinforcers for disability-congruent behaviors, and introducing contingency management interventions that promote the acquisition of disability-congruent behaviors.

Maintaining the output of disability-appropriate behaviors is the final and most important step in adjustment to disability. Rehabilitation is unsuccessful if the behaviors learned in the rehabilitation unit cannot be transferred to the patient's home environment (226). Although the patient may demonstrate the ability to perform a task, the probability of its occurrence depends on contingencies operating in the home environment. Disability-congruent behaviors, such as propelling a wheelchair, maintaining a fluid schedule, and using gait aids are unlikely to be reinforcing in themselves.
Two strategies for improving generalization are bringing disability-congruent behaviors under the control of reinforcers occurring naturally in the environment and reprogramming the patient’s home environment to deliver appropriate reinforcement contingently. The first strategy is promoted through interventions designed to reengage the patient in meaningful vocational and avocational activities after dismissal. Therefore, vocational counseling and therapeutic recreation are important as part of inpatient rehabilitation. Gradual and systematic rehearsal of newly learned skills in the home environment during weekend visits is an additional method encouraging generalization. The second strategy is promoted through such interventions as home modifications, assigning a family member to monitor and reinforce home therapy programs, and contracting with the patient for continued compliance. Unfortunately, powerful contingencies may be operating to prevent generalization. For example, the patient may receive reinforcers in the form of increased attention or financial rewards from litigation, a condition also known as secondary gain. Inability to control sources of secondary gain may prevent generalization of disability-congruent behaviors to the home environment. Family interventions are critical to prevent these problems.

Coping Skills Model
The coping skills model (227), which emphasizes both cognitive and behavioral factors, is based on the crisis theory originally formulated by Lindemann (228). Crisis theory asserts that people require a sense of social and psychological equilibrium. After a traumatic event, a state of crisis and disorganization occurs. At the time of the crisis, a person’s characteristic patterns of behavior are ineffectual in establishing equilibrium. This state of disequilibrium is always temporary, and a new balance is achieved within days to weeks. Veninga provides a practical guide for negotiating the dilemma of a crisis (229). Snyder’s edited text on coping research is a must read on how any of a variety of possible future stressors will be minimized.

Denying or Minimizing the Seriousness of a Crisis
This coping skill may be directed at the illness or at its significance and helps to reduce negative emotions to manageable levels. This reduction enhances the mental clarity needed for effective action in emergency situations. The likelihood of implementing a greater range of coping responses is also increased.

Seeking Relevant Information
Often, misunderstanding of medical diagnoses and procedures causes emotional distress. Understanding often reduces anxiety and provides a sense of control. Gathering information gives the patient and family a concrete task and the accompanying feeling of purposefulness. One longitudinal study of people with chronic illness showed that information-seeking has salubrious effects on adjustment (232).

Requesting Reassurance and Emotional Support
The literature shows that perceived social support, adjustment during a crisis, and improved health outcomes are interrelated (233–235). Component parts of social support include perceiving that one is cared for, being encouraged to openly express beliefs and feelings, and being provided material aid. Social support may enhance coping by reducing counterproductive emotional states, building self-esteem, and increasing receptivity to new information. Cobb suggested that social support enhances health outcome either directly through neuroendocrine pathways or indirectly through increased patient compliance (236). He cited evidence showing that patients who receive social support are more likely to stay in treatment and follow their physicians’ recommendations. Turner found a reliable association between social support and psychological well-being, especially during stressful circumstances (237).

Learning Specific Illness-related Procedures
Learning specific illness-related procedures is a skill that reaffirms personal competence and enhances self-esteem, which is often undermined by physical disability. Bulman and Wortman asked social workers and nurses on a rehabilitation unit to define good and poor coping in patients with SCI (238). Both groups agreed that good coping included the willingness to learn physical skills that would minimize disability. Conversely, the definition of poor coping included an unwillingness to improve the condition or attend physical therapy.

Setting Concrete Limited Goals
Limited goal setting breaks a large task into small and more readily mastered components. As each component is mastered, self-reinforcement accrues and sets the stage for further learning. Limited goal-setting decreases feelings of being overwhelmed and enhances the opportunity to achieve something considered meaningful.

Rehearsing Alternative Outcomes
Activities such as mental rehearsal, anticipation, discussions with significant others, and incorporation of medical information are involved in this skill. Here the patient considers possible outcomes and determines the most fruitful manner of handling each. Recalling previous periods of stress and how these were successfully managed is an example of this coping skill. The patient engages in behaviors that alleviate feelings of anxiety, tension, fear, and uncertainty. A cognitive road map is delineated to provide guidance on how any of a variety of possible future stressors will be minimized.

Finding a General Purpose or Pattern of Meaning in the Course of Events
Physical disability is a crisis that can destroy a person’s belief that the world is a predictable, meaningful, and understandable place. There appears to be a compelling psychological need to
believe that the world is just (239) and to make sense out of a crisis experience. The previously discussed concept of worldview, with its focus on coherence and meaning, is relevant in this context. Some theorists claim that the search for meaning is a basic human motivation (240). Bulman and Wortman studied 29 subjects with SCI and concluded that the “ability to perceive an orderly relationship between one’s behaviors and one’s outcomes is important for effective coping” (238). Krause, in a 15-year prospective study of persons with SCI, found that survival was directly related to higher activity levels and being employed (241).

SUMMARY

This chapter reviewed the history and current status of rehabilitation psychology, followed by an overview of services offered by the rehabilitation psychologist and theories of adjustment to disability. Although the rehabilitation psychologist provides a wide variety of direct and indirect services, certain skills are particularly relevant to rehabilitation, including psychological assessment, behavior modification, and research. Rehabilitation environments represent settings in which people under physical and emotional distress are asked to learn. Many of these people not only are emotionally upset but may have brain injuries that further impair learning efficiency. Standardized measurement of personality, mood, intellectual ability, academic achievement, neuropsychological integrity, and chemical health provide a reliable base on which to set rehabilitation goals.

Rehabilitation is concerned with the functional performance of a person. Rehabilitation team members provide diverse interventions to ensure that the person can physically perform specific activities. Whether this person will actually do so is determined by contingencies in the rehabilitation unit and home environment. The rehabilitation psychologist’s behavioral modification skills permit the careful assessment and harnessing of these contingencies in the service of the patient.

Progress in any scientific field depends on quality research. Such research is of particular concern for rehabilitation because outcomes are determined by a complex set of physical and social variables. Doctoral-level psychologists are typically the only rehabilitation team members with training in research. Traditionally, this training stresses asking practical research questions relevant to clinical problems.

Theories of adjustment to disability can be grouped along a continuum stressing internal cognitive events on the one end and external social and behavioral events on the other. Stage theory is a widely held but largely unsubstantiated model that stresses internal events. Alternative models worth considering include the behavioral model and the coping skills model. Contemporary research on the concepts of coping and hope provide a compelling theoretical basis for understanding the process of adjustment to disability and chronic illness.

REFERENCES

2. Eisenberg MG, Jansen MA. Rehabilitation psychology: state of the art. 


41. Lomay VT, Hinkebein JH. Cultural considerations when providing rehabilitation services to American Indians. Rehabil Psychol. 2006;51(1):36–42.


INTRODUCTION

Human communication involves the exchange of information and ideas through various processes commonly referred to as speech, language, cognition, and hearing. It is thought that the study of speech communication is one of the oldest academic disciplines with roots dating back to the time of Aristotle. Despite technological advancements, human communication has retained its value because of its practical nature in the transfer of meaningful information amongst individuals. The components of human communication are dynamic and multidimensional, which are continuously influenced by physiological, psychological, and environmental factors. Human communication occurs through a series of autonomic processes mediated by the central neurological system and monitored via a sensory feedback loop. Although complex, the components of communication can be simplified into a four-step process:

- Encoding: the speaker creates the message in his or her mind
- Transmittal: the speaker sends the message
- Reception: the listener receives the message
- Decoding: the listener breaks down the message in his or her mind

This chapter serves to improve the understanding of human communication by describing the processes of speech, language, cognition, and hearing. Additionally, oropharyngeal swallowing is reviewed due to its anatomical and physiological similarities to speech production. Descriptions of commonly recognized acquired disorders in the adult population are reviewed to facilitate understanding of the rehabilitative assessments and management.

COMMUNICATION

Human communication is often described as a well-coordinated, timed, multidimensional group of processes used to share thoughts, ideas, and emotions. The processes include speech, language, cognition, and hearing. Speech is the system that produces individual sounds, which, when placed together, develop meaningful messages. It is the verbal means of communication often thought of as the motor vehicle for message delivery. Language may be defined as an arbitrary system of signs or symbols used according to prescribed rules to convey meaning within a linguistic community (1). Language forms the set of rules shared by communicators consisting of morphology, syntax, semantics, and pragmatics. Cognition refers to the mental processes by which one recognizes, manipulates, and exchanges information to better “understand and make sense of the world” (http://www, December, 2008). Cognition gives rise to the thoughts and ideas for which one uses speech and language. Hearing is the ability to detect and assign meaning to sounds. In the case of speech, hearing serves as the system that closes the communication loop.

COMMUNICATION PATHOLOGY

A pathologic condition that affects any one of the organs involved in the process of speech or language influences the total final product. Frequently, the pathologic condition often is limited to a single speech or language component, and the dysfunction can be detected in only one specific parameter such as an isolated voice or articulation impairment. More commonly, however, the pathologic condition of a single organ influences other elements of the communication process in ways that are predictable when one considers the integrated nature of speech and language. For example, severe obstructive pulmonary disease does not just impair the respiratory support for speech but results in alterations in vocal pitch, vocal intensity, and phrasing or prosody, as the speaker compensates for an impaired ability to sustain airflow. Certain disease states, involving organs that are not directly involved in speech or language, can affect the final vocal product in a secondary manner. For example, some endocrine disorders, such as hypothyroidism, can influence voice quality as an isolated component of speech. It can also lead to language confusion and impaired memory. Because of the complexity and interactive nature of the speech and language processes, whenever one evaluates or treats a patient with a communication disorder, some considerations must be given to overall human physiology, as well as to the dynamics of speech and language systems. Many professionals including audiologists, physiatrists, and otolaryngologists contribute to the assessment and treatment of communication. Speech-language pathologists have the specialty training in the communication impairments and are
credentialed to assess, treat, and manage the communication disorders with the ultimate goal of developing and maintaining a functional communicative status for individuals.

**SPEECH PRODUCTION**

Speech production requires multiple physiological systems working together to create an intelligible message. The five systems comprising speech production include respiration, phonation, resonation, articulation, and hearing. Although each system works concurrently with each other to achieve the end result of speech, it is often easier to understand the complexity of the systems by reviewing them as separate entities.

**Respiration**

Two forms of respiration are recognized: chemical and mechanical. Chemical respiration is concerned with the exchange of oxygen and carbon dioxide to and from the blood, whereas mechanical respiration (or ventilation) is concerned with the tidal movement of air in and out of the lungs. For speech production, mechanical ventilation will be the primary focus.

To produce sound, an energy source is needed to initiate the process. For speech, this energy source is the respiratory system, causing air pressure variations that result in airflow out of the respiratory tract and into the vocal tract. During inhalation, the diaphragm contracts and lowers allowing the thoracic cavity to expand and the lungs to fill with air. This causes a low pressure gradient, allowing the air to rush into the lungs. The pressure then equalizes as air is sustained in the lungs. Exhalation is the more passive mechanism of the inhalation/exhalation process. Exhalation causes relaxation of the diaphragm causing it to return to its normal position and shape. This in turn causes increased pressure gradient compared to outside the lungs, which ultimately results in air flow out of the lung cavity. The entire process of breathing (inspiration and expiration) is known as one respiratory cycle. At rest, this takes place 12 to 15 times a minute. Speech production occurs upon exhalation. During speech production, the time for inhalation is often reduced while the exhalation time is increased.

**Phonation**

The phonatory process, or voicing, occurs when air is expelled from the lungs through the glottis (space between the vocal folds). As air passes through this area, a lower pressure gradient occurs within the larynx. When this drop becomes sufficiently large, the vocal folds begin oscillating. This oscillation causes vibration, which is often referred to as the engine of the voicing mechanism. Vocal fold motion occurs in a wavelike fashion with lateral undulations as well as superior and inferior movements. The oscillation of the vocal folds serves to modulate the pressure and flow of the air through the larynx. This airflow is modulated and is the main source of voiced sounds. A normal voice is best described as the product of a controlled exhalation of air, steady maintenance of subglottic air pressures, and delicately balanced vocal folds movement. Any perturbation to this complex process can result in a perceived vocal impairment.

**Resonation**

The raw vocal tone is modified and amplified by resonance within the pharyngeal, oral, and nasal cavities, which are referred to collectively as the vocal tract. The innumerable configurations of the shape of the vocal tract provide the human voice with a tremendous range of variation in perceived quality.

Following transit through the larynx, the air is modulated by (a) changing the tension of the pharyngeal walls; (b) raising or depressing the larynx; (c) modifying the position of the jaw, tongue, and lips; and (d) occluding or lowering the soft palate. Valving of the velopharyngeal region requires the closure of the nasal cavity from the oral cavity for production of all sounds other than the /m/, /n/, and /ng/ sounds. This valving is achieved in a three-dimensional pattern of closure typically described as retraction and elevation of the soft palate/velum, anterior movement of the posterior pharyngeal wall, and inward movement of the lateral pharyngeal walls. Various patterns of velopharyngeal closure were initially described by Skolnick in 1973 and adopted by others to better understand the closure mechanism (2,3).

**Articulation**

The final production of sound occurs as air is expelled through and manipulated within the oral cavity by the movements of the lips, teeth, and tongue. Coordinated actions of the tongue, lips, jaw, and soft palate regulate the air stream and produce the meaningful sounds of speech called phonemes. These structures, often referred to as articulators, relax, compress, or momentarily stop the air stream in varied ways to produce specific vowels or consonants. The addition or elimination of a voicing component from the larynx, during sound production, dictates whether a consonant is voiced or voiceless.

**Neural Control**

Motor control for respiration, phonation, resonation, and articulation requires complex neural networking, regulation, and monitoring of muscle activity of both the central nervous system and the peripheral nervous system. Many of the speech systems and musculature used for speech production are also used for oropharyngeal swallowing and mastication. It is natural that a large amount of overlap exists in the neural networking, although affecting different behaviors. The central nervous system and peripheral nervous system are the primary controls for the neural networking, though the extrapyramidal nervous system also plays a role. Although studied since the 17th century, many questions remain in determining the brain regions involved in speech production. Neuroimaging advances continue to attempt to solve this dilemma; however, inconsistencies in normal subject’s activation patterns continue to be found (4). Speech involves simultaneous activity in various regions of the brain that are largely spread over the cortical and subcortical areas. Additionally, clinical data of patients with cerebellar lesions or Parkinson’s disease suggest that the cerebellum and the brainstem are engaged in the motor control of speech (5–7).
MOTOR SPEECH DISORDERS

Dysarthria

The dysarthrias are a group of motor speech disorders characterized by slow, weak, imprecise, or uncoordinated movements of speech musculature. Rather than a single neurological disorder, the dysarthrias vary along a number of different dimensions. As stated above, the neuroanatomical site of a lesion causing dysarthria can be one or a combination of the cerebrum, cerebellum, brainstem, and cranial nerves. The prevalence and incidence of dysarthria are not precisely known. Dysarthria can be a symptom of a neurological disease process with a constellation of other symptoms, or it can stand alone with a disease. Approximately one third of individuals with traumatic brain injury (TBI) may be dysarthric with nearly double that prevalence during the acute phase of recovery (8,9). Dysarthria is frequently seen in 50% to 90% of parkinsonian patients with the increased prevalence as the disease progresses (10). Additionally, dysarthria is often a preliminary sign of amyotrophic lateral sclerosis or can become present as the disease progresses. In larger tertiary care medical centers, dysarthria was the primary communication impairment with acquired neurological disease seen for speech-language pathology evaluations over a 4-year period (11).

The dysarthrias can be classified by time of onset, site of lesion, and etiology; however, the most widely used classification was first described by Darley et al. (12–14). This is often referred to as the Mayo Clinic Classification System with each type representing a perceived and distinguishable grouping of speech characteristics with a presumed underlying pathophysiology or locus of lesion.

As a first step, differential diagnosis involves distinguishing the dysarthrias from other neurogenic communication disorders. The dysarthrias are distinct from aphasia in that language function (i.e., word retrieval, comprehension of both verbal and written language) is preserved in dysarthria but impaired in aphasia. Although both apraxia and dysarthria are considered motor speech disorders, they can be distinguished on the basis of several clinical features. In apraxia, automatic (i.e., nonspeech) movements are intact, whereas in most dysarthrias, they are not. Highly consistent articulatory errors are characteristic of dysarthria, whereas inconsistent errors are a hallmark of apraxia. Finally, in most dysarthrias, all speech systems, including respiration, phonation, resonance, and articulation, are involved; whereas in apraxia, respiratory or phonatory involvement is rare. It should be recognized that patients often can have elements of both dysarthria and apraxia, particularly those with bilateral brain damage.

Differential diagnosis among the dysarthrias seems to have received more systematic attention than any other aspect of the disorder. Table 15-1 summarizes information related to the various dysarthrias (15). In studies conducted at the Mayo Clinic, the perceptual features of the speech of seven groups of dysarthric patients were examined (12,13). These groups consisted of patients who were unequivocally diagnosed as having one of the following conditions: pseudobulbar palsy, bulbar palsy, amyotrophic lateral sclerosis, cerebellar lesions, parkinsonism, slow ataxia, myoclonus, and tremors.

### TABLE 15.1 Summary of the Etiologies, Neuropathologies, and Neuromuscular Deficits Characteristic of the Common Dysarthrias

<table>
<thead>
<tr>
<th>Type</th>
<th>Example</th>
<th>Location of Neuropathology</th>
<th>Neuromuscular Deficit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flaccid</td>
<td>Bulbar palsy</td>
<td>Lower motor neuron</td>
<td>Muscular weakness; hypotonia</td>
</tr>
<tr>
<td>Spastic</td>
<td>Pseudobulbar palsy</td>
<td>Bilateral upper motor neuron</td>
<td>Reduced range, force, speed; hypertonia</td>
</tr>
<tr>
<td>Ataxic</td>
<td>Cerebellar ataxia</td>
<td>Cerebellum (or tracts)</td>
<td>Hypotonia; reduced speed; inaccurate range, timing, direction</td>
</tr>
<tr>
<td>Hypokinetic</td>
<td>Parkinsonism</td>
<td>Extrapyramidal system</td>
<td>Markedly reduced range; variable speed of repetitive movements; movement arresting rigidity</td>
</tr>
<tr>
<td>Quick</td>
<td>Chorea, Myoclonus, Gilles de la Tourette’s syndrome</td>
<td>Extrapyramidal system</td>
<td>Quick, unsustained, random, involuntary movements</td>
</tr>
<tr>
<td>Slow</td>
<td>Athetosis, Dyskinesias, Dystonia</td>
<td>Extrapyramidal system</td>
<td>Sustained, distorted movements and postures; slowness; variable Hypotonus</td>
</tr>
<tr>
<td>Tremors</td>
<td>Organic voice tremor, Amyotrophic lateral sclerosis, Multiple sclerosis, Wilson’s disease</td>
<td>Extrapyramidal system</td>
<td>Involuntary, rhythmic, purposeless, oscillatory movements</td>
</tr>
<tr>
<td>Mixed</td>
<td>Multiple motor systems</td>
<td></td>
<td>Muscular weakness, limited range and speed</td>
</tr>
</tbody>
</table>

dystonia, and choreoathetosis. Speech samples were rated along 38 dimensions that described pitch characteristics, loudness, vocal quality, respiration, prosody, articulation, and general impression. Results of these studies indicated that each of the seven neurological disorders could be characterized by a unique set of clusters of deviant speech dimensions and that no two disorders had the same set of clusters. Thus, differential diagnosis among the dysarthrias can be made, in part, on the basis that one type of dysarthria sounds different from the others. However, single features such as imprecise consonant production or nasal emission may not be sufficient to distinguish one type of dysarthria from another. Instead, differential diagnosis is made on the basis of clusters of features reflecting underlying pathophysiology and the findings on examination of the musculature. The following summarizes the major types of dysarthrias, their primary distinguishing perceptual attribute, and their presumed underlying localization and distinguishable deficit.

Flaccid Dysarthria
Flaccid dysarthria is due to weakness in cranial or spinal column nerve innervations to the speech systems. Its specific characteristics depend on which nerve is involved. Nerves affecting articulation include trigeminal, facial, or hypoglossal. The vagus nerve contributes to voice and resonatory dysarthria primarily seen with breathy, hoarse, and diplophonic speech. Spinal respiratory nerves can also be affected with deficits in breath patterning for speech, often causing production of short phrases. Most common speech characteristics include hypernasality, breathiness, diplophonia, nasal emission, audible inspiration (stridor), short phrases, and rapid deterioration and recovery with rest. Brainstem stroke or brain injury is a common cause of flaccid dysarthria (14).

Spastic Dysarthria
Spastic dysarthria is usually associated with bilateral lesions of upper motor neuron pathways that innervate the relevant cranial nerve and spinal nerve. Its distinguishable characteristics are attributed to spasticity, and they often include a harsh, strained vocal quality, slow speech rate, pitch breaks, and variable loudness. All speech systems are typically affected with this classification (14).

Ataxic Dysarthria
Ataxic dysarthria is associated with lesion’s disturbances of the cerebellum or its controls. As it is known for its incoordination as a characteristic, individuals are perceived with articulation and prosodic feature disturbances. Irregular articulatory breakdowns, distorted vowels, and inappropriate variations in pitch, loudness, and stress often are the classic features (14).

This type of dysarthria has been coined with the term “drunken speech.” Ataxic dysarthria is often a result of cerebellar stroke or spinocerebellar ataxia.

Hypokinetic Dysarthria
Hypokinetic dysarthria is associated with basal ganglia control of the central nervous system. Its features are mostly related to reduce range of motion and rigidity. Additionally, reduced loudness, short rushes of speech, breathy-tight dysphonia, and monoloudness and monopitch are notable characteristics. Often dysfluency and word repetition are reported. Parkinson’s disease and its syndromes are the most noted disorder for hypokinetic dysarthria (15).

Hyperkinetic Dysarthria
Hyperkinetic dysarthria is also associated with basal ganglia control; however, unlike hypokinetic dysarthrias, it is distinguished by abnormal involuntary movements that affect the intended speech movements. It is classified in many of the movement disorders and subcategorized into slow and fast hyperkinetic dysarthrias. Slow includes dystonia and athetosis, while fast includes tic and chorea. They may affect any of the speech systems and area usually distinguished by unpredictable variability in voice and articulation. Distorted vowels, excess loudness variations, sudden forced inspiration/expiration, voice stoppages/arrests, transient breathiness, intermittent hypernasality, and inappropriate vocal noises are just some of the common characteristics. Etiologies include Huntington’s chorea, Tourette’s syndrome, cerebral palsy, and side effects of neuroleptic drugs (16).

Unilateral Upper Motor Neuron Dysarthria
This dysarthria has an anatomical rather than a pathophysiological label. It typically results from stroke affecting upper neuron pathways. Damage is unilateral; severity usually is rarely worse than mild to moderate. Often its characteristics overlap with flaccid, spastic, or ataxic dysarthrias (16).

Mixed Dysarthria
Mixed dysarthria reflects combinations of two or more of the single dysarthria type. It occurs more frequently than any single dysarthria type. A common diagnosis involving mixed dysarthria is amyotrophic lateral sclerosis, which, in advanced stages, has features of both flaccid and spastic dysarthrias (15).

Assessment
Traditionally, the assessment of dysarthric speech has mainly involved the use of perceptual evaluation measures. The inadequacies of this method of assessment have resulted in the development of a range of instrumental assessment techniques to provide more objective analyses of the underlying physiological impairments of the speech mechanism. Generally, the assessment of dysarthria relies on a combination of perceptual and instrumental analysis, including acoustic analyses. Most of the perceptual scales are based on the studies by Darley et al., using a number of speech dimensions to be rated on a severity rating scale (14). Intelligibility scores has been one of the main measurement parameters to indicate communicative functional status. The interactive product of respiration, phonation, articulation, resonance, and prosody helps to define the speech impairment and requires full attention during assessments. Often, little attention is placed on connected speech intelligibility in the dysarthria population; however, it may
be a more valid indicator of functional communication than assessment at the word or simple phrase level (17).

Studies have demonstrated acoustic variations in dysarthria; however, many of the studies have focused on one element in acoustic analysis rather than a broadly directed analysis of the speech output by systems, for example the articulatory, resonatory, phonatory, and respiratory components (18–22).

A dysarthric speech assessment must involve perceptual acoustic description of all the systems. Clinicians must understand the interrelationship of the weakness, slowness, discoordination, or abnormal tone of the speech musculature and its effect on the systems of the speech mechanism, including respiratory, phonatory, velopharyngeal, or oral articulatory subsystems (23). Both perceptual and instrumental tools are available for measuring speech performance (24,25). The perceptual tools are those that rely on the trained eyes and ears of the clinician, whereas instrumental approaches to assessment include devices that provide information about the acoustic, aerodynamic movement, or phonatory aspects of speech.

Assessment of the respiratory subsystem begins with perceptual measures, including ratings of the number of words produced per breath, the loudness of samples of connected speech, or visual observations of the presence of clavicular breathing. Instrumental approaches to the measurement of respiratory function may include acoustic measures of vocal intensity and utterance durations. Respiratory performance may also be assessed by estimating the subglottic air pressure generated by the speaker (26,27). Respiratory inductive plethysmography, commercially available as the Respirtrace, is an instrument capable of obtaining information about the movements of the rib cage and abdomen during breathing and speech.

Phonatory (laryngeal subsystem) assessment typically begins with perceptual ratings of pitch characteristics (e.g., pitch level, pitch breaks, monopitch, and voice tremor), loudness (e.g., monoloudness, excess loudness, variation of volume), and voice quality (e.g., harsh voice, hoarseness, wet voice, breathiness, strained-strangled voice). Acoustic analysis can be performed by deriving vocal fundamental frequency and intensity (28,29). Measures of laryngeal resistance to airflow can also be obtained (30) along with laryngeal visualization with endoscopy or stroboscopy (30).

Assessment of the velopharyngeal mechanism can be made with perceptual judgments of hypernasality or the occurrence of nasal air emission. Nasalization also can be measured acoustically. Precise inferences can be made about the timing of velopharyngeal closure by obtaining simultaneous pneumatic measures of air pressure and air flow during selected speech samples (31,32). Movement of the velopharyngeal mechanism can be observed through cineradiographic techniques and/or endoscopic visualizations.

Assessment of oral articulation can be made by the rating of consonant and vowel precisions and coordination of articulatory movement. Although movements can be recorded using cineradiographic technique and myoelectric activity with electromyographic recordings, these techniques are not used in routine clinical practice.

**Treatment Considerations**

Clinical decisions regarding treatment of dysarthria should include behavioral objectives for reaching short- and long-term goals. An overall goal to improve functional communication for the patient is necessary and should be identified and described with the patient and his or her significant family members. Generally, treatment goals can vary with the severity of the speech impairment and the overall medical disorder.

For severely involved speakers, whose intelligibility is so poor that they are unable to communicate verbally in some or all situations, the general goal of treatment involves establishing an immediate functional means of communication. This may include use of augmentative approaches. The term communication augmentation refers to any device designed to augment, supplement, or replace verbal communication for someone who is not an independent verbal communicator. These augmentative or alternative communication (AAC) systems can be as low tech as writing or communication boards, or higher tech such as talk-back switches or computer-based speech synthesis. The selection of an appropriate augmentation system requires a thorough evaluation of the individual’s communication needs. To determine an individual’s needs, the clinician needs to consider the patient’s physical and cognitive capabilities, including cognition, language, memory, physical control, vision, hearing, and overall medical condition. This assessment may require additional co-evaluators, such as an occupational therapist or a rehabilitation engineer. Once the individual’s capabilities have been ascertained, augmentative system components can be selected, and appropriate system modifications can be made.

For those moderately involved speakers who are able to use speech as their primary means of communication but whose intelligibility is compromised, the general goal of treatment involves improving intelligibility. Use of compensatory strategies with speech production or augmentative systems is common. Achieving compensated intelligibility may take a variety of forms, depending on the speaker and the nature of the underlying impairment. For some with greater involvement, use of an alphabet supplementation system, in which they point to the first letter of each word as they say the word, assists in the transition to intelligible speech (9). For others, the treatment involves an attempt to decrease the impairment by exercises that will improve performance on selected aspects of speech production. For example, exercises may involve developing more adequate respiratory support for speech (33) or training to establish an appropriate speech rate (34). For the mildly involved dysarthric speaker whose speech is characterized as intelligible but less efficient and less natural than normal, treatment planning should consider the patient’s needs in communication within his or her home or job setting. For some speakers, these mild reductions in speech efficiency pose no problems. For other mildly involved speakers, however, treatment is warranted. The general goals of treatment...
for dysarthric people with mild disabilities include maximizing communication efficiency while maintaining intelligibility and maximizing speech naturalness. Maximizing naturalness is accomplished by teaching appropriate phrasing, stressing patterning, and intonation (35).

Often partnerships with other medical professionals are appropriate. For example, a maxillofacial prosthodontist may assist with managing resonant disorders such as velopharyngeal incompetence. Fabrication of a prosthetic device such as a palatal lift can assist with management of impaired velopharyngeal closure. Assessment of candidacy for the device is a coordinated effort between the speech-language pathologist and the prosthodontist. An appropriately fitted palatal lift will allow certain dysarthric speakers to better produce speech sounds that require the buildup air pressure and can maximize intelligibility by improving prosody and contextual breath support.

Treatment approaches for patients with progressive disorders such as parkinsonism, multiple sclerosis, and amyotrophic lateral sclerosis are different from those used with the dysarthric speaker who is recovering from a single medical event (36). Initially, the patients are encouraged to maximize the functional communication level by paying specific attention to the clarity and precision of their speech. At some point, the patients will need to modify their speaking patterns by controlling rate and consonant emphasis and by reducing the number of words per breath. Some patients with progressive dysarthria make the adjustments in their speech pattern without specific treatment; others may need to practice these modifications with a speech pathologist or trained family member until the changes become habitual. In severe cases, a communication augmentation system may be considered. These augmentation systems usually are chosen or designed to accommodate the lifestyle of the patient while serving his or her anticipated communication needs over the longest period of time.

Apraxia of speech (AOS) occurs in the absence of significant weakness and incoordination of muscles, with automatic and reflexive movements undisturbed. Lesions in the premotor cortex are a frequent finding (37). AOS is characterized by labored and dysprosodic productions, resulting in errors of omission, substitution, and repetition. There is debate as to whether AOS is a pure motor or linguistic (i.e., phonemic) disturbance (38–40). Patients have difficulty programming the positioning of the speech musculature and sequencing the movements necessary for speech. It is seen by some as a distinct condition that often coexists and complicates aphasia, whereas others regard the characteristics as part of the nonfluent Broca’s aphasia. AOS carries a negative prognosis for recovery when there is a moderate to severe aphasia in tandem. When it occurs without the concomitant language disturbance, therapy can focus on retraining the patient’s ability to program sound patterns, to shift from one sound to another, and to use preserved melodic and rhythmic patterns to facilitate speech.

Head and Neck Cancer

The diagnosis of head and neck cancer often presents with speech and/or swallowing impairments. The location of the malignancy often dictates the speech impairment being phonatory, resonance, articulatory, or a combination of the three. Treatment of the disease can be a sole modality such as surgery or radiation therapy or multiple modalities, including surgery, irradiation, and chemotheraphy. Over the last 20 years, treatment has focused on organ preservation with the goal of maintaining function. This type of treatment is often recommended for tumors of the tongue, oral cavity, tonsil, base of tongue, and pharynx and larynx. Depending on tumor size, spread of disease, and nodal involvement, the treatment often involves radiation with or without chemotherapy. Speech impairments thus would typically be seen in the acute phase of treatment and in the immediate posttreatment acute phase (6 months posttreatment). Chronic effects are often seen with scarring and fibrosis greater than 6 months posttreatment.

Often surgical intervention is needed with advanced disease or recurring disease. This may involve removal of the larynx alone or with other organs. Removal of the larynx or total laryngectomy is a common procedure. With loss of the voicing component for speech production, there is an obvious need for speech rehabilitation, beginning with presurgical teaching and continuing with communication alternatives and rehabilitation postsurgery.

Several options for speaking are available to postlaryngectomy patients including artificial larynx, esophageal speech, or tracheal esophageal puncture, a voice restoration procedure. Artificial larynx use or electrolarynx offers individuals the opportunity to speak within days of surgery. Commercially available electrolarynges are designed to introduce air vibrations either directly into the oral cavity through a catheter or indirectly through the neck tissues. In each case, the tones resonate within the oral and pharyngeal cavities and are modified by articulation into audible, intelligible words. Often an intraoral electrolarynx can be used 2 or 3 days after surgery, providing the patient with an immediate means of communication. Intraoral devices can also be used a long term for patients with necks that are unsuitable for direct transmission of vibration, usually because of pain, edema, or scar tissue. Experience suggests that good speech is slower to develop using an intraoral device than with a neck device. Therefore, care must be taken to help the patient avoid early frustration associated with not being understood immediately.

A second form of speech alternative is esophageal speech. Esophageal speech is accomplished by training the patient to move air from the oral and pharyngeal cavities into the esophagus by injection or inhalation methods. The air is then trapped within the cricopharyngeal segment and released. The vibration of the release from the segment is the phonatory sound that is then modulated with resonatory and articulatory modifications to produce speech. Accomplished esophageal speakers can speak clearly and effortlessly; however, many laryngectomy patients are unable to learn this technique. Failure to learn esophageal speech may represent insufficient or excessive
Fluency

Fluency disorders are characterized by a disruption in the ease and flow of connected speech. The most common and well-known type of fluency disorder is stuttering. Stuttering is a very complex, dynamic, and somewhat controversial disorder in that theories and opinions abound regarding etiology, diagnosis, and treatment. Theories regarding underlying cause include genetic, cognitive, psycholinguistic, neuromuscular, as well as multifactorial. It is generally accepted that primary characteristics of stuttering include blocks (absence of sound), repetitions (of sounds or words), and/or prolonged sounds. Secondary behaviors often observed include struggle (body movement, eye blinks, lip/jaw tremors) and avoidance of sounds, words, conversation partners or environments that trigger the dysfluency. The evaluating clinician must be very cautious with diagnosis for several reasons. One, stuttering resolves in 75% to 90% of small children who begin to stutter. Additionally, there is a spectrum of normal dysfluency that resolves in 75% to 90% of small children who begin to stutter. Labeling one as a “stutterer” when he or she is within the normal spectrum of behavior can be quite damaging and can lead to exacerbated symptomatology.

Due to high variability of type, severity, response to treatment, and possibly cause, having a standard treatment strategy is nearly impossible. However, common approaches include environmental modifications, desensitization, fluency-shaping techniques (i.e., easy onset, continuous phonation, auditory/visual feedback), and possibly psychological intervention. One component that is quite consistently included is counseling and support by the speech-language pathologist with the goal of changing the patient’s mindset from anticipating stuttering to anticipating fluency. Goals may vary from absence of stuttering to reducing severity and eliminating avoidance behaviors so as to improve quality and effectiveness of overall communication. It is generally believed that any effective treatment program must also have transference into nondisabled environments and maintenance as a goal (46,47).

LANGUAGE

As described earlier, language is an arbitrary system of signs or symbols used according to prescribed rules to convey meaning within a linguistic community (1). An understanding of the mechanisms responsible for the processing and formulation of language is critical to good rehabilitation practice. Various theories have been developed to define language. Linguistic theory of language attempts to account for the rules and structure of language stored in memory; whereas the psycholinguistic theories of language attempt to account for how language is stored and accessed to produce the various verbal forms such as sounds, words, and phrases and discourse. Additionally, the role of attention is needed and has been studied extensively in language processing (48).

Traditional study of language has focused on specified sites being responsible for specific functions. There are many lines of evidence to support the concept that the left hemisphere plays a crucial role in language processing. Goodglass, in 1993, reported that approximately 95% of cases with aphasia had focal lesions in the left hemisphere (49). The traditional approach to language views damage to the third frontal convolution. Broca’s area, as causing language production problems, whereas damage to the first temporal convolution, Wernicke’s area, is associated with language comprehension. This approach is simplistic; more recent research has proven correlation of neurologic lesion with language function to be much more complex. Contributions from the right hemisphere have been noted to give input regarding prosody and comprehension of complex syntax (50). Additionally, other subcortical structures such as the basal ganglia and thalamus have been identified as having roles in language function (51). For example, thalamic hemorrhage has been reported to affect language to varying degrees, with some individuals having almost normal language performance and others demonstrating marked paraphasias and periods of fluctuating consciousness (52,53). These fluctuations may be related more to the role of the thalamus in arousal and selective attention as prerequisites to communication than to its role in actual deficits of language (52). Although most studies confirm the notion that the symptoms associated with subcortical disease may be transitory, individuals who evidence attention and arousal deficits...
beyond the acute stage of illness have diminished ability to learn compensations for their communication failures in spite of nearly normal comprehension skills.

**Aphasia**

Aphasia is an acquired disorder of all language modalities, including verbal expression, auditory comprehension, written expression, and reading comprehension. It interferes with the ability to manipulate the meaning (i.e., semantics) or order (i.e., syntax) of words, spoken or written. Three important points to emphasize in this definition include

1. The term *aphasia* implies impairment in both receptive and expressive language modalities. Expression may be more severely involved than reception, and reception can appear grossly intact. If the testing instrument is sensitive to subtle change in language behavior, pathology can be identified in the more intact modality.

2. Aphasia is consistent only with focal disease, usually of the left hemisphere. Aphasic symptoms may be part of a diffuse pathologic condition. However, these patients evidence more than disruptions in their ability to manipulate linguistic symbols, such as disorientation. Prognosis and recovery for this group are markedly different from those who evidence aphasia alone.

3. Although it is well-known that aphasic disturbances are usually a consequence of cortical disease, the identification and classification of more atypical aphasic syndromes are also associated with subcortical infarction and hemorrhage (54).

**Language Characteristics**

Verbal expressive characteristics of aphasia include anomia, agrammatism, paragrammatism, or paraphasia or the production of jargon, stereotypic, or echolalic language patterns. Although most aphasics display an overall reduction in word classes available for production, they show particular deficits in the retrieval of nouns (i.e., anomia). Because nouns carry a large part of the meaning during an intended message, the language of the anomic patient is described as "empty" because sentences often lack a subject or referent.

In their attempts to retrieve words, aphasics may make "paraphasic" language errors. When the substitution for the intended word is from the same word class, such as *chair* for *table*, it is a semantic paraphasia. The substitution of like sounds or syllables, such as *flair* for *chair*, is classified as a phonemic paraphasia. A final class of grammatical error is the neologism. Neologisms are attempts at the target that bear no phonemic or semantic relationship to that target, such as, "I want to brush my ploker." Patients who find word retrieval difficult may also circumlocute, or talk around the intended noun, such as saying, "I wear it on my wrist" instead of *watch*.

Agrammatism is a form of expressive deficit characterized by reliance on nouns and verbs (i.e., content words) to the exclusion of articles, verb auxiliaries, pronouns, and prepositions (i.e., function words). Agrammatic productions often are described as telegraphic. Paragrammatic language is characterized by the misuse, rather than the omission, of grammatical elements.

Individuals who have expressive output that is largely incomprehensible, even though the utterance is well articulated and verbose, may display a form of expressive deficit called jargon. Concentrations of neologisms are called neologistic jargon and may be associated with stereotypes such as "blam, blam, blam" substituted for all attempts at verbalization. A preponderance of unrelated semantic paraphasias is semantic jargon. Finally, some individuals evidence echolalia, typified by the patient echoing back the same utterance he or she has just heard.

Written expression also is impaired in typical aphasia, characterized by letter or word substitutions, omissions, or additions (paragraphias) or errors in syntax. The ability to use gestures as a substitute form of expression can be impaired as well (38).

Table 15-2 provides a summary of the terminology used to describe expressive language deficits in aphasia.

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agrammatism</td>
<td>The absence of recognized grammatical elements during speech attempts</td>
</tr>
<tr>
<td>Anomia</td>
<td>Difficulty producing nouns</td>
</tr>
<tr>
<td>Circumlocution</td>
<td>Attempts at word retrieval and in descriptions or associations related to the word</td>
</tr>
<tr>
<td>Echolalia</td>
<td>An accurate repetition of a preceding utterance when repetition is not required</td>
</tr>
<tr>
<td>Empty speech</td>
<td>A fluent utterance that lacks substantive parts of language, such as nouns and verbs</td>
</tr>
<tr>
<td>Jargon</td>
<td>Mostly incomprehensible but well-articulated language</td>
</tr>
<tr>
<td>Neologic jargon</td>
<td>Mostly incomprehensible, some words are partially recognizable, others are contrived or “new”</td>
</tr>
<tr>
<td>Paragrammatism</td>
<td>Misuse of grammatical elements, usually during fluent utterances</td>
</tr>
<tr>
<td>Phonemic paraphasia</td>
<td>Flair for <em>chair</em>, also called literal paraphasia</td>
</tr>
<tr>
<td>Press for speech</td>
<td>Excessively lengthy, often incomprehensible, well-articulated language</td>
</tr>
<tr>
<td>Semantic jargon</td>
<td>A combination of unrelated semantic and phonemic paraphasia, together with recognizable words</td>
</tr>
<tr>
<td>Semantic paraphasia</td>
<td>Table for <em>chair</em>, also called nominal paraphasia</td>
</tr>
<tr>
<td>Stereotypes</td>
<td>Nonsensical repetition of similar syllables for all communicative attempts, such as <em>dee dee dee</em></td>
</tr>
<tr>
<td>Telegraphic speech</td>
<td>Language similar to a telegram, mostly nouns and verbs</td>
</tr>
</tbody>
</table>
Auditory comprehension characteristics of aphasia include deficits of auditory perception and auditory retention. Auditory misperceptions are characterized by a tendency to confuse words that are similar in either meaning or sound. These confusions create a distorted message resulting in errors of comprehension. Most aphasics will experience more errors in comprehension as the length of the auditory input increases. In general, the speed of auditory input, combined with increased length, leads to errors in auditory retention. In addition, increased sentence length often presupposes a more difficult syntax and vocabulary, combining to make comprehension more liable to error. It has been demonstrated that some aphasics retain more information from the beginning of an utterance, whereas others retain information from the end (55). Evaluation of this aspect of the patient’s auditory capacity is especially important if rehabilitation is to succeed. Comprehension of reading material also is impaired. The severity of this impairment often is greater than that of the linguistic deficits in other modalities. However, for some, reading comprehension is a relative strength and can be used to augment or facilitate comprehension of the spoken word by pairing the spoken message with a written message.

**Aphasia: The Classical Descriptions**

Historically, there have been many attempts to place pathologic language symptoms into homogenous groups, permitting reference to specific aphasic subtypes. The Boston classification system standardizes terminology by classifying disorders into those in which expressive skills are predominantly fluent and those in which they are predominantly nonfluent (56). Although such a distinction might be useful clinically, it often can be difficult to make this classification, as in the case of a conduction aphasic (a fluent aphasia) who may have long pauses and expressive struggle (nonfluency) during speech. The eight major types of aphasia in the Boston system include the more common forms of Broca’s aphasia, Wernicke’s aphasia, anomia, conduction aphasia, and global aphasia, as well as the less frequent transcortical types, transcortical motor and transcortical sensory (Table 15-3). Each of these aphasic syndromes has discrete symptoms and is correlated with a specific localized cortical lesion, some with subcortical extension.

**Nonfluent Aphasias**

**Broca’s Aphasia**

One of the classic localization theory aphasias, with site of lesion in Broca’s area (or area 44), located in the third frontal convolution anterior to the precentral gyrus. Often called “expressive aphasia,” this aphasia is characterized by expressive skills more greatly impaired than receptive skills. Agrammatic verbal output is a hallmark symptom, with word retrieval more intact than sentence formulation. Error awareness is often good, making the patient a relatively good communicator, as compared with those with other types of aphasias (57). Repetition is typically poor. Reading and writing also show a range of impairments. AOS frequently accompanies Broca’s aphasia.

**Transcortical Motor Aphasia**

Currently, this type of aphasia is believed to be of a smaller site of lesion, located in the frontal lobe, superior and anterior to Broca’s area (57). Language function is similar to those of Broca’s aphasia with the exception that repetition is relatively preserved.

**Global Aphasia**

Associated with a large left hemisphere lesion, typically including both Broca’s and Wernicke’s areas. It is considered the most severe aphasia with significant deficits in all language modalities. Often, automatic expressions such as counting or profanity are preserved. Also, frequently these patients can use other modalities, such as facial expression and/or gesture, to communicate basic wants, needs, or feelings (57).

**Fluent Aphasias**

**Wernicke’s Aphasia**

The other classic localization theory aphasia, Wernicke’s aphasia is associated with lesion in the Wernicke’s area (area 22), the posterior portion of the superior temporal gyrus (57). Often called “receptive aphasia,” this aphasia is characterized by receptive skills being more severe than expressive skills. Patients with this type of aphasia often produce sentences with intact grammar and rhythm of speech but frequent paraphasias of both types and/or frequent neologisms or jargon. Due to poor auditory comprehension, error awareness is also poor, usually making for a less effective communicator than one with Broca’s aphasia. Repetition is also impaired.

---

**TABLE 15.3 Summary of the Boston Classification System of Aphasia**

<table>
<thead>
<tr>
<th>Type</th>
<th>Language Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Nonfluent</strong></td>
<td></td>
</tr>
<tr>
<td>Broca’s</td>
<td>Telegraphic, agrammatic expression often associated with apraxia; good comprehension except on more abstract tasks</td>
</tr>
<tr>
<td>Transcortical motor</td>
<td>Limited language output; fair naming; intact repetition; fair comprehension</td>
</tr>
<tr>
<td>Global</td>
<td>Severe expressive and receptive reduction in language</td>
</tr>
<tr>
<td>Mixed transcortical</td>
<td>Severe reduction in expression and reception; repetition intact</td>
</tr>
<tr>
<td><strong>Fluent</strong></td>
<td></td>
</tr>
<tr>
<td>Anomia</td>
<td>Word-finding difficulty without other serious linguistic deficits</td>
</tr>
<tr>
<td>Conduction</td>
<td>Phonemic paraphasic errors; good comprehension; fluency in bursts; deficits in repetition of low-probability phrases</td>
</tr>
<tr>
<td>Wernicke’s</td>
<td>Phonemic and semantic paraphasias; poor comprehension</td>
</tr>
<tr>
<td>Transcortical sensory</td>
<td>Fluent neologic language; poor comprehension; intact repetition</td>
</tr>
</tbody>
</table>


Transcortical Sensory Aphasia
Lesion is typically in the inferior temporo-occipital border area. It is similar to Wernicke’s aphasia, but repetition is relatively preserved.

Conduction Aphasia
Thought to be related to subcortical lesions in the arcuate fasciculus, an association tract running beneath the cortex, connecting temporal and parietal lobes and carrying impulses between Wernicke’s area and Broca’s area. In conduction aphasia, repetition is disproportionately impaired relative to auditory comprehension and verbal expression (57). Verbal output is generally grammatical and fluent but has episodes of halting speech during moments of word retrieval difficulty.

Anomic Aphasia
It is associated with lesion site of the posterior parietotemporal juncture. It typically is the mildest form of aphasia and is marked by word retrieval difficulties, with syntax and fluency generally intact. Verbal output is characterized by frequent semantic paraphasias or overgeneralizations for the intended words. Comprehension impairment is mild (57).

Differentiation from Other Disorders
Aphasia, particularly in the acute stages, may be difficult to differentiate from other disorders that compromise communication. Accurate differentiation is necessary because each communication disorder requires separate treatment and management approaches. It should be noted that aphasia may occur in conjunction with other syndromes. Below are some disorders that compromise communication but should be differentiated from aphasia.

Agnosia
Agnosia is the inability to interpret or recognize information when the end organ is intact. For example, a patient with auditory agnosia would have normal audiometric hearing thresholds but cannot interpret speech signals at the cortical level. Hence, auditory comprehension will be severely compromised. Patients with agnosia can be differentiated from those with aphasia because they will be impaired in only one modality. For example, the patient with auditory agnosia who has severe comprehension deficits can read the same words through the intact visual modality.

Dementia
Dementia is a syndrome of progressive cognitive deterioration that adversely affects the ability to communicate (58). Although specific expressive and receptive language disturbances can present as part of an underlying disease process, the aphasic patient does not show evidence of cognitive deficits in such areas as orientation, judgment, self-care, and visual-perceptual skills. The distinction between those patients with language deficits secondary to aphasia and those with diffuse disease is particularly relevant in rehabilitation because the prognosis for retraining specific skills and developing independence is more favorable for the patient with aphasia alone.

Language of Confusion
Language of confusion is characterized by reduced recognition, reduced understanding of and responsiveness to the environment, faulty memory, unclear thinking, and disorientation (58). It often is associated with head trauma. In contrast to the language disorders of dementia, the prognosis for recovery after traumatic injury is more favorable, and the course is not progressive.

Aphasia Assessments
Tests for aphasia measure the patient’s receptive and expressive language capacities by sampling different types of language skills through systematically controlled channels. For example, an examination of reception via the visual input system might begin with a concrete task such as copying or matching and then proceed to more difficult tasks such as reading sentences or paragraphs for comprehension. Tests of expression might range from simple repetition to naming to providing definitions or picture descriptions. Most test batteries currently in use provide a representative sample from which inferences can be made about performance in similar linguistic situations. Although most tests of aphasia do sample linguistic competencies, they are not equipped to measure either the least severe or the most severe disorders. Therefore, the examination will have to be supplemented by other specialized formal and informal measures in selected cases. Additionally, most isolated aphasia test batteries do not depict an individual’s functional communication performance. To determine one’s overall communicative functional status, additional testing is often required.

Some of the more commonly utilized aphasia assessments include the Western Aphasia Battery-Revised (WAB-R), Boston Diagnostic Aphasia Examination (BDAE) third edition, and the Communication Activities of Daily Living-2 (CADL-2) (56,59,60). Also frequently given is the Boston Naming Test (BNT), which is actually a subtest of the BDAE. To assess reading comprehension in varied contexts, portions of the Reading Comprehension Battery of Aphasia-2 (RCBA-2) are often given (61). It should be noted that these assessments are rarely given in their entirety. Depending on the degree of the patient’s impairment and the area of deficit the evaluating clinician deems most informative regarding patient function, specific subtests are selected and administered.

Approaches to Treatment
Aphasia treatment should be patient dependent and maximize communication in actual interactive situations. Aphasics, even those with similar types of lesions, represent a heterogeneous group. Because of this heterogeneity, outcomes of treatment are often unpredictable. In addition to building on the patient’s communicative strengths, remediation should be directed toward helping the patient, family, and friends accept and adapt to the person’s impairment.

Traditionally, the focus of aphasia treatment has been on the stimulation-facilitation approach, in which the patient and clinician interact within a stimulus-response framework on tasks that are related to deficit areas determined during assessment
Communication Impairment After Right Hemisphere Damage

Communication impairments with right hemisphere–damaged patients largely relate to cognitive skills that impact communication. A cognitive-communication disorder is defined by the American Speech-Language-Hearing Association (ASHA) as a disorder that encompasses difficulty with any aspect of communication that is affected by disruption of cognition (65). Communication may be verbal or nonverbal and includes listening, speaking, gesturing, reading, and writing in all domains of language (phonologic, morphologic, syntactic, semantic, and pragmatic). Cognition includes cognitive processes and systems (e.g., attention, perception, memory, organization, executive function). Areas of function affected by cognitive impairments include behavioral self-regulation, social interaction, activities of daily living, learning and academic performance, and vocational performance (65).

Due to the speech-language pathologist’s knowledge about normal and abnormal development, brain-behavior relationships, pathophysiology, and neuropsychological processes as related to the cognitive aspects of communication, American Speech Language Hearing Association (ASHA) states that the Speech-Language Pathologist (SLP) is qualified to assume a habilitative and rehabilitation role in managing cognitive-communication disorders (65). Skills affected by right hemisphere damage include attention, orientation, visual or auditory perceptual skills, memory, social communication (pragmetics), organization, reasoning and problem solving (executive functioning), safety awareness, and judgment. Some patients experience anosognosia, an unawareness of the problems experienced, which makes remediation more challenging (65).

The pragmatic aspects, such as the lack of facial expression while speaking, failure to maintain eye contact, failure to use gesture, and lack of vocal inflection, have all been described in the patient with right hemisphere damage (66). Loss of inflectional patterns that signal anger or frustration is not surprising, given the role the right hemisphere plays in decoding and recalling melodies. The clinician must be alert to the possibility of such a pathologic process because loss of these affective cues often is mistaken for the functional states of rudeness or unconcern for another’s feelings.

Additional problems with extraverbal aspects of communication also help to explain the “difficult-to-get-along-with” personality of the patient with right hemisphere damage. Extraverbal skills include behaviors such as appreciation of humor and figurative language and the use of pragmatics such as the ability to maintain a conversational topic and turn-taking during conversation. These patients have difficulty in organizing information, and they fail to make use of contextual cues.
 Failure to recognize these cues results in many conversational irrelevancies, as well as poor monitoring of rate and amount of expressive language when they stray from or completely lose the thread of the conversation. Failure to recognize humor and metaphorical language structures may be misunderstood by family members as a noncaring, depressed personality. Problems in interpreting incoming visual and auditory information, long known to be a right hemisphere function, make it difficult for these patients to get the main idea of discourse, leading to further frustration of friends and family.

**Assessment**

Before the Rice Evaluation of Communication Problems in Right Hemisphere Dysfunction (67), no formal assessment tool was available for the systematic evaluation of the communication skills of the patient with right hemisphere damage (68). This test offers a rating system for the use of pragmatics, nonverbal skills such as eye contact and facial gestures, interpretation of metaphorical language, memory skills, writing, visual scanning and tracking, and analysis of conversation, including topic maintenance, verbosity, and reference. Approaches to treatment based on the analysis are presented.

Informal assessment tools for communication ability were described by West et al. (69). Their screening battery includes exploration of basic language skills, an analysis of single-word responses to part-whole tasks, oral opposites, written opposites, oral analogies, and printed analogies. Other sections include the interpretation of idioms and proverbs, effects of imagery, and evaluation of the patient’s ability to appreciate humor. Clinicians frequently will gather a variety of subtests from different sources to assess areas of cognitive communication, including attention, memory, executive functioning, and visual-spatial skills. Understanding patient personality and functional status prior to neurologic injury is important; therefore, interview of close friends or family members is a critical element in assessment.

**Treatment**

Treatment should concentrate on three broad areas as follows:

1. The communication specialist should develop tasks that help the patient attend to contextual cues in an effort to reduce verbosity and improve topic maintenance, retell stories in a fashion that highlights the main points, and produce language that follows a logical sequence. Although these tasks appear related to the impairment, no data support success or failure using such strategies.

2. Education of the patient’s family members on how the loss of pragmatic and affective languages can affect their perception of the patient’s personality is important. Rehabilitation successes in other areas will be diminished or lost if the patient’s family members do not understand the reasons for the adjustment to a new personality.

3. Many speech-language pathologists also address the cognitive processes that affect communication, such as attention, memory, executive functioning, and safety awareness/judgment. They do this through tasks that are individualized to address the specific deficit areas and can translate to optimizing function in real-world situations. For example, answering questions regarding a map can address visual attention, visual perceptual skills, left attention, and, depending on the complexity of the map and the questions, executive functions such as planning, problem solving, and reasoning. Even more functional is the actual use of a map to route find within an unfamiliar setting. This will address the aforementioned skills but will more greatly challenge attention, executive functioning, and safety awareness.

**Communication Impairment After Traumatic Brain Injury**

Patients with TBI will suffer some degree of cognitive-communication impairment. Those with mild TBI may not initially have symptoms. Instead, difficulties may develop over weeks or months and/or may be subtle and thus may be missed by medical personnel. Such symptoms include memory loss, poor attention/concentration, emotional irritability, easily getting lost or confused, and reduced speed of thinking. The moderate to severe TBI patients will have more evident deficits, usually from the beginning stages. These include:

- Attention
- Concentration
- Distractibility
- Memory
- Speed of processing
- Confusion
- Perseveration
- Impulsiveness
- Language processing
- “Executive functions” (70)

The Rancho Los Amigos Scale—Revised is a tool that describes recovery from TBI in terms of increasing arousal and responsiveness, with reducing confusion and inappropriateness. As such progress takes place, improved cognitive communication and language processing also develops. An interval scale from lowest function to highest is the means for scoring and is depicted in the following manner. Levels I, II, and III require total assistance. Level I describes when the patient has no response to any type of stimuli—visual, auditory, proprioceptive, vestibular, or painful. Level II is when generalized responses (such as increased respiratory rate, whole body movements) occur after the presentation of a stimulus. Level III is when localized responses begin to emerge. For example, the patient will turn toward or away from a specific stimulus and may respond inconsistently to simple commands. During level IV, the Confused-Agitated: Maximal Assistance stage, the patient is more alert and engages in simple purposeful behavior (i.e., scratching an itch) but becomes easily agitated. Verbalizations often are incoherent and/or inappropriate, and attention is very brief. During level V, the Confused, Inappropriate Non-Agitated: Maximal Assistance stage, the agitation reduces; however, it can be triggered easily during overstimulation. The person is generally not oriented and is unable to learn new information. Sustained
attention is slightly longer, but goal directed problem-solving or self-monitoring behavior is minimal to absent. Level VI represents the Confused, Appropriate: Moderate Assistance stage, where orientation is inconsistent, attention to highly familiar tasks improves, and ability to use external memory aids with assistance emerges. New declarative learning and awareness of impairment tend to be severely impaired. During level VII, the Automatic, Appropriate: Minimal Assistance for Daily Living Skills stage, carryover of new learning begins and ability to self-monitor and correct errors emerges. Executive functioning and social functioning are impaired, having a significant impact on function in home and community environments. Level VIII is the Purposeful, Appropriate: Stand-By Assistance stage, where one can use assistive memory devices to recall functional information, attend to familiar tasks for at least an hour in distracting settings, learn new tasks, verbalize awareness of deficits but may not appreciate the impact and take appropriate corrective action. Depression or irritability may be present. Level IX is the Purposeful, Appropriate: Stand-By Assistance of Request stage with more independent alternating attention, with heightened self-awareness and executive functioning, and with standby assistance as needed for optimal success in challenging environments. Frustration tolerance may be low. The final level is X, the Purposeful, Appropriate: Modified Independent stage, in which one may be able to multitask with breaks, independently complete tasks requiring high levels of executive functioning with extra time, and/or use compensatory strategies. Social behavior is appropriate. Frustration tolerance may be low when sick or fatigued (71). Recovery through the stages can happen quickly, in a matter of days or weeks, or happen over months. It is important to be constantly reevaluating the patient to determine current skills/deficits and modify treatment plan accordingly.

The most pervasive cognitive-communication deficits with chronic TBI patients include memory, executive function, and social behavior (72). Working memory, the ability to hold and manipulate information in the mind over short periods of time, and declarative memory (e.g., stored facts, memory for past events, and memory for words), one type of long-term memory, are most affected. Implicit, or procedural, memory, which is the other form of long-term memory and includes habits, skills, and emotional associations, typically remains preserved. Executive functions affect focused attention, planning and organizing, problem solving, verbal reasoning, and metacognition. Social functioning includes emotion recognition, social knowledge, social performance, and execution of these integrated functions in real-time interactions (72).

It should be noted that one of the more challenging aspects of rehabilitating individuals with TBI is the frequent presence of poor self-awareness, or insight. The patient often does not recognize or have an appreciation of the consequences of his or her deficits. This poses a safety risk and can impact patient participation in therapy.

Assessment
The Academy of Neurologic Communication Disorders and Sciences (ANCDS) has recently published practice guidelines for the assessment of persons with TBI. After review of standardized tests frequently given by speech-language pathologists and/or recommended by test publishers and distributors for assessment of communication ability in persons with TBI, a report was generated delineating best assessment tools with regard to test reliability and validity. These include the American Speech-Language and Hearing Association Functional Assessment of Communication Skills (ASHA-FACS), the Behavior Rating Inventory of Executive Function (BRIEF), the CADL-2, the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS), the Test of Language Competence—Expanded (TLC), and the Western Aphasia Battery (WAB) (60, 73–76).

The ANCDS, however, also strongly states that standardized tests are only one component of the evaluation process, which must include multiple sources of information, including the person’s preinjury characteristics, stage of development and recovery, and communication-related demands of personally meaningful everyday activities (72). Most clinicians will utilize a combination of formal and informal measures designed to capture the new onset of cognitive-communication deficits. Based on how these deficits impact the individual’s function in daily life, an individualized treatment plan will be developed.

It should be noted that assessment remains particularly challenging in the very early stages of recovery. Speech-language pathologists often become involved once the patient has begun to demonstrate some degree of localized response, becoming minimally aware of his or her immediate surroundings. The Western Neuro Sensory Stimulation Profile (WNSSP) can be administered to monitor change in patients in the Rancho Los Amigos Neurological Rehabilitation Scale (RLAS) II–V stages. This assessment can be repeated often to demonstrate progress toward increased motor and sensory responses. It also provides a structure in which to assess presence of and best method of a yes/no response to basic questions (77). Once the patient progresses into stage V or VI on the RLAS and has the attention and responsiveness to participate in more formal assessment, then some of the aforementioned tests approved by the ANCDS may become appropriate. However, particularly while the patient remains hospitalized, whether in acute care, acute rehabilitation, or subacute rehabilitation, focus in assessment may continue to be more informal, related to the ability to complete basic functional skills for the hospital environment.

Treatment
Treatment typically involves tasks that improve orientation and memory, help in developing selective attention and discrimination, and emphasize reasoning, executive functioning, and social functioning.

In early stages of recovery, RLAS II–III, the therapist will focus on elicitation of responses to sensory stimulation with the goal of progressing from generalized to localized to purposeful responses. The challenge to the patient in these stages is to recognize, process, and react to sensory input. Environmental management of distracting stimuli that may overload the patient’s senses is important. Next, in RLAS IV–V stages, the
discrimination, organization, and integration of the input become difficult, leading to restlessness and extreme confusion. The goal of treatment in these stages is keeping the patient oriented, improving attention to very basic tasks (card sorting, simple conversation, eating a meal), following simple directions, and developing a trusting relationship that will become particularly important in later stages. The patient usually will begin to speak in these stages. Language of confusion or anomia may be present and should be addressed. Again, managing the environment is also important. Too much stimulation can very much disrupt the integration of sensory input and increase agitation and restlessness. During RIAS VI and VII stages, the focus is on developing memory strategies such as the use of an external memory aid (memory book) to help the patient recall daily activities and important facts. Also addressed is improving attention in tasks of increasing length while controlling for environmental distractors. Executive functioning and social functioning begin to be targeted in these stages, but poor self-awareness can be a limiting factor.

In stages IX and X, the TBI patient is beginning to have the ability to metacognite, and self-awareness is improving. Treatment focus is on the memory, attention, executive functioning, and social functioning skills that the patient requires for returning to community or work environment. Tasks should integrate most or all skills and remain very functional, such as planning a dinner menu to then be executed at home. The patient can report back with analysis regarding successes and challenges. Computer-based therapy is also commonly used, via functional use of the internet or with specific software that addresses attention, memory, and executive functioning.

The ANCDS has reviewed all Class I, II, and III research studies related to cognitive rehabilitation and published evidenced-based practice guidelines with regards to attention, memory, behavioral intervention, problem solving, and self-awareness.

In acute rehabilitation, there are currently no well-controlled studies showing that direct therapy improves outcomes beyond what can be accounted for in spontaneous recovery. However, given that procedural learning remains intact, there is evidence that certain skills and strategies can be learned in this early stage. In the chronic stages, there is good evidence that attention training, intervention for behavioral problems, use of external memory aids, and treatment of problem solving are beneficial in certain contexts.

- Practice guidelines for treatment of attention indicate that individualized attention exercises in at least weekly treatment sessions 1 hour in duration lead to positive outcome measures in tests sensitive to attention and working memory as well as outcome measures that include activity-based self-report (72).
- Practice guidelines related to external memory aids indicate that they do help people with memory problems. Instructional techniques for any new learning, but particularly with use of memory aids, should rely on procedural learning. This technique is commonly referred to as spaced retrieval training (78) and relies on high amounts of errorless, distributed practice, with prompts faded over time (72,78).
- Practice guidelines for training of self-awareness in young and middle-aged adults indicate that by increasing understanding of the individual’s injury and its implications and providing feedback to improve self-awareness, self-monitoring, and self-control of disruptive behaviors in specific contexts, improvement is possible. There is no evidence of generalization; therefore, the treatment needs to be in a context as specific to the target task as possible. Likewise, in young and middle-aged adults, there is evidence that problem solving and behavior disorders can be improved in personally relevant activities (72).

An important role of the speech-language pathologist throughout the continuum of care is family training. How to speak to the patient, utilize external aids, and modify the environment by managing stimulation are all important. Also, how best to intervene to manage disruptive social behavior can be helpful information. Family members are sometimes required to live with someone with a new personality and ability to adapt can be very difficult. Referrals to psychology or social work are often necessary.

**Language Deficits and Dementia**

Patients with dementia present with a generalized intellectual impairment that compromises communication efficiency. In general, those patients who are labeled “demented” have acquired memory impairment and other disorders of cognition. The severity of the dementia may correlate with the degree of communication impairment. Wertz characterized the linguistic performance of these patients as a deterioration of capacities in all communication modalities consistent with the deterioration of all other mental functions (79). Because dementia usually results from a progressive, diffuse pathologic process, patients rarely improve, and it is difficult to identify any learning strengths. There is also recent evidence to show that language may be a barometer of change in charting the deterioration in dementia (80). Knowledge of these changes may be useful in family counseling and subsequent management of the communication disorder.

The inability of a demented patient to communicate a message that is relevant, timely, and completely understood by the listener often depends on the severity of the disruption of those cognitive constructs that subserve language, such as general orientation, attention, memory, and visual-perceptual integration. Although some patients in the early stages of dementia complain of an inability to produce nouns as their major communication deficit, this is not a consistent finding in all patients. Most sensitive to linguistic compromise is the patient's ability to name items in a category of either semantically or letter-related items (81). Progressive deterioration of memory skills particularly coincides with interruptions in normal communicative discourse. Inability to monitor (i.e., identify, sort, remember, or integrate) auditory
Assessment
Assessment of linguistic skills usually is informal but should be accomplished at regular intervals to chart progression. Obler suggests the following categories for evaluation: orientation, naming (i.e., confrontation and responsive), analysis of discourse, comprehension, repetition, verbal fluency, idioms and proverbs, sentence construction, number facts, automatic speech, and reading and writing (83). Tests of memory, because of its close relationship to linguistic ability, also need to be administered at regular intervals. The Dementia Rating Scale (84), although not specific to communication, is a psychometrically sound tool used to assess general deficits of cognition. The five subscales assessed include attention, initiation, perseveration, construction, conceptualization, and memory. The impact of memory on linguistic communication can be measured by the Arizona Battery for Communication Disorders of Dementia (85). The test is comprised of 16 sub-tests, four of which—story retelling, delayed; word learning, free recall; word learning, total recall; and word learning, recognition—can be used as a screening tool to help differentiate the linguistic communicative deficits of patients with Alzheimer’s disease from those with normal skills.

Treatment
Treatment for patients with communication disorders secondary to dementia should be supportive. Treatment goals should include environmental controls, capitalization on the preserved procedural memory, and family education. Environmental structure provides orientation and reinforces memory that, in turn, will improve the accuracy of linguistic attempts. Interactions with the patient should be structured to reduce demands beyond the limits of the cognitive system as established by psychometric evaluation. Frequent measures of linguistic function are part of the treatment and are necessary so that family and friends can be informed about how much to expect from the patient both receptively and expressively. This knowledge will reduce frustration for the patient and family. Some patients retain skills longer than others do, and knowledge of the best input and expected output modalities is useful in management. Because dementia often results in diminution of all linguistic modalities, performance may be strengthened through the use of multiple channels such as the combination of gestural, written, and verbal language. This multimodality strategy is a process that should be taught to family members.

Recent evidence suggests that combining exercise activities with cognitive stimulation tasks may facilitate the communicative process (86). This may be related either to increased blood flow to the brain during periods of aerobic exercise or to stress management. Therapy to enhance communicative skills applied during rest periods from aerobic exercise also may be of importance for patients with dementia (87). Although the evidence that certain medications facilitate memory in patients with dementia remains controversial, no data are available to support the use of facilitative drugs to enhance language.

SWALLOWING IMPAIRMENT
Swallowing is an essential life function that begins in utero. It is necessary to survival both because it is the source of hydration and alimentation and because it has a crucial role in maintaining airway integrity by clearing residue from the oral cavity and pharyngeal tract. Abnormal swallowing (dysphagia) may lead to dehydration, starvation, aspiration pneumonia, airway obstruction, and—in the worst cases—death. Dysphagia is frequently associated with cerebrovascular disease, TBI, head and neck cancer, and other conditions common in rehabilitation patients. In many cases, swallowing impairments are amenable to rehabilitation treatment.

PHYSIOLOGY OF SWALLOWING
Swallowing is coordinated with other patterned behaviors, including respiration and mastication. Swallowing is divided into several phases in an effort to simplify description. The phases (i.e., oral preparatory, oral propulsive, pharyngeal, and esophageal) are based on anatomic location and flow of the bolus (88–91). These phases are neither discrete nor isolated from each other and may overlap in time. Events in one phase may impact events in another phase or in multiple phases, ultimately affecting the physiology of the whole swallow.

Anticipation Before the Swallow
The concept of anticipation is a relatively new construct in the physiology of swallowing. This includes the perception and cognition regarding nutritional materials prior to introduction
to the mouth. Specifics include the size of the bolus prepared for intake, the fine motor control for placing the bolus on a fork or spoon or sipping from a straw or cup, instructions for feeding, and the speed at which food or drink is presented to the mouth, all of which can affect the physiology of eating, drinking, and swallowing. Some investigators have proposed calling this an Anticipatory Phase of swallowing (89,90,92–94).

**Oral Preparatory Phase**
The oral preparatory phase is a necessary step prior to the initiation of the oropharyngeal swallow. The manner in which the bolus is prepared for swallowing varies, depending on the consistency of the material. The preparation for solid food entails several distinct but overlapping processes. Ingestion occurs first and then passage of food through the lips and into the mouth by biting or manual placement. This is followed immediately by stage I transport, in which food is propelled from the anterior to the posterior oral cavity. If food particles are still too large or coarse for swallowing, they remain in the mouth. During mastication, food is softened, and food particles are reduced in size by chewing (i.e., incising, crushing, and grinding) and mixing with saliva. Food in the mouth stimulates mechanoreceptors for the trigeminal nerve (cranial nerve V) located in the periodontal membrane and palate. Stimulation of these receptors activates the central pattern generator for mastication, producing sequential contraction and relaxation of the elevator and depressor muscles of the mandible and resulting in cyclic opening and closing of the mouth. This cyclic grinding motion of the jaws is coordinated with rotation of the tongue, pushing the food between the upper and lower teeth. Saliva is excreted from the salivary glands, helping to break down the food and stimulate the taste buds. The physical consistency of the food is monitored continuously by oral mechanoreceptors. Liquids can be held either between the tongue and palate or in the lingual sulcus, or the oral propulsive phase is initiated immediately on ingestion. Soft foods can be held immediately between the tongue and anterior hard palate or lateralized for mastication before resuming a midline position for swallowing.

**Oral Propulsive Phase**
The oral propulsive phase begins once the bolus has been prepared and is ready for transport posteriorly into the pharynx. Once a small portion of solid food is fully prepared (triturated), a cycle of stage II transport is initiated. The tongue pushes upward and forward in the mouth, contacting the anterior portion of the hard palate. The area of tongue-palate contact expands backward, propelling the small portion of the triturated food through the faucial arches and into the pharynx. A small portion of food may remain in the oropharynx while chewing continues for several more jaw cycles. Additional small portions of triturated food may be propelled into the pharynx multiple times, depending on the bolus size (95). When a large enough bolus has been prepared, a swallow is initiated. The pharyngeal phase of swallowing follows immediately. With liquids, the mechanism for transport to the pharynx is as described for solid foods, but the onset of the pharyngeal phase is more rapid after oral propulsion.

**Pharyngeal Phase**
When appropriate sensory input reaches the medullary central pattern generator for swallowing, a complex motor sequence is elicited to propel a bolus through the pharynx, around the larynx, through the upper esophageal sphincter (UES) containing the cricopharyngeus, and into the esophagus, all within a duration of approximately 1 second in healthy adult humans.

When the pharyngeal phase is initiated, respiration ceases and the palatopharyngeal isthmus closes to seal off the nasopharynx, preventing entry of the bolus into the nasal cavity. The larynx closes by adduction of the vocal folds and anterior tilting of the arytenoid cartilages to meet the epiglottic petiole. At nearly the same time, the hyoid bone is elevated and pulled forward by the suprahypohoid muscles. In turn, the hyoid bone and supralaryngeal muscles pull the larynx superiorly and anteriorly. The cricopharyngeus muscle relaxes, becoming compliant and allowing the UES to open (96). Opening of the UES is essential to the pharyngeal phase. Relaxation of the cricopharyngeus is permissive, but the primary force for UES opening is the anterior traction of the hyolaryngeal complex provided by the submental muscles. The tongue moves backward and downward, pushing the bolus through the pharynx like a plunger. The epiglottis inverts, deflecting the bolus around the larynx and away from the airway. The pharyngeal constrictors contract sequentially with a peristaltic wave from top to bottom, clearing the pharynx of residue, following the tongue’s downward motion. Following the passage of the bolus into the esophagus, the larynx opens, the UES closes, and the cricopharyngeus muscle resumes its tonic contraction, sealing the UES (67,97,98).

**Esophageal Phase**
The wave of pharyngeal constriction that cleared the bolus into the esophagus continues throughout the esophagus as a primary peristaltic wave that propels the bolus through the lower esophageal sphincter (LES) and into the stomach. Esophageal clearance is assisted by gravity but also requires relaxation of the LES. Reflux of stomach contents is prevented by tonic contraction of the LES and reflex esophageal swallowing that is triggered by esophageal distension (secondary peristalsis).

**Evaluation of Swallowing**
When a complaint about swallowing is registered or dysphagia is recognized, an evaluation is required with two primary goals in the clinician’s mind: (a) to maintain patient safety during nutritional intake and (b) to determine the best manner for patients to obtain adequate nutrition while maintaining patient safety (99,100). Such an evaluation should consist of a detailed description of the complaint, complete medical history, and a physical examination of the peripheral deglutitory motor and sensory system, including trial swallows under observation. Instrumental diagnostic studies, including videofluorography, manometry, electromyography, and fiberoptic endoscopy, are indicated in selected cases.
Description of the Complaint
In many instances, the subjective complaint gives the examiner clues to the cause of the swallowing problem. Critical data include a sensation of food sticking in the throat or chest, difficulty initiating swallowing, occurrence of coughing or choking spells associated with eating, drooling or difficulty clearing oral secretions, weight loss, change in diet or eating habits, episodes of aspiration pneumonia, and symptoms referable to gastroesophageal reflux. Difficulties swallowing solids and liquids should be contrasted and compared.

With liquids, patients may complain of coughing or choking during drinking. These symptoms are suggestive of laryngeal penetration (entrance of the bolus into the larynx but remaining above the level of the vocal folds) and/or aspiration (passage of the bolus through the vocal folds and into the lower airway) and should be taken seriously. Aspiration is common in patients with swallowing disorders and can lead to serious medical complications. For example, between 8,000 and 12,000 Americans choke to death each year, many with undiagnosed swallowing disorders (101,102). Aspiration of tiny amounts is normal, and the response to aspiration varies widely. It is clear, however, that aspiration in patients with dysphagia significantly increases the risk of pneumonia, so its detection is critically important (103–109).

A complaint of food sticking in the throat or chest is common with solid food and raises the possibility of food pathway obstruction. It may have many causes, however, including bulbar palsy, pharyngoesophageal diverticula, tumor, stricture, or esophageal dysmotility. The sensation of food sticking in the chest (thoracic dysphagia) is usually associated with disease of the esophagus or LES. A complaint of cervical dysphagia (food sticking in the neck), however, has poor localizing value and may be caused by dysfunction of the pharynx, esophagus, or either esophageal sphincter. Nasal regurgitation is associated with weakness or incompetence of the palatopharyngeal mechanism. Oral malodor is associated with tumor, infection, mastication problems, poor oral hygiene, or pharyngeal retention of food and can suggest the presence of a diverticulum of the pharynx or esophagus. Odynophagia is a worrisome symptom, often associated with cancer of the esophagus. Heartburn, acid or sour regurgitation, and regurgitation of digested food suggest gastroesophageal reflux disease. Reflux or vomiting of stomach contents can lead to respiratory complications such as aspiration pneumonitis. Instances of aspiration pneumonia can be indicative of severe dysphagia. The same anatomic and neuromuscular systems are shared by voice, speech, and swallowing, so dysfunction in any of these systems should be carefully noted and described. Weight loss or changes in eating habits often reflect an underlying problem with swallowing as well.

History
Data should be compiled from a review of the patient’s general health and social histories, including what led to the complaint of dysphagia. Attention should be paid to the neurologic history that might suggest contributing factors such as stroke, head trauma, neuromuscular disorders, or degenerative diseases (parkinsonism, multiple sclerosis, dementia) any of which can cause dysphagia. Respiratory disorders and use of artificial airways can also contribute to dysphagia. All prior operations should be noted, especially those involving the head and neck. Current prescription and nonprescription medications should be listed. Those that have side effects of sedation, muscle weakness, drying of mucous membranes, disorientation, or dyskinesia may contribute to dysphagia. Anticholinergic and psychoactive medications are specially noted. Imaging studies are useful for understanding structural changes of the airway and food pathway as well as the brain. Relevant laboratory studies may include important evidence for infection, nutritional deficiency (especially iron deficiency, which may cause esophageal webs), connective tissue disease, or muscle inflammation. Current diet, route of alimentation and hydration, oral health, and disorders affecting behavior or level of consciousness are important to dysphagia rehabilitation.

Psychosocial factors may have a significant impact on swallowing, especially for elderly individuals. An individual living alone may be unable to obtain supervision during meals. For a nursing home resident, prescribing a special diet may be unrealistic. Swallowing must also be considered in the context of feeding. Feeding dependency is an enormous problem for the elderly. Problems with feeding may be difficult to differentiate from impairments of swallowing per se (110).

Clinical Examination and Bedside Swallow Evaluation

Clinical Examination
A general physical examination is essential to look for evidence of cardiopulmonary, gastrointestinal, neurologic, and/or metabolic disease that may impair swallowing. This exam includes an assessment of mental status, including the patient’s ability to follow directions, and a determination of his or her willingness to participate in therapy. Simply talking with a patient can reveal disorders of voice (dysphonia), speech (dysarthria, apraxia), and/or spoken or written language (aphasia), in addition to cognitive deficits in the areas of orientation, memory, reasoning, and visual-motor-perceptual function (111). A quick, informal screening for these functions provides useful information. Cranial nerves should be assessed carefully. The respiratory system is examined for signs of obstruction or restriction such as tachypnea, stridor, use of accessory muscles (such as clavicular breathing), paradoxical motion of the chest wall, or labored breathing. The head and neck are inspected for abnormal tissue (e.g., discoloration, appearance changes in tissue) and palpated for structural lesions. The hyoid bone and laryngeal cartilages are palpated carefully and gently mobilized. Facial sensation is checked bilaterally. The muscles of the face, mouth, and neck are examined beginning with the muscles of facial expression, carefully comparing movement of the two sides of the face for signs of asymmetric weakness. The masseter and temporalis muscles are palpated as the patient bites and chews. Movements of the lower jaw are assessed in three directions of movement, observing and noting any clicking of temporomandibular joint, pain, weakness, and/or asymmetry. The
examination proceeds to inspection of the intraoral mucosa. Careful attention should be paid to the presence of lesions, changes in tissue color, oral debris, abnormal movement, and dryness. Palpation with gloved hand on the floor of the mouth, gum lines, tonsillar fossa, and tongue serves to help rule out neoplastic growth. Tongue strength can be assessed by placing fingers against the outer cheek and resisting the patient’s tongue as it is pushed into the inner cheek. Atrophy, weakness, and fasciculations of the tongue should be noted. The palate is inspected for symmetry at rest and during phonation and gagging. Each side of the pharynx is stimulated to elicit gag reflexes, observing whether the soft palate and pharyngeal walls contract briskly and symmetrically. However, gag reflexes may be difficult to elicit in some normal individuals. The presence of a gag is by no means indicative of a normal swallow, nor is the absence of a gag indicative of dysphagia (112,113). The presence of primitive reflexes associated with chewing and swallowing (such as the sucking, biting, or snout reflexes) should be noted. These pathologic reflexes are often found in patients with damage to both cerebral hemispheres and can indicate impairments of oral motor control.

**Bedside Swallowing Evaluation**

The bedside swallowing evaluation (BSE) is a screening examination that includes observing the patient eating and drinking (99,99a). The purpose of a BSE is to determine whether a patient exhibits signs and symptoms of dysphagia or aspiration and whether the patient is appropriate for instrumental examination or further clinical assessment. The BSE is not a “pass/fail” test; rather, it helps determine the next step in a patient’s evaluation and treatment.

Trial swallows are an essential portion of the evaluation, but they do carry the risk of aspiration. A risk-benefit ratio for the BSE should be considered for each patient. The BSE is the safest when the patient is able to generate a voluntary or reflex cough and is able to follow simple commands. Unstable respiratory status is a relative contraindication. If the determination is made to proceed with a BSE, we recommend using a substance that is relatively safe if aspirated. Clear water and/or small ice chips are relatively innocuous for screening trials at the bedside. Upon presentation of small volumes (1/4 to 1 teaspoon of water/ice chips), the examiner observes and palpates the neck as a gross measure for the promptness of the swallow and adequacy of laryngeal elevation. Clinical signs and symptoms of overt aspiration include coughing, choking, gurgling, wet or hoarse vocal quality, throat clearing, and stridor following the swallow. More subtle signs of aspiration include teary eyes or sniffing. Other clinical observations include drooling, expectoration of food/liquid, a slow rate of eating, residual food in the mouth after swallowing, impulsive eating or drinking, and abnormal posturing of the head and neck with swallowing. The absence of these overt clinical signs or symptoms does not necessarily mean that there is no aspiration. Silent aspiration (i.e., aspiration in the absence of overt signs or symptoms) is common and a high index of suspicion is imperative (108,114–118).

A half teaspoon of crushed ice may elicit chewing because of its texture and temperature. The examiner can observe the chewing action and feel for the laryngeal elevation to indicate that a swallow has occurred. Once it has been determined that the patient adequately elevates the larynx and that there is an adequate protective cough, other substances with varying textures and consistencies can be tried, again considering the risk-benefit ratio mentioned earlier. Soft, solid foods will also elicit chewing and allow the examiner to observe oral coordination and feel for laryngeal elevation. The mouth should be inspected for retention of food after swallowing. The BSE also presents an opportunity to assess the patient’s ability to perform compensatory maneuvers, such as neck flexion or breath holding, that can be later assessed during an instrumental examination.

The purposes of the history and physical examination are to observe actions of the swallowing mechanism, to characterize the nature and severity of the swallowing deficit, to screen the patient’s ability to perform compensatory maneuvers, and to determine whether instrumental examination(s) are necessary. The physical examination is neither sensitive nor specific for identifying aspiration and cannot prove or disprove that a patient aspirates (119). In this regard, any suspicion of difficulty should be immediately referred for further evaluation.

**Diagnostic Studies**

Similar to a BSE, the instrumental examination of swallowing is not a “pass/fail” test. The purpose of instrumental examination is to determine the mechanism of swallowing dysfunction while maintaining a safe method for alimentation and hydration. Additional goals include detection of structural defects, determining which physiologic components (e.g., bolus preparation, tongue control, initiation of the pharyngeal swallow, tongue base retraction, laryngeal closure, UES opening, esophageal clearance) are impaired, detecting the presence of and mechanism for aspiration, and testing of therapeutic and compensatory techniques based on the physiologic impairments that are recognized. Indications for instrumental examination include frequent choking episodes, difficulty managing secretions, wet vocal quality after a swallow, respiratory complications, and unexplained weight loss. Relative contraindications include inability to cooperate with the examination and severe respiratory dysfunction. Two methods are discussed in some detail below—the videofluorographic swallowing study (VFSS) and the fiberoptic endoscopic swallow study.

**Videofluorography**

The VFSS, sometimes referred to as a modified barium swallow study, is considered the gold standard for evaluation of oral or pharyngeal dysphagia. This evaluation is not to be confused with a routine clinical barium swallow or barium esophagram (evaluation of the esophagus). The rehabilitation approach to videofluorography is to have the patient swallow radiopaque liquids and solid food with varying consistencies toward the goals of detecting abnormalities of the swallowing mechanism while establishing a safe and efficient method of
eating and drinking. An empirical approach is used to identify the factors associated with safe and unsafe swallowing such as the consistency of liquids/food, volume of liquids/food, posture of the patient (especially position of the head and neck), and the means for presenting the food. When abnormalities of swallowing (such as aspiration or retention in the pharynx after swallowing) are recognized, compensatory techniques (such as alteration of bolus consistency or posture, or respiratory maneuvers) are attempted and their effectiveness tested empirically (91,120).

Limitations of the VFSS include concerns for radiation exposure and need to modify food consistency with barium. Advantages include viewing the entire swallow from the oral cavity to the esophagus (even to the stomach); viewing aspiration before, during, and after the swallow; and viewing the physiologic effects of rehabilitation techniques in detail. These advantages are unique to VFSS and make it an invaluable tool for assessment of dysphagia. Risks of the VFSS are minimal in comparison to its benefits.

A complete VFSS examination begins with a small amount of thin liquid barium. Both lateral and anterior-posterior views should be obtained with the patient in an upright posture, either standing or seated, depending on posture and physical ability. The oral and pharyngeal stages of swallowing should be studied with the camera focused anteriorly on the lips, superiorly on the hard palate, posteriorly on the cervical spine, and inferiorly on the cervical esophagus. A variety of textures, including thin and thick liquid barium, pudding consistency (barium mixed with pudding), and cookie, are used during the examination. Because concomitant disorders of the pharynx and esophagus are frequent, the examination should include views of the esophagus whenever possible. The esophagus and gastroesophageal junction can be visualized with the patient in an upright, anterior-posterior position, but esophageal function is best assessed in the recumbent position. Many clinicians prefer to separate the evaluation of esophageal function (barium esophagram) from the VFSS, but this requires performance of an additional radiological procedure. Although static x-ray films may be valuable to detect morphologic changes in the pharynx or esophagus, they are not useful in studying the dynamics of swallowing.

**Fiberoptic Endoscopic Examination**

A second instrumental evaluation, fiberoptic endoscopic examination of swallowing (FEES), was first described by Langmore et al. (121,122). This method is a practical means for evaluating swallowing when a patient cannot tolerate transport to radiology (e.g., due to critical illness) or is too large to fit in the fluoroscopy system. FEES carries the added benefit of directly visualizing the pharynx and larynx, to inspect for mucosal lesions or motion impairment of the vocal folds. FEES can detect aspiration in patients for whom radiographic studies are difficult, but it should be noted that aspiration during a swallow cannot be directly visualized. Furthermore, the FEES does not provide views of the esophagus or its sphincters. This limits its applicability for assessing the anatomy and physiology of dysfunctional swallowing. Risks of FEES include patient discomfort, bleeding from passage of the endoscope through the nares, gagging, laryngospasm, and vasovagal response. Despite these limitations, FEES is the instrument of choice for visualization of vocal fold closure, medial and anterior arytenoid movement, amount and location of secretions, and anatomical defects of the pharynx and larynx (123).

**Other Diagnostic Studies**

Esophagoscopy is essential for detecting a variety of esophageal and LES disorders and provides the opportunity for diagnostic biopsy. Electrodagnostic studies may be helpful for detecting motor unit dysfunction of the larynx, pharynx, and oral musculature but cannot substitute for VFSS (124). Manometry of the esophagus is used to identify and differentiate esophageal motor disorders.

**MANAGEMENT OF SWALLOWING IMPAIRMENT**

Once the patient’s swallowing has been described, the impairment(s) identified, and the compensatory strengths recognized, recommendations for management are generated. The recommendations will include diet, route of alimentation and hydration (i.e., oral or nonoral), the need for compensatory strategies (e.g., special postures, respiratory maneuvers), and the need for therapy, including therapeutic exercises. Additionally, the swallowing clinician may make recommendations regarding medication administration (e.g., with liquid, crushing in applesauce or pudding, or via a feeding tube). When oral feeding is recommended, a plan is needed that will maintain optimum alimentation and hydration intake while minimizing risk of choking, aspiration, and malnutrition. The plan must be individualized and based on an understanding of the patient’s swallowing mechanism, the etiology of the dysphagia, and the prognosis for recovery (100). Following recommendations and implementation of the clinical plan, the patient should be monitored for continued signs and symptoms of dysphagia, evidence of pneumonia (fever, cough, and dyspnea), nutritional status (including hydration), and alteration in cognition or mood (125). Table 15-4 lists some of the therapeutic and compensatory techniques that can be employed for patients with oral-pharyngeal forms of dysphagia.

**HEARING**

Hearing loss is the most common type of acquired sensory impairment affecting between 29 and 36 million adults in the United States (126,127). Following stroke and arthritis, hearing loss is the third most prevalent major chronic disability reported by the elderly (128). The prevalence of hearing loss increases substantially with age. Approximately 18% of American adults aged 45 to 64 years report hearing loss.
Therapeutic and Compensatory Techniques for Managing Patients with Dysphagia

This rises to 30% in those aged 65 to 74 years old and 47% for those aged 75 years or older (127). With increasing life expectancy as well as increasing noise exposure, the prevalence of hearing loss is anticipated to become even greater in the upcoming decades. Hearing loss is a potentially disabling disorder that can affect communication with a resultant decline in psychosocial functioning, health-related quality of life, work productivity, and overall quality of life (129). Hearing loss is frequently unrecognized, misunderstood, and all too often neglected, both by those who are affected and by health care providers.

Psychosocial Implications of Hearing Loss

The psychosocial impact of hearing loss is poorly understood and underappreciated by most people. Indeed, even those who live with a person with hearing impairment rarely grasp the full extent to which this all-pervasive, invisible disability affects daily living. Until hearing loss is experienced directly or unless a very close acquaintance has impaired hearing, the ways in which a person depends on hearing remain largely unrecognized. Most aspects of daily life are affected by hearing loss in some way. Hearing loss impairs communication, often subtly at first, and increasingly so as the magnitude of the hearing loss increases. In addition, hearing loss restricts environmental awareness and can be a safety concern when it affects the ability to hear warning signals such as sirens and alarms. Hearing loss impacts relationships, employment opportunities, academics, and learning.

The two most commonly reported consequences of hearing loss are depression and social isolation. In addition, adverse effects on general well-being and on physical, cognitive, emotional, behavioral, and social functions have been reported. Social and emotional handicaps are present even in those with only mild to moderate hearing loss. Inability to hear all or part of a conversation may create frustration, anger, and even paranoia. Over time, social connectivity and relationships may deteriorate, leaving the individual with hearing impairment in isolation and with a diminished quality of life. Hearing loss is associated with reduced abilities to perform activities of daily living as well as instrumental activities of daily living (129).

Misunderstanding and lack of sympathy for the hearing-impaired person seems to be built into our social bias. These attitudes are certainly quite different from our perceptions and treatment of blindness. Often the symptoms of hearing loss (e.g., not answering when spoken to, answering inappropriately, or requiring repetition) encourage other people to talk more and treat the person with hearing impairment as if his or her cognitive abilities were also diminished. This often leads to frustrating or truncated communication attempts.

A major goal of the health care practitioner is to help patients maintain or regain function. The capacity for independent living requires maintenance of functional health. Functional health refers not only to physical health but also to emotional, cognitive, and social health. Physicians are in an excellent position to identify treatable conditions that may compromise their patients’ functional performance. Unfortunately, many health care professionals tend to view hearing loss as a benign problem that does not threaten functional health. Regular hearing assessments and management for persons aged 60 years and older should be considered a standard of care.

Anatomy and Physiology of the Ear

Anatomically, the ear is a complex organ, capable of transforming airborne sound waves into mechanical energy. This mechanical energy is then converted into electrochemical signals and then into neural impulses that are processed as auditory information. The peripheral portion of the auditory system is divided into three sections: the outer, middle, and inner ear. The outer ear is comprised of the auricle (pinna) and the external auditory canal. The auricle is a skin-covered cartilaginous structure and is the most visible portion of the outer ear. The

---

**TABLE 15.4 Therapeutic and Compensatory Techniques for Managing Patients with Dysphagia**

<table>
<thead>
<tr>
<th>Technique</th>
<th>Physiologic Swallowing Impairment</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effortful/hard swallow</td>
<td>Reduced tongue base retraction, pharyngeal weakness</td>
<td>Improve bolus flow through the pharynx by improving tongue base-to-pharyngeal wall contact</td>
</tr>
<tr>
<td>Chin down/chin tuck</td>
<td>Delayed initiation of the pharyngeal swallow</td>
<td>Prevent advancement of the bolus beyond the valleculae</td>
</tr>
<tr>
<td>Head turn</td>
<td>Tongue base retraction, Pharyngeal weakness</td>
<td>Reduce distance of tongue base-to-pharyngeal wall contact</td>
</tr>
<tr>
<td>Head tilt</td>
<td>Oral weakness, reduced sensation</td>
<td>Head is tilted toward the stronger/more sensate side to improve oral control of the bolus</td>
</tr>
<tr>
<td>Supraglottic swallow</td>
<td>Impaired/delayed vocal fold closure</td>
<td>Improve vocal fold closure to prevent aspiration</td>
</tr>
<tr>
<td>Super-supraglottic swallow</td>
<td>Impaired/reduced vocal fold and laryngeal closure</td>
<td>Improve vocal fold closure and arytenoids-to-epiglottic petiole contact to improve airway closure during the swallow</td>
</tr>
<tr>
<td>Thickened liquids</td>
<td>Bolus control</td>
<td>Improve bolus control by slowing the bolus</td>
</tr>
<tr>
<td>Thin liquids</td>
<td>Tight UES</td>
<td>Reduce friction, improve bolus passage to esophagus</td>
</tr>
</tbody>
</table>

---
When a patient presents with a hearing complaint, diagnosis of the type and degree of hearing loss and the underlying etiology is typically the first goal in the assessment and management process. This begins with a detailed hearing history that includes questions about the duration of the hearing loss, whether the onset was sudden or gradual, and whether it is stable, progressive, or fluctuant. The patient should be queried about associated symptoms such as tinnitus, dizziness, ear pain or drainage, and aural fullness. In addition, questioning should review family history of hearing loss, as well as loud noise exposure, head trauma, ear surgery, and ototoxic drug use. The physical examination should include an otoscopic evaluation to ensure that the external ear canal is healthy and unobstructed and that the tympanic membrane is translucent and intact. The status of the patient's hearing is determined through a series of audiologic tests. Based on the results of the audiologic examination, the patient's history, and physical examination findings, additional audiologic assessments as well as associated neurologic, laboratory, and imaging studies may be indicated. Only after the assessment has been completed and a diagnosis has been established should rehabilitation of the hearing loss be initiated.

**Behavioral Hearing Assessment**

**Pure-Tone Audiometry**

Determination of pure-tone thresholds across a wide frequency range for air conduction and bone conduction stimuli is the most basic component of the hearing test. Threshold is defined as the lowest level at which a person can detect an auditory stimulus 50% of the time. This information is used to establish the type, degree, and configuration of the hearing loss. Air conduction thresholds are determined in octave, and sometimes interoctave, intervals over the frequency range of 250 through 8,000 Hz. Air conduction signals are presented via supra-aural or insert earphones, and the signal is transferred along the entire auditory pathway, including the outer ear, middle ear, and inner ear. Bone conduction thresholds are established for octave intervals between 250 and 4,000 Hz. For bone conduction measures, a bone oscillator is typically placed on the mastoid prominence of the test ear, while masking sounds are delivered to the nontest ear. The bone conduction oscillator causes the skull to vibrate resulting in direct stimulation of the cochlea, with minimal contribution of the outer and middle ears. Bone conduction thresholds reflect an individual's sensory hearing ability and are unencumbered by disorders affecting the outer and/or middle ear.

Pure-tone thresholds are represented graphically on an audiogram (Fig. 15-1) or in a tabular form. The frequency scale along the abscissa is measured in Hertz (Hz) for the octave and interoctave frequencies of 250 through 8,000 Hz. The intensity scale on the ordinate of the audiogram displays the hearing level (HL) of the signal in decibels (dB) ranging from a very faint level of −10 dB HL up to a very loud level of 110 to 120 dB HL.

**Speech Audiometry**

In addition to pure-tone thresholds, the basic audiologic evaluation includes measurement of the speech recognition threshold (SRT) and the assessment of word recognition performance.
The SRT serves primarily as a reliability check of the pure-tone threshold levels. Familiar two-syllable spondee words (e.g., hotdog, cowboy, baseball) are presented to each ear, and the patient is asked to repeat back these words until they can no longer understand or hear them. The intensity level at which 50% of the words are correctly repeated is defined as the SRT. The SRT should be within 10 dB of the pure-tone average (PTA) of 500, 1,000, and 2,000 Hz or the average of the best two of these frequencies. The SRT provides a means to cross-check the validity of the pure-tone thresholds. When there is a difference between the SRT and the PTA of 10 dB or greater, the source of this discrepancy should be determined. Causes include functional hearing loss, unusual audiometric configurations, language or cognitive disorders, patient misunderstanding of test instructions, and audiometer calibration or malfunction.

A hearing impairment may be reflected not only in a sensitivity loss but also in the reduced ability to understand speech, even when speech is sufficiently loud. Word recognition testing (speech discrimination) analyzes the patient’s ability to understand speech using standardized lists of 25 to 50 single-syllable words. These lists are balanced to represent the phonetic content of everyday English speech. Word lists are presented to each ear separately and the patient is asked to repeat back the test words. The percentage of correct responses is the word recognition score. The test is initially presented at a comfortably loud level. Higher levels may be presented in an effort to determine the patient’s best word recognition performance.

**Physiologic Hearing Assessment**

**Immittance Measurements**

Physiologic assessment of the middle ear function is a routine part of the basic audiologic evaluation. Collectively referred to as immittance measurements, middle ear assessment includes tympanometry and acoustic reflex testing. Tympanometry provides information about the admittance, or compliance, of the middle ear transmission system and an estimate of middle ear pressure. This information is particularly useful in identifying and differentiating among middle ear disorders. The acoustic reflex occurs when a sufficiently intense sound is presented to the ear causing contraction of the stapedius muscle and stiffening of the ossicular chain. Evaluating the integrity of the acoustic reflex provides information regarding not only the middle ear function but also other components of the reflex arc including the cochlea, seventh and eighth cranial nerves, and the lower brainstem.

**Otoacoustic Emissions**

A more recent addition to the audiologists’ armamentarium, evoked otoacoustic emissions (OAEs) are low level sounds that emanate from the cochlea in response to sound stimulation. They are generated by healthy cochlear outer hair cells and impart insight into the functional integrity of the cochlea. In general, OAEs are observed in ears with normal to near-normal hearing sensitivity and can be greatly impacted by the presence of outer, middle, or inner ear disorders. OAEs are used in universal newborn hearing screening programs and in other cases in which patients are unable or unwilling to provide behavioral responses to sounds. They also provide confirmation of behavioral test results, serve as an objective method of monitoring ototoxicity, and assist in the identification of both functional hearing loss and neurologic disorders of the auditory system such as auditory neuropathy. While OAEs provide adjunctive information in hearing evaluation, they do not replace the audiogram.

**Auditory-Evoked Potentials**

When the ear is stimulated with sounds, a series of bioelectric events occur, that begin in the cochlea and quickly reach cortical structures. Much like electroencephalography (EEG), auditory-evoked potentials can be recorded from surface electrodes placed on the scalp and in or near the ear canal. The auditory responses are extracted from ongoing EEG activity by time-locked computerized averaging during repeated stimulation with transient or tonal stimuli. The evoked potentials most commonly used for audiologic assessment include electrocochleography (ECochG), the auditory brainstem response (ABR),
auditory steady-state response (ASSR), and auditory middle latency response (AMLR). The primary clinical application of ECochG, which measures electrical potentials generated in the cochlea, is identification and monitoring of Meniere’s disease. The ABR is used in newborn hearing screening programs to estimate peripheral hearing sensitivity in pediatric and uncooperative patients, in otoneurologic diagnosis, and for intraoperative monitoring during cerebellopontine angle surgery in which the auditory system is at risk for iatrogenic insult. The ASSR is used to estimate the degree, type, and configuration of peripheral hearing loss and appears to have a good predictive capacity in relation to the behavioral audiogram, especially in cases of severe and profound hearing loss. Less frequently used, the AMLR is thought to be generated in the auditory thalamocortical pathway. It can be used to predict auditory thresholds and may provide insight into neurologic diseases involving primary auditory pathways and auditory processing disorders.

**Audiometric Test Interpretation**

**Type of Hearing Loss**

The general anatomic location of a hearing impairment can be determined by comparing the air conduction and bone conduction thresholds for each ear individually. A conductive hearing loss is present when air conduction results demonstrate a hearing loss but bone conduction results are within the normal range (Fig. 15-2). The difference between air and bone conduction thresholds reflects the amount of conductive involvement and is called the *air-bone gap*. A conductive hearing loss may be caused by any disorder or dysfunction of the sound-conducting mechanisms from the external auditory canal through the middle ear. Although otoscopic examination may provide evidence of cerumen impaction, tympanic membrane perforation, or serous otitis media as the cause of a conductive hearing loss, there are also conductive pathologies that present with normal otoscopic examinations, such as otosclerosis or ossicular discontinuity. Tympanometry can provide useful information by describing the status of the middle ear system. Patients with conductive hearing loss typically have normal word recognition scores, because the sensorineural system is intact. Speech needs only to be presented at louder levels than normal to compensate for the conductive deficit.

When an equal amount of hearing loss is present by air conduction and bone conduction, the hearing loss is called a sensorineural loss (Fig. 15-3). The hearing disorder could be located in the cochlea, the associated neural pathways, or both. The specific etiology of the sensorineural hearing loss cannot be determined by the audiometric results alone. Word recognition test results often provide important diagnostic
insight into the site of lesion of sensorineural hearing loss. In general, cochlear involvement demonstrates speech discrimination scores that are reduced to a degree compatible with the degree of hearing loss. The greater the degree of sensory (cochlear) hearing loss, the poorer the word recognition. On the other hand, neural auditory disorders often yield speech discrimination scores disproportionately poorer than would be expected from the pure-tone thresholds. That is, a 40-dB HL sensorineural hearing loss with a 72% word recognition score would be consistent with cochlear involvement, whereas a similar amount of hearing loss with only a 10% word recognition score would suggest the possibility of eighth cranial nerve involvement. From a rehabilitative standpoint, the better the speech discrimination score, the better the prognosis for hearing aid success, as there is less distortion in the auditory system. Tympanometry is most often normal in cases with sensorineural hearing loss, while acoustic reflex patterns may help to differentiate sensory from neural causes.

A loss of hearing sensitivity for bone conduction with a greater loss for air conduction represents a mixed hearing loss (Fig. 15-4). A sensorineural hearing loss is present, as reflected by reduced bone conduction thresholds, and conductive loss also is present, as reflected by the air-bone gaps. Tympanograms may provide insight into middle ear problems contributing to the conductive component of a mixed hearing loss.

In addition to hearing loss resulting from damage or dysfunction of the peripheral hearing mechanisms, damage or dysfunction of the central auditory pathways can impact hearing abilities. Central auditory processing disorders most often are manifested as difficulty understanding speech in noise or situations in which the signal is degraded, such as a reverberant room. Auditory test results will frequently show normal hearing sensitivity and normal speech recognition ability in quiet. Performance declines are observed on degraded or difficult speech recognition tests including speech in noise and filtered speech, measures of frequency and temporal processing, dichotic listening, and/or on electrophysiologic measures of the central auditory pathways such as the ABR and AMLR. Common causes include head trauma, stroke, neurodevelopment disorders, genetic factors, tumors, neurologic disease, aging, and long-term sound deprivation (130). While less amenable to traditional hearing aid amplification, persons with auditory processing disorders may benefit from other types of assistive listening systems, as well as therapeutic auditory regimens.

Degree and Configuration of Hearing Loss

The results obtained from the air conduction evaluation provide quantitative information as to the magnitude of hearing loss. Classification systems have been devised in an effort to relate the amount of air conduction hearing loss to the expected degree of handicap imposed by a hearing loss. Such systems typically use the PTA to estimate various hearing loss categories and the expected impact on speech understanding. The most commonly used PTA is a three-frequency average of 500, 1,000, and 2,000 Hz, although a four-frequency PTA that also includes 4,000 Hz may be more reflective of functional hearing. Table 15-5 gives an example of such a classification system (131). In addition to type and degree of hearing loss, the configuration or shape of the pure-tone audiogram is helpful in describing and understanding the impact of a hearing loss. Classification schemes include flat, sloping, and rising configurations. A sloping audiogram has a greater amount of hearing loss in the high frequencies, whereas a rising configuration refers to a greater amount of hearing loss in the low frequencies.

---

**FIGURE 15-4.** This audiogram shows that the patient has a mixed hearing loss. Although both the air conduction and the bone conduction thresholds are reduced, a greater impairment is evident for air conduction.

---

**TABLE 15.5 Classification System for Degree of Hearing Loss**

<table>
<thead>
<tr>
<th>Pure Tone Average</th>
<th>Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–15 dB HL</td>
<td>Normal</td>
</tr>
<tr>
<td>16–25 dB HL</td>
<td>Slight</td>
</tr>
<tr>
<td>26–40 dB HL</td>
<td>Mild</td>
</tr>
<tr>
<td>41–55 dB HL</td>
<td>Moderate</td>
</tr>
<tr>
<td>56–70 dB HL</td>
<td>Moderately severe</td>
</tr>
<tr>
<td>71–90 dB HL</td>
<td>Severe</td>
</tr>
<tr>
<td>91 dB HL and above</td>
<td>Profound</td>
</tr>
</tbody>
</table>
Any type of audiometric classification system must be interpreted with caution because most are based on pure-tone air conduction thresholds alone and do not incorporate the effects of speech discrimination difficulties, etiologic factors, the type of hearing loss, or hearing loss configuration. In addition, those with similar amounts of pure-tone hearing loss may be affected in very different ways, depending on lifestyles, hearing demands, and other psychosocial factors.

**Causes of Hearing Loss**

Hearing loss can be caused by or occur in conjunction with a variety of diseases and disorders, genetics, aging, trauma, or ototoxicity. The following description provides an overview of the major causes of hearing loss including conditions affecting the outer, middle, and inner ear as well as the central auditory pathways.

**Hereditary Hearing Loss**

Hereditary hearing loss accounts for approximately 50% of hearing loss identified in newborns. Approximately 70% of genetic hearing loss is nonsyndromic, and approximately 30% is syndromic. Nonsyndromic hearing loss is most often sensorineural. It is typically classified by its mode of transmission, which is recessive in approximately 70%, dominant in approximately 30%, x-linked in approximately 1% and mitochondrial in approximately 1%. Nonsyndromic recessive hearing loss is most often congenital and profound, whereas nonsyndromic dominant hearing loss is more variable in presentation with congenital or delayed onset. There are at least 400 syndromes with hearing loss as a phenotypic feature. Syndromic hearing loss is quite variable in its presentation. It can be conductive, mixed, and sensorineural; it can range from slight to profound; it can be stable or progressive; and it can involve one or both ears (132). The Hereditary Hearing Loss Homepage (133) provides a comprehensive review of the genetics of hereditary hearing loss, which is beyond the scope of this chapter.

**Outer Ear**

When hearing loss originates in the external auditory canal, it usually is related to a mechanical obstruction in the form of cerumen (ear wax) or a foreign body. Accumulation and impaction of excess cerumen in the external ear canal is one of the most common reasons that patients seek care for hearing- and ear-related symptoms. Elderly, pediatric, and cognitively impaired individuals are at higher risk for cerumen impaction. This is especially common in elderly residents of nursing homes where the prevalence of cerumen impaction may be as high as 65% (134). The resultant conductive hearing loss ranges from slight to mild and is most often resolved after removal of the impacted cerumen from the ear canal. In addition to hearing loss, excessive ear wax may prevent benefit from or cause damage to hearing aids.

Infections of the ear canal (otitis externa) can occur in acute, chronic, and malignant forms. They are most commonly caused by bacterial infection and more rarely may have a viral or fungal origin. Signs and symptoms include ear pain, drainage, edema, inflammation, itching, and aural fullness. If the infectious process causes the ear canal to become swollen and shut, there is a resultant conductive hearing loss. Medical management includes treatment of pain and control of infection. Hearing aid use may need to be discontinued during periods of active infection (135).

**Middle Ear**

Otitis media is an infection, or inflammation, of the middle ear usually with fluid (effusion) behind an intact tympanic membrane. Acute otitis media (AOM) occurs frequently in children and is the most common infection for which antibacterial agents are prescribed in the United States. When treated successfully with medications or middle ear surgery, AOM is alleviated or controlled, but when left untreated, complications of the disease may spread to involve intracranial or intratemporal structures. Otitis media with effusion (OME) indicates a situation where the middle ear is filled with fluid but typically without signs or symptoms of acute ear infection other than hearing loss. The audiometric presentation of a person with otitis media ranges from a slight to moderate conductive hearing loss, typically with a rising configuration. The tympanogram is flat indicating absence of middle ear system mobility, and the stapedial reflex is absent.

In addition to hearing loss, complications of otitis media include cholesteatoma, an inclusion of keratinizing squamous epithelium that proliferates within the temporal bone. A cholesteatoma can block sound transmission or cause erosion of the middle ear ossicles, resulting in a conductive hearing loss. When a cholesteatoma erodes the temporal bone, it puts the patient at risk for mixed or sensorineural hearing loss, vertigo, and/or meningitis. Mastoid surgery is the means by which cholesteatoma is eradicated. The primary surgical goal is to rid the ear of the cholesteatoma itself. Secondarily, reconstruction of the tympanic membrane and/or ossicles may be necessary to preserve or improve hearing. In many cases, hearing aids will be required following the eradication of the cholesteatoma (137).

Glomus tumors are the most common neoplasm of the middle ear and the second most common tumor of the temporal bone in adults (138). The most frequent initial symptoms are pulsatile tinnitus and hearing loss. On otoscopic examination, a glomus tympanicum appears as a red-blue mass behind or involving the tympanic membrane. Conductive hearing loss is the most common audiometric finding occurring in up to 52% of people with these tumors as the result of reduced ossicular mobility or destruction. As the tumor grows medially, there is cochlear involvement resulting in a mixed or sensorineural hearing loss. Surgical excision is the treatment of choice unless contraindicated by the patient’s medical condition or tumor location and biology (138).

Otosclerosis is a progressive, focal disease of the ear in which there is excessive resorption of bone, which is replaced by soft new bone that gradually changes into a dense sclerotic mass. In its early stages, otosclerosis hinders transmission of sound by the ossicular chain due to fixation of the stapes footplate in the oval
There are a number of drugs that may be toxic to the inner ear and may cause hearing loss, tinnitus, balance dysfunction, or a combination of these symptoms. The most common ototoxic drugs include aminoglycoside antibiotics, cisplatin, high-dose salicylates, and loop diuretics. The typical ototoxic hearing loss is bilateral and sensorineural. It begins in the high frequencies and, with continued administration of the ototoxic agent, spreads to include lower frequencies while high-frequency hearing continues to decline. Eventually, the patient may experience significant communication difficulties necessitating use of amplification. In general, aminoglycosides cause a gradual onset of hearing loss, while cisplatin can cause a significant hearing loss after a single treatment. Aminoglycosides and cisplatin cause permanent hearing loss, while the hearing loss resulting from high-dose salicylates and loop diuretics is often reversible.

**Inner Ear**

**Presbycusis**

Presbycusis is the term used to describe age-related hearing loss. The most common audiologic presentation of age-related hearing loss is a sloping, high-frequency sensorineural hearing loss. Hearing thresholds decline gradually with age, beginning as early as the third decade in men and the fifth decade in women. Concomitant to the loss of sensitivity is a reduction in the ability to understand speech, especially in background noise and reverberant rooms or when the speaker talks at a rapid rate or with accented speech. Reduced cognitive abilities can further compound these problems. Age-related changes in the ear include deterioration of structures within the inner ear, including the stria vascularis, cochlear hair cells, and dendritic fibers of the eighth nerve. In most cases, age-related changes occur within the milieu of other potential causes of hearing loss, including genetic susceptibility, noise exposure, ototoxic drugs, and ear disease. Most persons with presbycusis are able to obtain benefit from hearing aids used alone or in conjunction with other amplification devices.

**Noise-Induced Hearing Loss**

Exposure to high levels of sound can cause damage to the hair cells in the cochlea resulting in noise-induced hearing loss (NIHL). Approximately 26 million Americans aged 20 to 69 have hearing loss attributable to noise exposure (130). More recently, NIHL in pediatric populations has become a concern. Harmful sounds include both impulse and continuous noises. Hearing loss can occur suddenly after a single noise exposure but more commonly occurs gradually as the result of the accumulative effect of years of noise exposure. The hearing loss begins in the high frequencies with a characteristic notch around 4,000 Hz. Over time with repeated exposure, NIHL spreads to the mid and low frequencies with deleterious effects on the ability to hear and understand speech, resulting in the need for amplification. Categories of activities associated with noise exposure include military (e.g., weapons, explosives), occupational (e.g., loud machinery, heavy equipment), and recreational (e.g., hunting, loud music, power tools). While personal hearing protection devices such as ear plugs and ear muffs provide some reduction in sound levels, use of these among individuals exposed to harmful noise levels is not always consistent.

**Neoplasms**

The most common tumor of the temporal bone is the acoustic neuroma (vestibular schwannoma) that accounts for 80% of
intracranial tumors in the cerebellopontine angle. Vestibular schwannomas make up almost 10% of all intracranial tumors and most commonly arise from the vestibular portion of the eighth cranial nerve (127). Patients typically develop unilateral sensorineural hearing loss with poor speech discrimination, tinnitus, and sometimes vertigo. Treatment options include surgical resection to remove the tumor or highly focused radiation therapy. Although hearing preservation is possible in some patients, the primary goal is the removal of the neoplasm. Post-treatment benefit from a hearing aid may be possible in some cases.

Medical-Surgical Rehabilitation
Medical or surgical treatment of hearing loss is most often available for people with impairments involving the conductive mechanisms of the auditory system. Hearing impairment originating in the middle ear system may be treated with otologic surgery. Surgical procedures such as myringoplasty (i.e., repair of tympanic membrane perforation), tympanoplasty (i.e., ossicular reconstruction), stapedectomy for otosclerosis, and myringotomy with placement of ventilating tubes for middle ear effusion often can correct the conductive hearing loss.

Otolaryngic surgery sometimes is required for treatment of life-threatening disease and not for hearing improvement. Pathologic conditions such as cholesteatoma, glomus tumor, or chronic middle ear disease often necessitate surgery. In addition, otoneurosurgery is necessary for sensorineural impairment caused by acoustic neuromas. Otolaryngic treatment of congenital or hereditary sensorineural hearing loss, NIHL, presbycusis, and most other types of sensorineural impairments is not possible at this time. Perhaps the greatest advance in the surgical rehabilitation of sensorineural impairments is now available with the development of the cochlear implant for severe to profound hearing loss.

Hearing Aid Amplification
Although hearing loss due to conductive disorders may often be successfully alleviated with medical or surgical intervention, the most common site of lesion that causes hearing loss is the cochlea. Of all patients with cochlear hearing loss, fewer than 5% can be helped medically. Hearing aids are the principal resource for improving communication and reducing hearing handicaps in persons with sensorineural hearing loss. Significant improvements in hearing aid technology have resulted in greater flexibility in selecting and fitting hearing aids for most hearing losses. The past decade has witnessed innovations resulting in a wide assortment of hearing aids, varying both in physical size and in technological sophistication. Yet, despite the substantial benefits provided by hearing aids, only 23.5% of people with hearing loss own hearing instruments (132).

Hearing Aid Candidacy
It is a disservice to the patient to discourage a hearing aid trial because of the misconception that a hearing aid does not help "nerve deafness." While some physicians may tell their patients that they can probably "get by" without a hearing aid and should wait until the hearing loss progresses, consultation with an audiologist will provide the patient with an individualized opportunity to make amplification decisions that best suit their lifestyle and listening needs. Discouraging a hearing aid trial for an individual with communicative difficulties and a potentially remediable hearing loss serves only to invite isolation and frustration. Unless assurance and support are provided, hearing-impaired patients may unfortunately postpone and avoid the use of amplification. Referral to an audiologist for hearing assessment and hearing aid evaluation provides patients an opportunity to understand their hearing loss and its impact on daily function, as well as candidacy for the many types of amplification systems suitable for their individual needs.

Hearing Aid Selection and Fitting
Current hearing aid styles include devices that fit behind-the-ear, in-the-ear, in-the-canal, and completely-in-the-canal. Because many people unfortunately associate hearing loss with the stigma of aging, the introduction of smaller devices, which fit entirely in the ear canal, has had obvious cosmetic appeal. Newer and smaller behind-the-ear hearing aids that feature a slim tube and an open-fitting have resulted in a recent swing of the pendulum back to behind-the-ear hearing aids. This type of fitting helps to eliminate the sensation of an occluded ear and has proven to be advantageous for those with mild to moderate hearing loss as well as hearing loss confined to the high frequencies. Selection of appropriate hearing aids must take into account degree and type of hearing loss, loudness tolerance, communication needs, manual dexterity, vision impairments, and personal preference.

Hearing aids no longer merely provide analog, linear amplification. The vast majority of current hearing aids are digital, and include many features that enhance the amplified signal and tailor the sound to the specific needs of the individual hearing aid user. Compression circuitry reduces the amplification of loud sounds to comfortable levels without removing valuable informational content. In addition, automatic loudness-adjusting circuits automatically decrease amplification of continuous background sounds in an effort to improve speech understanding in noise. Hearing aids without volume controls automatically increase the loudness of soft sounds while decreasing loud level inputs. Frequency compression, a more recent hearing aid feature, modifies the bandwidth of the amplified signal so that it is within the patient’s frequency range of audibility. There are also hearing aids that offer several programs for use at the listener’s discretion, because a single amplification paradigm may not be optimal for all listening situations. One may choose a wide frequency amplification for quiet environments, another program that eliminates low frequencies for noisier situations, or a program specifically suited for telephone use or listening to music. Most hearing aids are digitally programmed by a computer through the use of a microchip within the hearing aid itself. Individual programming of
the hearing aid enables the audiologist to make significant modifications to meet the patient's needs. Multiprogram hearing aids may include a user-operated remote control to facilitate program changes or volume changes. Fixed or adaptive directional microphones help to improve speech understanding by improving the signal-to-noise ratio. Bluetooth technology enabling wireless communication between the hearing-aid and other electronic devices such as cell phones and MP3 players helps to keep the technology-savvy hearing aid user connected.

After the initial history, examination, and evaluation, the audiologist will determine the need for and recommend the type of hearing aids, specify the acoustical and technological requirements of the hearing aids, program and fit the hearing aids, and provide training in the use of amplification. The potential advantages and limitations of hearing aids are reviewed with the patient, and follow-up is provided during the adjustment period after the hearing-aid fitting. Most states mandate a rights-to-return period, which provides patients with the opportunity to return the hearing aids in their own everyday listening environments to evaluate benefit, comfort, and overall satisfaction with the instruments. If they are unsatisfied for any reason, they can return the hearing aids at little or no cost during this period of time. The audiologist assists the patient during this transition and may make further adjustments or changes to the devices to ensure that the hearing-aid fitting is optimal. In addition to fitting and programming the hearing aids, the audiologist provides counseling in effective communication and auditory training to enhance and maximize listening skills.

**Assistive Listening Devices**

Although substantial improvements have been achieved in hearing-aid technology, factors such as distance from a speaker, background noise, reverberation, and declines in central auditory processing have deleterious effects on speech intelligibility that cannot be overcome through the use of hearing aids alone. Assistive listening devices (ALDs) comprise a number of situation-specific amplification systems designed for use in difficult listening environments. ALDs commonly use a microphone placed close to the desired sound source (e.g., a television, theater stage, or speaker's lectern from which sound is directly transmitted to the listener). Transmission methods include infrared light, FM radio waves, and induction loops. Such transmission of sound directly to the listener improves the signal-to-noise ratio. That is, the desired sounds are enhanced while competing extraneous noises are decreased, thus improving understanding. Public Law 101-336, the Americans with Disabilities Act (1990), requires assistive listening technology in places of public accommodation. ALDs in many churches, theaters, and classrooms enable hard-of-hearing persons to avoid the isolation imposed by the inability to hear a sermon, play, or public address.

Amplified telephones, low-frequency doorbells, amplified ringers, and closed-captioned TV decoders are just a few examples of the number of devices currently available for the hearing impaired for everyday use. Alarms are available with low-frequency signals for those persons with high-frequency hearing losses who cannot hear the higher pitch alarms. Flashing alarm clocks, alarm bed vibrators, and flashing smoke detectors are other alerting options for severely hearing-impaired individuals.

**Cochlear Implants**

The cochlear implant is a surgically implanted auditory prosthesis designed to provide acoustic stimulation for persons with bilateral severe to profound sensorineural hearing loss who derive little or no benefit from hearing aids. The device bypasses the cochlear hair cells and provides direct electrical stimulation of the residual eighth nerve fibers in the cochlea. The FDA approved cochlear implants for adults in 1984 and for children as young as 2 years in 1990. In 2000, the age of eligibility was lowered to 12 months. As of 2009, more than 41,500 adults and 25,500 children in the United States have received cochlear implants (127). Candidacy for a cochlear implant is evaluated by a team, including an otolaryngologist and audiologist. Prospective implant candidates require an extensive audiologic evaluation to document that hearing aids are not beneficial and an otologic evaluation to determine implant and surgical candidacy. The implant operation involves placing an electrode array into the cochlea and connecting it to an internal coil, which is placed under the skin behind the pinna and is aligned with an external coil placed behind the ear. After the healing period, the patient is fitted with an external microphone, speech processor, and transmitter. The microphone, usually worn at ear level, feeds electrical impulses to the speech processor that resembles a body-type or behind-the-ear hearing aid. The processor digitally encodes the sound and sends these electrical signals to the external coil, which transmits the signal to the internal coil and then to the electrodes in the cochlea. Current flows between the active electrodes, stimulating remaining eighth nerve fibers and producing a sensation of sound. The audiologist programs the speech processor to identify and establish a sound map that works best for the individual patient. Rehabilitation includes repeated visits to program and fine-tune the sound map and intensive auditory training. Patient commitment to the full rehabilitative process is vital to success with the implant.

While a cochlear implant does not provide normal hearing, users may derive many benefits including recognition of environmental sounds and ability to understand spoken language, and in some cases without, speech reading cues. Benefits for children implanted early in life may include enhanced development of speech and language skills as well as improvement in overall literacy. Current evidence suggests that binaural implantation and bimodal fittings (combining a cochlear implant and hearing aid in opposite ears) may become the standard of care for pediatric cases.

**Speech Reading and Auditory Training**

Hearing aids and other amplification systems provide the hearing-impaired individual with access to auditory information, but this alone does not guarantee comprehension. Supplementary information obtained through speech reading (lip reading) and auditory training to improve listening skills is an important component of successful hearing rehabilitation.
Speech reading is the use of visual cues in the recognition of speech and incorporates the interpretation of facial expressions, body movements, and gestures. Everyone uses speech reading to some extent, although usually we are not conscious of the importance of visual input in helping us to recognize what is being said. Many people with hearing impairment, particularly those with gradually progressive hearing loss, develop this skill through necessity. Although a considerable amount of the speech signal can be perceived visually, only about one third of English speech sounds are clearly visible. Certain sounds (e.g., /f/ and /θ/) are relatively easy to see on the lips, whereas others, such as /k/ and /g/, are not visible, and some (e.g., /p/ and /b/) are indistinguishable from one another.

Auditory training teaches the hearing-impaired individual to make the most effective use of the limited auditory cues imposed by the hearing loss as well as the additional auditory information delivered by the hearing aid. While the benefits of auditory training programs have been documented, such programs have not been time- or cost-effective. Recently, computer-based, adaptive, auditory training programs have become available (e.g., Listening and Communication Enhancement) that provide a home-based, individualized approach to aural rehabilitation.

Aural rehabilitation strategies also teach the hearing-impaired person to become a more assertive listener. Those who quietly accept hearing loss or not understanding speech merely invite continued social isolation. The hearing-impaired listener needs to inform others of his or her hearing impairment and advise them as to the most effective means of communication. Self-help groups are available, most notably the Hearing Loss Association of America, which offers local groups, as well as an active national organization and journals.

Although management of hearing loss with hearing aid amplification, ALDs, cochlear implants, and aural rehabilitation therapy does not cure the impairment or restore hearing and communicative efficiency to normal, such approaches represent the best treatments available at this time. They will improve the ability of most people to communicate effectively and reduce the handicapping consequences of hearing loss.

**ACKNOWLEDGMENTS**

The authors wish to acknowledge the contributions of Robert M. Miller, Michael E. Groher, Kathryn M. Yorkston, Thomas S. Rees, and Jeffrey B. Palmer (authors of Speech, Language, Swallowing, and Auditory Rehabilitation of the previous edition). Much of their original published work has remained in this updated chapter.

**REFERENCES**


443

CHAPTER 15 | SPEECH, LANGUAGE, SWALLOWING, AND AUDITORY REHABILITATION


121. Van Camp G, Smith R. Available at: http://webh01.ua.ac.be/hhh/.


137. Ching T, Massie R, Wanstrow E, et al. Bimodal fitting or bilateral implanta

Sexuality and Disability

INTRODUCTION

Sexuality is an integral part of being human and is a vehicle to demonstrate attraction, intimacy, and commitment. Because of this, sexuality persists beyond reproductive years and/or good health. Persons with disabilities often feel that since significant changes to their sexuality are not life threatening, sexual concerns do not merit attention by health care professionals. Nothing can be further from the truth. Sexuality is highly important and medically legitimate and greatly affects quality of life. When surveyed about what “gain of function” was most important to their quality of life, men and women with spinal cord injury (SCI) stated that sexuality was a major priority even above the return of sensation, ability to walk, and normal bladder and bowel function (1), that SCI altered their sexual sense of self, and that improving their sexual function would improve their quality of life (2,3).

In this chapter, we focus on sexuality in those persons with physical disability versus mental, cognitive, or developmental disabilities. When a condition becomes chronic or gradually debilitating, expectations of recovery must give way to pursuit of adaptation (4). “Sexual rehabilitation” not only implies salvaging and restoring remaining function but also implies remaking and readjusting. Only in the arena of sexuality can rehabilitation go beyond the local affected area. Loss of physical, sensory and motor options forces an appreciation of the power of “brain or cerebral sex” and the development of evolved sexual experience through the process of “neuroplasticity.” This is best explained by the analogy of how, despite the “hardware” being altered by disability, the “software” can still be intact and adaptable (5). Research is just starting to address the potential of “sensory substitution” in persons with disability (6). In the medical treatment of sexuality after disability, the mind-body interaction cannot be forgotten by the clinician or the therapeutic value of this potential is lost.

THE COMPLEXITY OF SEXUAL FUNCTIONING

Models of Sexual Function

In 1966, Masters and Johnson (M&J) proposed a physiological model of sexual response after studying over 600 able-bodied men and women in a laboratory setting where details of genital, breast, skin, muscle, and cardiovascular functions were observed (7). Since the M&J model, many more models have been proposed embracing the biopsychosocial approach to sexuality, especially for women (8). A recent, clinically applicable model of human sexual response useful for both sexes (4) is well suited to the complexity of the disability population, who may not be neurologically intact. However, the M&J model is still helpful in understanding the neurophysiology of sex and in defining the sexual dysfunctions.

The M&J model describes a four-phase “sexual mountain” of rising and declining sexual arousal: (a) excitement, (b) plateau (high arousal before orgasm), (c) orgasm, and (d) resolution (the reversal and/or dissipation of phase 1). Pelvic vasocongestion and neuromuscular tension results in tumescence and/or erection of the erectile tissues in both men and women and, additionally in women, vaginal lubrication and accommodation (lengthening and uterine lifting). Cardiovascular and respiration parameters increase, and sweating appears. During the “preorgasmic” phase, men have maximal erection and rigidity and approach ejaculatory inevitability, and in women, the outer third of the vagina forms a thickening called the “orgasmic platform.” Orgasm is a pleasant experience recognized in the genital area, brain, or total body and is accompanied by involuntary rhythmic contractions of the pelvic floor muscles in both men and women as well as smooth muscle contractions of internal accessory sexual organ and structures. The orgasmic phase has the maximal heart rate (HR), blood pressure (BP), and respiratory rate in neurologically intact persons. Ejaculation, the process of forward (antegrade) seminal fluid expulsion through the urethra, is usually concomitant with orgasm in men but can be separated in neurological conditions. The final resolution phase is characterized by gradual reversal of the tumescence, pelvic vasocongestion, neuromuscular tension, and cardiovascular parameters noted above. Men, but not women, have a physiological refractory period (7), and women have the capacity to have extended, repeated (multiple), or compounded orgasms. However, some men also report the capacity for multiple orgasm (9).

Definitions of Sexual Dysfunctions

Male and female sexual dysfunctions are multifactorial, involving physiological, psychological, social, and emotional components. Sexual desire or libido is especially complex and under central neurophysiologic control. Kaplan (10) defines it as “the experience of specific sensations that motivate the individual to initiate or become responsive to sexual stimulation.” If a person with disability experiences a change in drive, it is most
often a reduction, known as hypoactive sexual desire disorder. In rarer cases, sexual drive may be increased due to brain stimulation from injury (i.e., hypersexuality of the Klüver-Bucy syndrome) or medications (replacement of the sex excitatory neurotransmitter dopamine [DA] in a patient with Parkinson’s disease). Three other disorders of sexual dysfunction are clinically identified using the M&J model: arousal disorders (disorders of male penile erection [or erectile dysfunction, ED] and female subjective and objective arousal disorders), ejaculatory dysfunction, and orgasmic disorders. Problems within these phases in both sexes can include fertility issues and pain specific to sexual activities (dyspareunia). A full definition of sexual dysfunctions can be found elsewhere (11).

**SEXUAL NEUROPHYSIOLOGY**

Both somatic and autonomic nerves provide important sexual afferent and efferent communication between the brain and periphery. Autonomic nerves are activated by stretch or lack of oxygen, rather than by touch or temperature (12). In the somatic system, tactile inputs include light touch, temperature, pressure, vibration, and pain. It appears that both somatic and autonomic nerves appear critical to recognize stimuli as “sexual” (12). The cerebral evaluation of skin and visceral stimulation; of visual, gustatory, and auditory inputs; and of fantasy and emotion forms either a sexually excitatory or inhibitory signal. This generated neuronal “trigger,” coordinated in the limbic system, hypothalamus, and other midbrain structures, is carried distally through the brainstem and spinal tracts and can be modulated by mood, hormones, emotions, and physical factors (4). In everyday life, this signal is usually inhibitory, until engagement of sexual activity is deemed appropriate and excitatory signals dominate, instigating the triggering of spinal cord reflexes for sexual function. Performance anxiety, distraction, and fear of negative consequences of being sexual (i.e., ED) can cause the supratentorial inhibition on the spinal cord reflexes to remain. In chronic disease or neurological conditions, signals may be directly disrupted by nerve injury or nerve degeneration, the physiology of an end organ and its ability to respond to the stimulus can be altered, or pain or other limitations can cause distraction away from the sexual focus, rendering sexual response unreliable.

Once the descending signal has passed down the spinal cord, the pelvic sexual organs receive their information from the spinal cord via three nerve pathways: (a) sacral parasympathetic (pelvic nerves and pelvic plexus), (b) thoracolumbar sympathetic (hypogastric nerves and lumbar sympathetic chain), and (c) somatic (bilateral pudendal nerves) (13). More attention is being paid to try and preserve these nerve tracts in pelvic surgery to reduce the amount of resulting sexual dysfunction (14,15). Sexual arousal leads to genital erectile tissue engorgement and pelvic vasocongestion via vasodilatation of arteries and smooth muscle relaxation. In women, lubrication depends on both intact innervation and normal estrogen levels (16), and in men, internal accessory organ function (including the production of semen) and erection are dependent on adequate testosterone levels.

There are two neurological pathways for genital arousal: reflexogenic and psychogenic. While psychogenic and reflexogenic pathways can act independently, they usually act synergistically to determine the genital response via a final common pathway involving a sacral parasympathetic route (13). The reflexogenic pathway, triggered by direct stimulation of the genital organs, has an afferent component conveyed by the pudendal nerve to the S2–4 segments of the spinal cord (17). The responding efferent component returns through the sacral parasympathetic center, contributing fibers to the pelvic nerve and onto the cavernosal nerves at the genitalia. The psychogenic pathway is of supraspinal origin (auditory, imaginative, visual, etc.) involving the medial preoptic nucleus (MPOA), paraventricular nucleus of the hypothalamus, and reticular activating systems (the latter involved with nocturnal arousal during REM sleep) (14). The long efferent tracts from the central nervous system between cortex, cord, and autonomic nervous system must be intact to elicit the thoracolumbar sympathetic center and the sacral parasympathetic center (17). Complete spinal cord injury (SCI) above the level of the psychogenic pathway eliminates the connection to it and the natural supratentorial inhibitory control, enhancing the reflexogenic mechanism initiated by touch (14). SCI involving the lumbosacral region results in loss of reflexogenic but not psychogenic capacity, since the pathway from the brain to the thoracolumbar center is still intact. The sympathetic nervous system can maintain genital arousal capacity after injury to parasympathetic pathways (13), and has a role in the development of psychogenic arousal (18). Both men and women undergo measurable genital arousal during rapid eye movement (REM) sleep (16,19). In men, the presence of REM sleep (morning) erections is a sign that daytime erection problems are more likely psychogenic in nature. An exception is in multiple sclerosis (MS), where, despite organic disturbance with daytime, erotic erections, nocturnal erections can be frustratingly preserved (20).

At the genital level, cavernosal smooth muscle relaxation results in vasocongestion, tumescence, and elongation of the erectile tissue in both men and women. The tunic albuginea, a fibroelastic stocking surrounding the corpora cavernosa, becomes stretched with tumescence, tightening its elastic fibers and kinking the emissary veins that pierce it. This occludes the venous blood outflow while high pressure arterial inflow continues. In men, this veno-occlusive mechanism, along with bulbocavernosus (BC) muscle contraction, results in a rigid erection and ischiocavernous (IC) muscle contraction helps propel ejaculated semen (14,21). In women, the veno-occlusive mechanism is less prominent due to a less effectual tunic (16). Vaginal lubrication (a plasma transudate from the blood circulating through the vessels of the vaginal epithelium), lengthening of the vagina, and uterine, urethral, labial, and pelvic ligament vasocongestion occurs with female arousal (15). Passive dilatation of the vagina results in a reflex contraction of both the BC and IC, indirectly affecting the clitoris and sensory perception of the clitoris (22). Spasm of either the BC or IC muscles, or injury to their pudendal innervation can
effect subjective sexual arousal and orgasm. Priapism can occur in both sexes, and a persistent arousal syndrome, potentially of organic origin, has been newly recognized in women (23,24).

While arousal is predominantly parasympathetic, ejaculation is primarily a sympathetic phenomenon. Preganglionic sympathetic fibers leave the spinal cord from the first and second lumbar segments, synapsing and eventually distributing to the vas deferens, seminal vesicles, and the prostate through the hypogastric nerves, stimulating smooth muscle contractions (12). Ejaculation occurs in two phases: seminal emission (sympathetic T10–L2) and propulsatile ejaculation or expulsion (parasympathetic S2–4 and somatic). Seminal emission involves transport of semen into the prostatic urethra via the ejaculatory ducts in the prostate. The sympathetic hypogastric nerve (L1, L2) activity closes the bladder neck to prevent retrograde ejaculation. A sense of impending “ejaculatory inevitability” proceeds propulsatile ejaculation and the seminal bolus is then propelled distally out the urethral meatus (14). Orgasm usually occurs with ejaculation, but they are not synonymous and are separate neurological entities.

Orgasm appears to be relayed through both the somatic and autonomic systems (23), but neurologically, orgasm is the least understood of the sexual phases. It is a complicated combination of local spinal cord reflexes and cerebral and autonomic influences, any of which could potentially dominate in any one orgasmic experience or be adequate within themselves. “Orgasm” after disability may include orgasmic attainment without genital stimulation (i.e. “orgasms” after SCI, orgasm arising from breast stimulation alone, etc). For example, about half of men and women with complete SCI can still experience orgasm, and a few neurophysiological theories have been proposed for this phenomenon (25). Strength of the pelvic floor contractions (somatic), the degree of engorgement of the internal genitalia (autonomic), subjective awareness of internal genitalia contractions (i.e., uterine), duration and degree of brain arousal, and interpretations of cardiovascular alterations with sexual activity are all factors in the subjective intensity of orgasmic release. While estrogen does not seem to influence the orgasmic potential in women, low androgen levels make orgasm more difficult to reach in both men and women (26). Oxytocin levels may rise during arousal and orgasm, and prolactin levels remain elevated after orgasm (23), but the significance of this is not known.

**DISABILITY–RELATED DISRUPTIONS TO SEXUAL FUNCTION**

Disability can affect sexual function through four basic mechanisms (4):

1. **Direct** effects of vascular, neurological (including pain), hormonal, anatomical, or other damage to any area functionally connected to the sex response

2. **Indirect** effects of the medical/psychological condition, such as changes to perception or judgment, sensory or motor alterations, bladder and bowel incontinence, spasticity, tremor, fatigue, anxiety, chronic pain, etc.

3. **Iatrogenic effects of treatment** (e.g., radiation, surgery, medication, and chemotherapy)

4. **Contextual factors**, that is, the biopsychosocial complexity and the situational components.

Contextual factors are especially critical in long-term care, where issues such as freedom of sexual expression, choice making capacity, right to privacy, and caretaker issues and responsibilities must be addressed (27).

**Medications**

It is well known that many drugs interfere with sexual functioning. The *central* neurotransmitters involved in the descending signal through the thalamospinal tracts are the excitatory neurotransmitters DA and noradrenaline (NA), whereas serotonin is generally inhibitory (4). Drugs that reduce NA, such as sympatholytic antihypertensives, negatively affect sexual function. DA blockers (e.g., antipsychotics) will impair libido, whereas the replenishment of DA agonists (e.g., levodopa in a patient with Parkinsonism) can elevate libido. Serotonin reuptake inhibitors (SSRIs), a class of antidepressants that cause increased levels of serotonin in the neural junction, are sexually adverse, as seen with the common side effects of reduced libido, as well as, ejaculatory and orgasmic delay (28), the latter effect which is used to treat premature ejaculation. Of the tricyclic antidepressants, amitriptyline, clomipramine, and doxepin appear to have the most negative impact on sexual function, especially orgasm and ejaculation (29). Reducing doses, short drug holidays, or switching to a “sex-friendly antidepressant” such as buproprion, nefazodone, trazodone, or mirtazapine are alternatives (30), but trazodone has been known to cause prolonged nocturnal erection and priapism (31). In the *periphery*, since genital arousal is primarily under parasympathetic control, sympathomimetic drugs cause vascular and smooth muscle constriction, inhibiting genital arousal. Drugs that antagonize cholinergic effects, such as tricyclic and SSRI antidepressants, phenothiazines, and butyrophenones, or those that decrease peripheral vascular resistance and genital blood flow, such as antihypertensives, can cause libido, erection and vaginal lubrication problems, and ejaculation issues (29,32) (Table 16-1). The use of phosphodiesterase V inhibitors (PDE5i’s) can often ameliorate the altered genital arousal side effects of such drugs as SSRIs and thiazide diuretics. PDE5i’s work less well in low testosterone milieus (32–40).

**Aging**

Age brings changes to elasticity of tissues, narrowing of vasculature, and decreased neural transmission, resulting in changes to the integrity and capacity of the sexual organs. Postmenopausal women experience less vaginal lubrication and decreased muscle spasm at orgasm. Men may take two to three times as long to achieve an erection, the erection will not be as rigid, and if the erection is lost with ejaculation, it may take even longer to regain it (7). Orgasm may be delayed and ejaculatory fluid volume
is reduced with age, likely due to reduced testosterone levels. Once ejaculation or orgasm occurs, penile detumescence and other genital changes occur rapidly and the refractory period is long (even days) in the elderly. Even with cross-cultural factors and medication issues taken into account, the prevalence of most sexual problems tends to increase with age; however, older age, net of other factors, consistently increases the likelihood of more sexual problems among men, but less so for women (the exception is poor lubrication, which is a significant issue for women as they age) (41). In women, sexual pain disorders appear to decline with age (42). Self-esteem may suffer with aging skin changes, weight gain, aches and pains, graying or loss of hair, fatigue, and declining athletic ability. Depression, a factor that blunts almost all aspects of sexual functioning, is far more common in the elderly. Dealing with a disability may lower expectations of both aging partners, but a previous healthy sexual relationship and open communication skills bode well for maintaining sexual intimacy despite health issues (4).

### Impending Mortality

Variable effects on sexuality can be expected with impending death. Some persons lose interest and focus on survival, and others embrace sexuality as a quality of life/intimacy issue that represents the positive aspects of living. Partnership support is obviously critical in this outlook. End stage disabilities and cancer share several assaults on sexuality, including medication effects, fatigue, change of body image with any disfiguring surgery or amputation, altered hormonal status, catheters, stomas, secondary paralysis, and relationship and psychological stress. Often sexual effects of life saving/prolonging procedures and surgery are inadequately discussed before and/or after the procedures, leading to sexual distress.

### Neurological Changes

The advent of noninvasive brain imaging techniques such as PET and fMRI have demonstrated areas in the brain associated with imagery, visual sexual stimulation, arousal, and orgasm in both sexes (23,26,43). Injuries to these specific brain areas, and to the spinal cord and peripheral nerves innervating and sending sensory information to the cognitive centers, will profoundly affect sexual functioning. Furthermore, the sequela of having a neurological disability that can affect cognitive, motor, sensory, and autonomic functions will have multiple effects on sexuality. Most acquired neurological damage must undergo 2 to 4 years of recovery before the remaining neurological sexual function is known.

#### Psychological Factors

Adjusting to disability is a highly unique matter, related to the severity of losses caused by the disability (not just physical, but emotional, relational, and financial), whether depression occurs and/or persists, the presence and strength of support systems, and premorbid personality factors and coping skills. For some persons, sexuality is a highly integral part of their self-esteem and/or self-image, and the disruption occurring with disability can be devastating. Whether the disability occurred acutely or slowly over time can also affect sexual adjustment, since readiness to approach the issue again has to be balanced with other life priorities such as physical recovery or financial losses. Age and sexual experience are critical factors. Willingness to reenter the sexual arena with past or new partners and openness to experiment with a changed body and/or utilize assistive aids takes a certain mindset and readiness. Anxiety can also play a part, resulting in arousal problems or premature ejaculation, and can also decrease interest or orgasmic attainment. Disability and chronic illness can affect both persons of the couple often with reversal of traditional masculine and feminine roles, loss of economic security, caretaking issues that depress sexual attraction, fatigue on the part of both persons, and forced alteration of life goals that may no longer make the couple compatible (40).

#### Cardiovascular Factors and Risks

Several studies address the cardiovascular response to sexual activity and the possible risk of an acute cardiac event.

---

**TABLE 16.1 Rehabilitation Medications Affecting Sexual Functioning**

<table>
<thead>
<tr>
<th>Medication/Class</th>
<th>Cause of Sexual Effect (28,29,32–40)</th>
<th>Dysfunction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Antihypertensives</td>
<td>Central sympatholytic/peripheral α-blocker</td>
<td>AD, ↓SD, EjD</td>
</tr>
<tr>
<td>SSRIs</td>
<td>Sex suppressant, especially orgasm</td>
<td>EjD, ↓SD, AD</td>
</tr>
<tr>
<td>Anticonvulsants</td>
<td>↓Arousal</td>
<td>AD, ↓SD</td>
</tr>
<tr>
<td>Opiates, marijuana, antipsychotics</td>
<td>↓Testosterone levels and/or ↑prolactin levels: with chronic use, sex suppressant</td>
<td>↓SD, AD, EjD</td>
</tr>
<tr>
<td>Antispasmodics</td>
<td>May suppress sexual reflexes</td>
<td>AD, EjD</td>
</tr>
<tr>
<td>Antilipidemins</td>
<td>Associated with ED</td>
<td></td>
</tr>
<tr>
<td>Indomethacin</td>
<td>Causing fertility problems</td>
<td></td>
</tr>
<tr>
<td>Gabapentin</td>
<td>Causing variable effects</td>
<td></td>
</tr>
<tr>
<td>Digoxin</td>
<td>↓Testosterone and LH and ↑estrogen</td>
<td>↓SD, AD</td>
</tr>
<tr>
<td>H2-blockers</td>
<td>Antiandrogen types affect arousal and can cause reversible spermatogenic suppression</td>
<td>↓SD, AD</td>
</tr>
</tbody>
</table>

↓SD, decreased sex drive; AD, arousal dysfunction; EjD, ejaculatory or orgasmic dysfunction.
An average peak HR of 110 to 130 bpm and peak systolic BP of 150 to 170 mm Hg was observed during the sexual activities of couples with longstanding relationships (44,45). Generally, the energy requirements during sexual activity do not exceed 4 to 5 METs (46). Furthermore, it was established that walking on a treadmill at 3 miles/hour at a 5% grade or climbing two flights of stairs at a rate of 20 steps in 10 seconds would require equivalent amounts of energy as sexual intercourse (44). This information on the two-flight test, which is equivalent to 6 METs, has been widely used to determine the risk of ischemia and/or safe return to sexual activity after cardiac events. However, even with no underlying cardiac muscle or blood vessel pathology, disordered autonomic control of the heart and/or blood vessels can result in significant alteration of cardiovascular responses during sexual activities. This is especially prominent in individuals with SCI (47), where abnormal autonomic control could result in the onset of autonomic dysreflexia (AD) during sexual activities (especially with ejaculation) and/or continuation of AD after activities (48) (Fig. 16-1). AD is a condition characterized by episodic hypertension, often associated with cardiac arrhythmias (49). Hypertensive crises during the episodes of AD can result in cardiac arrest, retinal or subarachnoid hemorrhages, cerebral vascular accidents, and death (50). AD symptoms with sexual activity often attenuate with time, but BP elevations can remain deceptively high.

Bladder and Bowel Concerns
Bladder and bowel issues are one of the most important factors in sexual quality of life in persons with disability (3,51,52), as they influence willingness to be social or to leave the security and privacy of one’s home where supplies are located. For some, life can seem dictated by washroom availability. The potential for odor from, or leakage or visibility of, urine or stool collection devices, can affect sexual self-esteem and social comfort, and interfere with the interest of a partner. The time required for a bowel program, for catheterization and/or attendant care, can be one of the main dissuaders from being spontaneous or social, and the risk of incontinence with sexual activity can severely limit sexual expression. To pursue a relationship or sexual partnership, the risk of major embarrassment has to be courageously accepted by the person with a disability. Bladder and bowel management must be assertively pursued to improve body image and independence, even if it means major measures, such as continent urinary diversions (53).

![BP and ECG recordings in individuals with C6 ASIA A SCI during VS procedure for sperm retrieval. Prior to VS (Rest, recording 1), there is relative hypotension (69/49 mm Hg) with regular HR of 82 bpm. With initiation of VS, arterial BP gradually increased to 148/105 mm Hg and the patient developed bradycardia (recording 2, HR 60 bpm) suggestive of typical episode of AD. Finally, at the time of ejaculation (recording 3) arterial BP surged to 240/130 mm Hg accompanied with a short period of cardiac arrhythmia. Interesting to mention is that in this case the patient did deny any symptoms typical for AD such as headache, blurred vision, sweating, and piloerection (personal observations).](image)
Pain Issues
In a survey of patients with chronic pain attending treatment centers, sexual difficulties were found to be twice that of the general population, and these difficulties were not simply related to mood or disability (54). Two types of pain conditions can affect sexuality. Pain localized to the genitals, pelvis, or other sexualized areas can directly interfere with sexual function and motivation, whereas chronic pain conditions can interfere with daily life and indirectly affect sexuality. Chronic pain conditions can result in depression, weight loss or gain, physical deconditioning, and loss of self-image, resulting in decreased mental, physical, and sexual functioning (54). The severity of pain and its location can also have a major role in sexual quality of life. Psychological variables are important and need to be incorporated into therapeutic approaches. Sexual dysfunction was found to be more frequently reported by those with greater disability and depression, shorter pain duration, and infrequent use of coping self-statements (55). With genital pain, because the issues are so complex, multimodal and multidisciplinary intervention is always required (56).

Fertility Concerns
With any disability, the question of fertility needs to be approached from four directions (57): (a) sexual functioning problems interfering with the ability to enter into the act of intercourse and for men, to ejaculate; (b) hormonal alterations secondary to the condition (i.e., head injury) or treatment (i.e., MS medications, brain irradiation); (c) altered oocyte or sperm quality secondary to the condition or treatment; and (d) inheritable genetic traits (58). In most cases, erection problems in men can be overcome with current available treatments, but ejaculation dysfunction may require sperm retrieval methods such as vibrostimulation (VS) and electroejaculation (EE), or surgical technology where sperm is aspirated from the epididymis or testicle (50). The less invasive VS is quite successful in those men with SCI whose injuries are higher than T10, with the more invasive EE being reserved for VS failures, but EE can also be utilized in conditions such as seminal emission failure secondary to retroperitoneal lymph node dissections for testicular cancer, or ejaculation failure in diabetes (59). Depending on the semen quality and quantity, obtained sperm can be placed in the vagina (intravaginal insemination), prepared and placed in the uterus (intrauterine insemination), or utilized with higher technology such as in vitro fertilization and intracytoplasmic sperm injection (IVF/ICSI). Women may not be able to have sexual intercourse due to a sexual pain disorder or other sexual dysfunction, or if adductor spasm is severe. In men, SCI per se negatively reduces sperm motility and increases abnormal morphology likely secondary to abnormalities within the seminal fluid (59). In women with SCI, pregnancy is associated with increased risks and complications such as increased incidence of urinary tract infections, changes in bladder management, increased risk for skin breakdown, difficulty with transfers due to increasing maternal weight, increased risk for deep vein thrombosis, delayed bowel emptying, pedal edema, vaginal spotting, fatigue, and thrombophlebitis. Special monitoring is necessary during labor and delivery as labor may not be felt or may present differently than in noninjured women. AD can occur, Caesarian sections and vacuum and forcep extractions are more common, and women with SCI also deliver lower birth-weight infants, the etiology of which is not clearly understood (50). Further information can be found in Chapters 72 and 80. Women with other chronic conditions also have to evaluate the safety and capacity to carry a pregnancy to term and undergo delivery.

SEXUALITY IN SPECIFIC DISABILITIES

Spinal Cord Injury
SCI has been the most extensively researched (animal and human) disability for understanding sexual neurophysiology and provides an excellent model for the holistic and reintegration approach that comprises sexual rehabilitation (18,40,50,60–63). After complete SCI above the lumbosacral spinal cord center (usually above T10), reflexogenic arousal should be preserved whereas psychogenic arousal will not (due to the lesion interrupting the pathways from the brain to the T10–L2 spinal segment, the level responsible for psychogenic arousal). Conversely, complete lesions interrupting the sacral reflexogenic pathways will result on the reliance of psychogenic arousal to promote genital erection in men and women. Unfortunately, in men with sacral SCI reliant on psychogenic arousal, the stimulation of the hypogastric and sympathetic pathways through mental arousal can trigger unwanted seminal emission and penile detumescence, a frustrating experience. Men and women with varying SCI levels and incompleteness will experience various capacities to tap into their psychogenic or reflexogenic pathways. The degree of preservation of sensation in the T11-L2 dermatomes is helpful in predicting those persons with SCI who are capable of psychogenic arousal (61,62), whereas the presence of a positive bulbocavernous reflex (BCR) is indicative of an intact sacral reflex, boding well for the reflexogenic arousal capacity. While over ¾ of men with SCI can obtain an erection, firmness and reliability in sexual situations can be a problem. Ejaculatory disorders (primarily anejaculation) are highly prevalent (over 90%) and thus fertility is a major issue for men with SCI (5). Natural ejaculation is most likely to occur in men with complete conus medullaris or cauda equina lesions, and least likely in men with complete supraconal lesions (64). However, VS-assisted ejaculation is more reliable in complete supraconal lesions (59). Orgasm is attainable in 40% to 45% of men (most often with ejaculation but not always) (51,61) and approximately 50% of women after SCI (62). Lack of genital sensation, and especially lower motor neuron injury affecting the sacral segments, make it significantly less likely to reach a genitally triggered orgasm (3,25). While sexual satisfaction after injury is lower in both men and women after SCI, pursuing intimacy remains a major quality of life issue for the majority living with SCI (3).
Multiple Sclerosis
It is estimated that more than 70% of men and women with MS experience some form of sexual dysfunction during their disease (65) even in the absence of severe disability (66,67). Sexual symptomatology can fluctuate with the changing nature and duration of MS, and will depend on the localization of lesions in the spinal cord or brain (43,68–70). Libido problems can be acute or chronic, and are linked to depression, presence of cerebral plaques, and the inability to experience orgasm (72). In ambulatory male patients with MS, 60% had ED, 50% had ejaculatory or orgasmic dysfunction or both, and 40% reduced desire (43). In men, if orgasm can be reached, ejaculatory fluid may flow out by gravity, minutes later. Women with MS report unsatisfactory sexual lives secondary to fatigue (68%), reduced genital sensation (48%), reduced vaginal lubrication and difficulty with arousal (35%), difficulty reaching orgasm and anorgasmia (72%), and sexual pain disorders (28). Furthermore, spasticity, fatigue, changes to genital sensation (71), muscle weakness, fear of fatigue with sexual activity and concerns about continence during sex (72) can affect the ability or willingness to engage in certain sexual acts, especially intercourse. Cognitive changes and psychological impact of the disorder can result in anxiety, depression, poor self-image, and lack of interest in sex in both patient and partner (40).

Limb Amputation
Sexuality in the amputee population has not attracted researchers’ attention since genital sexual function is often preserved (73). However, if the amputation is not traumatic but secondary to diabetes or cancer treatments, there will be added sexual complications from the underlying disorder. The medically indicated loss of a limb can have serious psychosexual impact, as can the surgical procedure itself, and result in perioperative and prosthetic device pain, deformity, phantom limb pain, and inability to perform customary sexual acts. Amputees have to adjust to a different body image, and depression may persist for years (74). Rehabilitative considerations for sexual acts are important, and physiotherapy assessment and education is necessary. Positioning options for sexual intercourse may need to be provided, and while preservation of the knee joint is helpful for balance, positioning with pillows or wedges, or the use of spoon or side-by-side intercourse may be necessary for upper limb or transfemoral amputations (74).

Parkinson’s Disease
Sexual dysfunction is common in Parkinson’s disease (PD). At least 75% of women with PD report difficulties with arousal and orgasm; and 50% experience low sexual desire. In men, 70% experience erectile difficulties, 40% have premature ejaculation, and 40% have delayed orgasm (75). Furthermore, bladder and bowel dysfunction that could interfere with sexual activity is prominent in patients with PD compared to controls of the same ages (76). An increase in sexual drive can be experienced with dopaminergic drugs, but often genital sexual dysfunctions can remain, leading to increased sexual frustration for the person with PD. Because of the psychomotor and disinhibiting cognitive changes, partner attraction and caretaker concerns can be an issue.

Neuromuscular Disorders
Though there is scant literature and research on the sexual consequences of neuromuscular diseases (including myopathies and neurological disorders), it is clear that sexuality is a very important topic for this population and their partners, but it is rarely discussed in medical settings. Sexual dysfunction is usually due to non-genital causes, e.g., physical limitations, privacy, and caregiver issues (60). Since these disorders vary tremendously from a moderate loss of stamina and strength, to life-threatening motor disability including loss of pulmonary capacity, the effect on sexual functioning can vary greatly. Age of onset (childhood or adult) of a neuromuscular disorder and its rate of progression are critical determinants of a patient's sexual practices and problems (77). Amyotrophic lateral sclerosis (ALS) spares sensory and autonomic neurons, maintaining sexual, bladder, and bowel functions while rapidly destroying motor neurons leading to almost total paralysis of the whole body. Sexuality was found to be an important issue, including in persons with ventilator-dependent ALS, and identified sexual problems included decreased libido, passivity of partner, and the patient’s own passivity due to physical weakness and body image changes due to ALS (78). Only in late stages did persons with ALS lose sexual desire (79). Factors such as independence in community living, choice of attendant care, and the ability to self-direct one’s personal affairs despite severe disability were found to support the high rate of successful marriage in persons with ALS (80).

Traumatic Brain Injury
Traumatic brain injury (TBI) can have variable effects on sexual function depending on the severity of the injury, the location of the lesion, and whether there are endocrinological disturbances and/or seizures secondary to the brain insult. Unfortunately, TBI often occurs in young persons between the ages of 15 and 24 (81), where sexual potential and experimentation with intimate companionships are at a peak. Decreased desire and arousal disorders are common, whereas hypersexuality and/or sexual deviancy post-brain injury is rare. Hypersexuality is seen in the Kluver-Bucy syndrome (82) and in bilateral thalamic infarction (83). In one study, compared with uninjured controls, men and women with brain injury reported lower energy and low interest in initiating sex (possibly due to decreased levels of testosterone), and difficulties with erection, vaginal dryness causing pain during sexual activities, and orgasm (84). While ED is common, it is variably reported from 10% to 71% (84), and can be organic in nature even if bladder or bowel problems are not evident. TBI will not only affect the cognitive status and personality (e.g., irritability and loss of anger control), but can also affect physical issues such as sensory loss to erogenous areas, paralysis, and/or spasticity. TBI does
not have to be severe in order to develop hypogonadotropic hypogonadism (43). A comprehensive review of brain injury lesions and their sexual sequela can be found in other texts (40,43,85). The problem of “inappropriate behavior” needs to be properly addressed, as it may be the result of a correctable concern (i.e., provision of sexual relief), or of values and beliefs of caregiver versus client (86).

**Stroke**

Despite the high incidence and prevalence of stroke, only limited numbers of studies with a small numbers of subjects have addressed issues of sexuality (87). Unfortunately, the real effect of stroke on sexuality is clouded by the fact that most individuals with stroke are elderly, are taking a number of medications that can affect sexuality, and already have a variety of concomitant medical problems that can also impact their sexuality. Men and women who suffer strokes frequently report a decrease in libido, decrease in lubrication and decline in erectile capability in comparison with prestroke, and a decrease in the frequency of sexual activity (88,89). Mutual verbal and nonverbal responsiveness, as well as frequency of caressing and touching with intention of being sexual, can decline (90). This may be due to inhibition or disinterest on the part of the person who had the stroke, but may also be due to reluctance of the partner when unattractive behaviors such as drooling or incontinence are present (89,90). Fortunately, practical guides to resuming activity after stroke are becoming more available (91).

**Coronary Artery Disease**

In contrast to stroke, a large body of literature is focused on the impact of cardiac disease on sexual functioning, accompanied by a great imbalance in the attention paid to male as opposed to female sexual dysfunctions following cardiac events. In one study, following a first cardiac event, up to 75% of middle-aged male patients either decreased or stopped sexual activity, and 80% of male patients with congestive heart failure reported either marked problems with or an inability to engage in sex (92). This dramatic decline in sexual activity occurs despite the fact that sexual activity requires fairly low energy expenditures as noted earlier (44). Similar to stroke survivors, the reasons patients and partners give for the decline in intimacy are more complicated than mere loss of interest or an inability to respond, instead they often reflect the multiple preoccupations and stresses of a chronic disease, as well as unwarranted fears of experiencing a heart attack during sex. However, clinical data suggest that sex is a comparatively weak precipitant of acute coronary events, accounting for only 0.5% to 1.0% of all such events (93).

**Pulmonary Disease**

Patients with emphysema and chronic bronchitis, or chronic obstructive pulmonary disease (COPD), often have concomitant issues with diabetes mellitus, endocrine abnormalities, cardiac disease, and autonomic nerve dysfunction that can affect sexual functioning, so it is difficult to differentiate those sexual concerns stemming from pulmonary problems alone and those of the comorbidities and their medications used to treat them. However, in COPD, dyspnea leads to diminished activity tolerance, and since full exhalation is not possible because of pre-existing obstruction, attaining the respiratory rate of 40 to 60 breaths per minute required for sexual activity can be a problem. Partner weight on the chest can also worsen the dyspnea. These stressors, along with fear and anxiety, lead to heightened sympathetic tone, hyperventilation, bronchospasm, coughing and further anxiety, such that a vicious cycle is created (94). Furthermore, testosterone levels are often reduced with chronic hypoxia, which contributes to poor genital arousal capacity. Specific breathing techniques, redirection to sexual practices and positioning to reduce dyspnea, use of inhalers, oxygen or cough drops prior to or during activity, and training with relaxation techniques can assist these couples (40,94).

**Diabetes**

Diabetes affects the core physiology of sexual function in both sexes. Small and large vessel disease, peripheral neuropathy, smooth muscle dysfunction in the genitalia, and psychological problems are variably responsible for libido, arousal, and orgasmic difficulties, as well as dyspareunia in women. Whereas the incidence of sexual dysfunction in men with diabetes approaches 50%, it is only slightly lower in diabetic women (28). ED also correlates with poor glycemic control (95), and therefore can be reversible in some cases. There are less frequent effects noted on ejaculation and orgasm, but the slow and steady onset of retrograde ejaculation is a known complication and can affect fertility (96). In both genders, neurovascular pathology can affect genital sensation, thus impairing orgasm (97). A significant positive association is found between the number of diabetic complications and the number of sexual complaints, and in women with Type 1 diabetes, sexual dysfunction is also related to depression and the quality of the relationship with the partner (98). Sexual dissatisfaction is further correlated with relationship and self-esteem problems, secondary medical conditions such as stroke or chronic renal failure (CRF), and reactions to catastrophic complications such as limb amputation (96).

**Chronic Renal Failure**

The comorbid etiologies of chronic renal failure (CRF) are responsible for ED in men, menstrual abnormalities in women, and the decreased libido and fertility in both sexes (28). Physiologic changes from uremia and the disturbances in the hypothalamic-pituitary-gonadal axis can occur before and/or after the initiation of hemodialysis or continuous ambulatory peritoneal dialysis. Central disturbances appear more prominent in uremic women, whereas the more subtle disturbances in the hypothalamic-pituitary-gonadal axis and impaired gonadal function are the main issues in men (99). Hyperprolactinemia in both men and women results in decreased libido, decreased frequency of sexual intercourse rates, orgasmic difficulties, and abbreviated longevity of sexual activity between couples (100).

**Arthritis and Connective Tissue Diseases**

Individuals with osteoarthritis and rheumatoid arthritis have many assaults on their sexuality, centered around pain, joint
stiffness and/or deformity, weakness, and depression. Poor hand function, loss of range of motion to the hips, adduction difficulties, knee pain, and back spasm can affect the mechanics for sexual acts (including pelvic thrusting), and positioning. Use of corticosteroids for treatment of rheumatoid disease can cause acne, truncal obesity, hirsutism, steroid facies (moon face), and other features that can affect sexual self-esteem. Other immunosuppressant medications can cause bladder problems and oral-genital ulceration (101). Approximately one third of patients with ankylosing spondylitis have some degree of sexual dysfunction, especially those with depression and joint mobility issues (101,102) and a significant degree of morning stiffness (> 4 hours) (103). Persons with systemic lupus erythematosus (SLE) have greater impairment with sexual dysfunction the greater the severity of disease (101). In women with systemic sclerosis, vaginal involvement, skin tightness, and muscle weakness affect sexual function, and decreased orgasmic function is prominent (104). Compared to healthy normal controls, women with fibromyalgia have similar overall sexual function, but have more problems with sexual desire, satisfaction, pain, and insensitivity of their genitals around the times of sexual activity (105). Another study in women with fibromyalgia showed coexistent major depression having no additional negative effect on sexual function (106).

SEXUAL REHABILITATION FRAMEWORK

Failing to address sexuality in a rehabilitation setting is no longer an option. With clear evidence that sexuality is very important in this population, and with patient-centered care, professional discomfort about discussing such issues should be secondary to the need to have patient’s sexual concerns addressed. A basic screening question such as “Many persons with [your disability/chronic illness] experience sexual concerns [examples can be given]. Do you have any concerns you would like to address?” allows the door to the topic to be opened, even if the person with the disability/chronic illness is not acutely in distress. Asking whether someone is “currently sexually active” may not be as helpful a screening question, since it assumes a willing partner is available, and sexual concerns can exist from solo or partnered activities. Addressing sexuality confirms that the issue is acceptable and medically legitimate, and allows freedom of choice to pursue therapeutic intervention.

One does not have to be a sexual medicine expert to take a sexual history. Below is a user-friendly framework adapted from Szasz (107) for clinicians to use, since it allows for an intuitive process to assess and manage patients. Patients are assessed holistically with regard to their disability/chronic illness since not all questions are genitally focused. A more detailed version is available (4):

1. Sexual interest (biological urge combined with the motivation and/or wish to be sexual)
2. Sexual response (mental and genital arousal, ejaculation in men, ability to attain orgasm and orgasmic quality)
3. Changes to genital sensation or other erogenous zones (loss of erotic zone sensitivity or hypersensitivities in specific areas, etc.)
4. Changes to motor function (hand function, balance, ability to transfer to a bed, hold a partner, etc.)
5. Bladder and bowel issues (management strategies, concerns with continence during social and sexual activities, and social implications)
6. Factors associated with the condition (medication effects, alteration in hormone status, pain, fatigue, AD, anemia, etc.)
7. Practical use of contraception, concerns about fertility, pregnancy, delivery
8. Parenting issues specific to the disability or illness
9. Relationship and partnership issues (the sexual context within which the person lives)
10. Sexual self-esteem and self-view issues (physical presentation, disfigurement, masculinity/femininity, changes in gender role, etc.)

The physical examination and any accompanying blood-work should appropriately address the issues elicited above. Serum-free or bioavailable testosterone levels should be checked in persons with head injury, especially if they have noticed a change in sexual desire. It is important to assess genital sensation (light touch and pinprick), rectal tone, voluntary anal contraction, and the presence or absence of a BCR. As noted earlier the BCR signifies an intact sacral reflex and the probability of reflexogenic arousal and ejaculation capacity. Detection of pinprick sensation around the glans penis or clitoris with the combined ability to voluntarily contract the anal opening is helpful in assessing the capacity for genitally triggered orgasm in persons with neurologic changes (with the exception of complete SCI) (107).

CURRENT THERAPIES FOR SEXUAL DYSFUNCTIONS

Sexual rehabilitation is multidisciplinary. While physician assessment must include inquiring about sexual concerns (as patients often do not bring up the topic), rehabilitation therapists, nurses, social workers and others may be asked questions regarding sexuality from patients who feel particularly comfortable with that particular clinician, or who gather the courage. If the inquiry may seem misplaced or inappropriate in some cases, a simple reflective “It seems you have some concerns about your sexual functioning, is that the case?” can put professionalism back into the situation without deflecting the legitimacy of the patient’s concern. Nonjudgmentally acknowledging the patient as a sexual human being, providing basic support, discussing options, or arranging a referral, can go a long way to alleviate patient anxiety. Obtaining assistance related to physical limitations for self and partner pleasuring may require the persistence of occupational or physical therapists, sexual health
Medical Options for Sexual Difficulties

Talk-oriented therapy is a critical component in the proper treatment of all sexual disorders, as the mind-body interaction cannot be separated into distinct entities. Reintegration of body image and formulations of new sensory “body maps” can happen alone or with a partner, but the presence of a long-term trusted partner appears key to maximally discovering sexual potential (108). While sex therapy per se is beyond the scope of this chapter, psychological and relationship issues must be addressed if medical therapy is to be productive (4).

In couple relationships, open communication around sexual wants, needs, and abilities, is even more critical when couples are faced with disability or illness. Persons without partners must also have sexual needs addressed, including the use of masturbation aids, as self-stimulation may be an important source of joy for this person (27).

Erection Enhancement

Oral medications that indirectly relax the penile smooth muscle and enhance an erection attained from psychogenic sexual stimulation are the first-line therapy and have been available for the last 10 years. The PDE5i’s, including Viagra (sildenafil), Levitra (vardenafil), and Cialis (tadalafil), work very well when there is a source of nitric oxide (NO) from either intact peripheral nerves (stimulated by intact arousal pathways) or healthy endothelial lining. PDE5i’s, therefore, do not work as well when there is poor neuronal or endothelial NO sources, such as in diabetes, atherosclerosis, hypertension, peripheral nerve injury, or autonomic neuropathy. In these cases, adjusting the medication to higher on demand dosing, or even using a lower dose of PDE5i on a daily basis, with or without additional PDE5i on demand, may prove beneficial. In other cases, where the PDE5i are more reliably effective, that is, ED secondary to SCI (50), psychogenic ED, and ED secondary to antidepressant use (109), lower dose on demand dosing is likely adequate since the endothelial and nerve generation of NO is presumed to be better. Since there is a slight BP lowering effect of the PDE5i, a lower starting dose of on demand PDE5i should be used in men at risk for hypotension. Lower doses should also be started in men with significant renal or hepatic impairment, or who may be taking drugs that increase the serum levels of PDE5i, such as antiretrovirals. PDE5i should be used with caution in patients with cardiac failure, within 6 months of acute myocardial infarction or stroke, and in patients with unstable angina pectoris. Co-administration of PDE5i with organic nitrates is contraindicated due to compounded lowering of BP. Relative contraindications include uncontrolled hypertension and impaired cardiac reserve, or symptomatic hypotension (such as in tetraplegia). There is relatively little interaction with other drugs, and the PDE5i’s have proven safe with no increased incidence of cardiovascular events or stroke. PDE5i have potentially beneficial effects on other body systems primarily through their smooth muscle relaxant effect. Some improvements have been seen with lower urinary tract symptoms and benign prostatic hyperplasia, and PDE5i may also be beneficial as well in vasoconstrictive disorders, may be protective against neurodegeneration and memory impairment, and are approved for use in pulmonary hypertension (109,110).

Injectable medication which directly relaxes the penile smooth muscle to create an erection is administered through intracavernosal injections (ICl) or penile injections. Prostaglandin E1 (PGE1), also called alprostadil (pharmacy compounded or commercially available as Caverject), can be injected into the side of the penis. PGE1 is most commonly used for ICI due to its...
higher safety profile, but other combinations typically of PGE1, papaverine and phentolamine, and occasionally atropine, can be compounded (111). ICI is the most effective treatment for ED, but proper technique and dosing must be taught. Men with neurological impairment require lower doses as they are at risk for prolonged erections or even priapism. Higher doses may be necessary in those with primarily vascular ED, such as diabetics and hypertensive patients. Good eyesight and adequate steady hand function is required for ICI, unless a needle injector is used or the partner agrees to administer ICI. Topical alprostadil is currently being investigated (112). An intrarethral preparation of PGE1 (MUSE) is available (113), but is less effective than ICI, especially in SCI patients (114).

Mechanical methods such as vacuum devices and penile rings are an effective nonmedicinal option but require some dexterity. After all reversible erection enhancement methods noted above have been exhausted, surgical implantation of a penile prosthesis can be considered, but since the cavernosal tissue is destroyed during placement of the implants, oral medications and ICI will no longer be effective. This permanency, plus the risks of surgery and infection, makes the placement of penile implants worthy of proper assessment (115).

Partner acceptance and cost also influence the choice of which ED therapy to use. Practical issues such as visual acuity, hand function, bladder management, whether the ED option can be performed independently or requires the assistance of a partner or caregiver, must be taken into consideration.

### Male Ejaculatory and Orgasmic Difficulties

Ejaculatory problems after injury or chronic illness (like diabetes) are most often a result of neurological changes as opposed to psychological causes. If the nerves responsible for ejaculation have been disrupted, few options are available to trigger ejaculation. Sometimes high dose pseudoephedrine can encourage seminal emission if sympathetic nerves have been compromised or if retrograde ejaculation is present (5,116–118). Obtaining an antegrade ejaculation for fertility purposes can be accomplished by functional or operative sperm retrieval methods. Those men who acquire secondary rapid or spontaneous ejaculation after spinal cord impairment may not be helped by the traditional use of SSRI medications used to treat premature ejaculation but slight improvement has been noted with phenoxybenzamine, terazosin, or prazosin (119).

Age and neurological disability can cause orgasmic delay, but it is important to remind men that an erection is not required to reach orgasm. Testosterone replacement in hypogonadal males and selection of sex-friendly antidepressant options in depressed men can be tried to assist delayed orgasm. The use of VS, especially when the vibrating attachment is cup shaped to enclose the glans penis, can assist with orgasmic attainment, even in neurologically intact men.

For fertility purposes, VS used in men with SCI involves the use of specialized powerful vibrators made with adjustable frequencies and amplitudes such that the uninjured sacral ejaculatory reflex can be triggered by placing the vibrator, or sometimes two in a sandwiching technique, on the glans penis. VS can often trigger AD, so it must initially be done in a clinic setting with experienced clinicians. If VS fails to elicit ejaculation, then EE can be tried. EE consists of placement of an electrical probe through the anal opening to approximate the location of the periprostatic nerves. Once current is applied, seminal emission is evoked. If there is genital sensory sparing, an incomplete injury or a low SCI, anesthesia may be required. Depending on the cost-benefit ratio of the patient’s health care benefits and the level of injury, it may be more efficient to do operative extraction, such as percutaneous epididymal sperm aspiration or testicular sperm extraction. However, VS and EE often provide enough sperm to allow the option of lower technologies such as intravaginal or intrauterine insemination, whereas utilization of operative techniques and the removal of only a few spermatozoa commits the partner of the man to more elaborate and expensive IVF/ICSI. Decisions on fertility technologies are also based on sperm quality, female factors such as age and tubal patency, and cost. In general, the higher the technology, the higher the chance of pregnancy per attempt.

### Female Arousal and Orgasmic Difficulties

In contrast to treatments for men with arousal difficulties, there are no approved medications for these sexual difficulties in women. However, clitoral cavernosal tissue responds as male cavernosal tissue does to PDE5i (especially in premenopausal women), but in healthy women without sexual dysfunction, it results only in enhancing vaginal engorgement during erotic stimuli but not in subjective sexual arousal (23). Some women, however, may benefit from a trial of lower dose PDE5i (120). Minor increases in vaginal lubrication were noted in randomized trials of sildenafil in women with SCI (63) and MS (121). Since direct clitoral stimulation has traditionally been the most effective method of attaining orgasm in women, VS on or near the clitoris may trigger the orgasmic reflex in women with genital sensory alterations. Some women with SCI have found vibration intravaginally or on the cervix to be more effective (122) than clitoral stimulation, but in any case, more prolonged stimulation is usually required compared to preinjury (62). The EROS-Clitoral Therapy Device (CTD), a vacuum device inducing clitoral vascular engorgement, is the only cleared-to-market device available by prescription to treat female sexual dysfunction, and it may prove to be of benefit in increasing orgasmic responses in women with altered pelvic neurophysiology (such as SCI) or with arterial insufficiency (25,123). Theoretically, initiating the BCR reflex and vasocongestion may also improve sensory awareness in women with some pelvic floor sensory preservation. In addition, by strengthening the pelvic floor, VS and/or CTD may prove helpful with urinary incontinence in women with disabilities, further enhancing their sexual self-esteem, but further trials need to be done.

### Nongenital Orgasmic Options

In spite of high arousal, orgasmic experiences can still be elusive to many persons with neurological-related disabilities. For men and women with SCI, the process of learning to enhance one’s sexual responsiveness increases with time, positive
partner sexual experience, emotional intimacy, and general open-mindedness (108). The philosophies of mindfulness and Tantric sexual practices, with their emphasis on total body stimulation, relaxation, and mind transcendence, have been particularly useful adjuncts for enhancing sexuality and the learned ability to achieve non-genital orgasm.

**CONCLUSIONS**

Sexual dysfunction is highly prevalent in the disability and chronic illness population. For many persons living with a disability, sexuality is a very important quality of life issue, and needs to be respected and treated as a legitimate medical concern. For rehabilitative clinicians, it is critical that specific sexual dysfunctions (sexual interest, arousal, ejaculation, and orgasm) not only be delineated but also be placed within the sexual, relationship, and disability-related contexts for treatment strategies to be successful. The use of a 10-point, user-friendly sexual rehabilitation framework can assist the rehabilitative clinician to approach, assess, and even manage the sexual concerns of their clients. The key to successful sexual rehabilitation is being physically and mentally proactive during recovery, to optimize physiological potential before adding therapies, to adapt if necessary, and to be open to new experiences (59). This may range from simple strategies such as the use of PDE5i, to enhancing genital stimulation with assistive devices in order to reinforce former sexual pathways and reflexes, to learning how to utilize nongenital sensate areas for sexual arousal and even orgasm. There may even be a role for pelvic electrostimulation (124) or neuromodulation with sacral implants (125) in persons with neurological impairment in the future.

Sexuality is an integral and meaningful part of being human. Embracing sexual rehabilitation with curiosity and respect is an essential part of comprehensive rehabilitation.

**ACKNOWLEDGMENTS**

The authors would like to express their appreciation for the support from ICORD, the Neurotrauma/Rick Hansen Foundation, Heart and Stroke Foundation, Canadian Institutes of Health Research, Christopher and Dana Reeve Foundation, Disabilities Health Research Network, and to all the colleagues and patients who have taught them the art and science of sexual rehabilitation.

**REFERENCES**


CHAPTER 17

Vocational Rehabilitation, Independent Living, and Consumerism

INTRODUCTION

According to the latest Harris poll of Americans with disabilities in 2004, approximately two thirds of working-age people with disabilities are unemployed (1). Other survey data indicate that, of those individuals who are unemployed, the vast majority (two out of three) wanted to work (2). Work plays a central role in life and is a common source of self-identify and financial independence in American society (3). Since its beginning, a major goal of rehabilitation in the United States has been to help people with disabilities become productive members of society through the activity of holding a job. Vocational rehabilitation (VR) programs have been specifically designed to promote work opportunities for people with disabilities. Historically, the focus of VR services has been to assist persons with disabilities who have vocational potential. It was not until 1978 that legislation defining VR services also included provision of services for individuals without clear vocational goals (4). Title VII, Comprehensive Services for Independent Living (IL), an amendment to the Rehabilitation Act of 1973 (PL 93-112), authorizes services for people with severe disabilities, those who require multiple services over an extended period of time and persons whose disability prevents them from working or participating in other major life activities (5).

Although VR and IL are comprised in the same legislation, they have not always worked in tandem, as VR professionals have tended to view consumers of IL services as being unable to achieve gainful employment and therefore not likely to benefit from the VR program (4). In recent years, however, the goals and principles of VR and IL have begun to converge and are likely to continue to do so, especially as newer understandings of rehabilitation and disability become more widely accepted with the growing international impact of the International Classification of Functioning, Disability and Health (ICF). The ICF, which is described more fully in Chapter 19 of this volume, was endorsed by the World Health Organization (WHO) in 2001 and provides a new framework for understanding health and health-related conditions (6). The ICF conceptual model provides a holistic perspective of health that is consistent with contemporary rehabilitation philosophy, in which disability is seen as a consequence of the interaction of the person with the environment (7). In a further development of the ICF, the Physical and Rehabilitation Medicine section of the European Union of Medical Specialists has recently adopted the ICF conceptual model, in which rehabilitation is understood as a health strategy to help people attain optimal functioning in their interaction with the environment (8,9). Within this model, individuals do not “have” disabilities; rather, they are “people with health conditions experiencing or likely to experience disability” (8). The United Nations Convention on the Rights of Persons with Disabilities has also endorsed an understanding of disability consistent with the ICF model, stating, “disability results from the interaction between persons with impairments and attitudinal and environmental barriers that hinders their full and effective participation in society on an equal basis with others,” including maximum independence and the right to work (10).

For the sake of clarity, terminology such as “persons with disability” is used in this chapter to refer to individuals with a health condition who are experiencing or likely to experience disability. This usage is in no way intended to imply that disability is an attribute of the person; rather, it is understood that disability is a consequence of the interaction between the person with the health condition and the environment.

This chapter provides an overview of both the VR and IL programs in the United States, including the legislative history and purpose of VR and IL services and the differences between these two service paradigms. The authors describe VR program settings and staff, as well as the services that help individuals with disabilities achieve their goals. This chapter also provides a brief review of key research findings that document the effectiveness of VR services and implications for practice.

PURPOSE AND CHARACTERISTICS OF VR AND IL SERVICES

Vocational Rehabilitation

The primary purpose of VR has been to help people with disabilities prepare for and obtain gainful employment, usually through competitive employment (e.g., paid work). VR programs provide rehabilitation services designed to maximize independence and employment and to promote full integration and participation in society (11). Rehabilitation
counselors work with individuals who have a wide range of disabilities. These include physical disabilities, such as spinal cord injury, stroke, arthritis, multiple sclerosis, congenital or orthopedic difficulties, chronic pain, or amputations; cognitive disabilities, such as traumatic brain injury (TBI), organic brain syndromes, developmental and learning disabilities; and psychiatric disorders, including major depression, bipolar disorder, and schizophrenia (5).

VR services may be provided in a variety of settings, including the state-federal VR program (a public agency under the U.S. Department of Education), private, nonprofit community-based programs (e.g., Goodwill Industries, Easter Seals), rehabilitation hospitals, Veterans Administration system, for-profit rehabilitation firms, psychiatric rehabilitation programs, insurance companies, and employer disability management programs (12). Within the state-federal system, the VR process involves a collaborative relationship in which the rehabilitation counselor and individual with a disability work together to identify a feasible vocational goal and the services needed to achieve employment. This process generally involves: (a) individual assessment and planning, which may include interviewing, paper-and-pencil tests, and performance evaluation in real or simulated work situations; (b) comprehensive services, which may include counseling, education, vocational training, physical therapy, speech therapy, and assistive technology (AT); and (c) job placement, which may include on-the-job training or job trials, job development, job search training, supported employment, placement in permanent employment, and postemployment services. Private rehabilitation companies that work with individuals who have work-related disabilities provide or plan services such as vocational assessment, work capacity evaluation, job analysis, work hardening and reconditioning, vocational training, job accommodations, job-seeking skills, job placement, and employer development.

In the state-federal VR program, the service provision plan is formalized with an Individualized Plan for Employment (IPE), which is jointly developed by the individual and counselor. Once job placement has been achieved, follow-up services are continued for a minimum of 90 days to provide support and consultation to the employee and to his or her employer. This helps to ensure that the employment situation is working out satisfactorily for all parties (5).

Independent Living
Living independently with a severe disability in a physical and social world that is often less than accommodating presents a lifetime of challenges. Attention to individual needs also is critical, as the impact of a severe disability may change at different life stages or in varying situations. For example, an individual who is relatively unhampered by disability in one area of life, or during one stage of development, may at another time or under different circumstances be completely overwhelmed by any one of the myriad challenges presented by a severe disability (5). The overriding goal of IL is the full inclusion and participation of individuals with disability in society.
its focus on providing services designed to “fix” the individual, rather than eliminate societal and environmental barriers that magnify or even create disability (15). The IL movement, in contrast, played a significant role in recognizing disability as being the consequence of external barriers, rather than a “problem” within the individual (16). In the IL process, the concept of independence is subject to various definitions, depending on the unique need and desires of the individual. Success is defined through maximizing self-sufficiency to the greatest extent possible for as long as possible, with an emphasis on consumer self-direction (17). The individual may be independent in some life situations but relatively “dependent” in other life situations in terms of the level of services needed. Within IL, self-determination is the guiding principle; autonomy and level of independence may vary, depending on one’s needs, but the individual maintains as much control as possible in decision making (18). For example, a person with quadriplegia may be able to independently perform tasks at work (e.g., by using a voice recognition program on a computer to compensate for upper mobility limitations) but need more extensive assistance from a personal attendant in performing activities of daily living (ADL) at home. Though dependent in ADL, the person maintains self-determination.

In order to be considered eligible for VR services, the consumer must have a disability that presents a significant barrier to employment and have a feasible vocational goal (4, 11). Assessment is an important part of the VR process and is a required initial step to determine if the individual is eligible for services. Once eligibility is determined, assessment is essential to understanding the functional impact of disability on the consumer. Based on results of the assessment, the rehabilitation counselor and consumer identify a vocational objective and begin planning services needed to attain that objective.

IL services, on the other hand, acknowledge the effect of disability on the client, but do not require a thorough analysis of the client, nor the disability, as a prerequisite to the provision of services. In addition, IL services are totally separate from a consumer’s eligibility for VR services. For example, a consumer may not desire to seek employment or VR services but be able to access IL services. The success of IL services depends on the people and resources in the community for direction and support. Consumer involvement is key and is assured through the governance structure of ILCs, which must be managed by persons with disabilities; in addition, the ILCs must ensure that a majority of the staff, including those responsible for decision making, as well as the governing Board, are persons with disabilities (19). IL services are aimed at addressing personal and environmental difficulties. In general, research has outlined the following areas for IL programs to address: self-image, well-being, functional limitations, health behaviors, interpersonal skills, and environmental barriers the system level (e.g., regulations, physical access) and within community (e.g., medical providers, social and family support). The appropriateness of IL services is based on the rights of people with disabilities for dignity, freedom, and control of their destiny. A major emphasis is that services assist with modifying the environment, not the person (4). Table 17-1 highlights some of the differences between VR services in the state-federal system and IL services provided by ILCs.

Consumer sovereignty and empowerment have always been the underpinnings of the IL movement. Persons with disabilities encounter an array of both physical and social discrimination. Empowerment is needed to battle discrimination in housing, employment, education, poverty, and social isolation (15). Empowerment is a form of self-determination where people with disabilities, via advocacy (self or institutional) have a right to determine their destiny. The Rehabilitation Act Amendments of 1998 also formalized consumer choice in the VR process and planning. Consumer sovereignty, sometimes referred as consumer involvement, asserts that people with disabilities can best judge their own interests and should ultimately determine what services are provided to them (16). This
current rise of consumerism directly challenges the traditional service delivery system. There has been a gradual de-emphasis on professional decision making with respect to case planning; accordingly, service provision plans are now drawn up jointly by the individual with the disability along with his or her counselor. Because of the increased awareness created by advocacy skills training at ILCs, many people with disabilities are better informed about their benefits and the regulations of the agencies with which they must deal (20).

**LEGISLATIVE HISTORY**

World Wars I and II as well as an array of legislation since the early 20th century have had a significant impact on VR programs and the IL movement (4). Refer to Table 17-2 for a summary of VR’s legislative history. The United State’s VR program began in 1918 with the passage of the Soldiers’ Rehabilitation Act. The Federal Board for Vocational Education, established in 1917 by the Smith-Hughes Act (PL 65-178), was authorized to create VR programs for veterans with disabilities and the U.S. Department of Labor’s task was to locate employment for these individuals (11). The Smith-Fess Act of 1920, vocational rehabilitation legislation, was then passed to serve civilians with physical disabilities who were either totally or partially incapable of remunerative employment. State-federal fund matching was used for services, which included vocational guidance, vocational education, occupational adjustment, and placement. Although physical restoration was not emphasized, a prosthetic device would be provided if it was necessary for the person with a disability to complete vocational training (4). The 1920 bill had to be reauthorized every few years and

<table>
<thead>
<tr>
<th>Year</th>
<th>Act/Amendment</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1917</td>
<td>Smith-Hughes Act</td>
<td>Established Federal Board for Vocational Education.</td>
</tr>
<tr>
<td>1918</td>
<td>Soldier’s Rehabilitation Act</td>
<td>Created VR programs for disabled veterans.</td>
</tr>
<tr>
<td>1920</td>
<td>Smith-Fess (Civilian Rehabilitation Act)</td>
<td>Established civilian rehabilitation programs.</td>
</tr>
<tr>
<td>1935</td>
<td>Social Security Act</td>
<td>VR became permanent federal program.</td>
</tr>
<tr>
<td>1936</td>
<td>Randolph Sheppard Act</td>
<td>Allowed blind individuals to operate vending stands on federal property.</td>
</tr>
<tr>
<td>1938</td>
<td>Wagner-O’Day Act</td>
<td>Required federal government to purchase products from workshops for the blind.</td>
</tr>
<tr>
<td>1943</td>
<td>Vocational Rehabilitation Act Amendments</td>
<td>Eligibility expanded to include people with emotional disturbance and developmental disabilities; medical services and income maintenance programs.</td>
</tr>
<tr>
<td>1944</td>
<td>The Serviceman’s Readjustment Act</td>
<td>Tuition and stipends for returning WWII veterans.</td>
</tr>
<tr>
<td>1954</td>
<td>Vocational Rehabilitation Act Amendment</td>
<td>Authorized federal funds to build and expand rehabilitation facilities; training grants to educational institutions for rehabilitation professionals.</td>
</tr>
<tr>
<td>1965</td>
<td>Vocational Rehabilitation Act Amendment</td>
<td>Expanded federal-state funding ratio; included “behavior disorder” as new category (but dropped in Rehabilitation Act of 1973).</td>
</tr>
<tr>
<td>1973</td>
<td>Rehabilitation Act, Title V</td>
<td>Creation of Individualized Written Rehabilitation Plan (now IPE) and consumer grievance procedures.</td>
</tr>
<tr>
<td>1974</td>
<td>Rehabilitation Act Amendments</td>
<td>Gave broader definition of “handicapped” emphasizing limitation in major life activities, not only employment.</td>
</tr>
<tr>
<td>1978</td>
<td>Title VII of the Rehabilitation Act Amendments</td>
<td>Comprehensive Services for Independent Living: Authorized grants to organizations receiving federal funds to provide IL services to those with little potential for employment.</td>
</tr>
<tr>
<td>1986</td>
<td>Rehabilitation Act Amendments</td>
<td>Authorized supported employment services to individuals who could not be placed in competitive employment; increased use of rehabilitation engineering services.</td>
</tr>
<tr>
<td>1990</td>
<td>Americans with Disabilities Act</td>
<td>Prohibited discrimination against people with disabilities in employment, public services, public transportation, public accommodation and telecommunication.</td>
</tr>
<tr>
<td>1992</td>
<td>Rehabilitation Act Amendments</td>
<td>Emphasized consumer involvement in the development of rehabilitation plans; mandated state rehabilitation agencies establish Rehabilitation Advisory Councils.</td>
</tr>
<tr>
<td>1996</td>
<td>Telecommunications Act</td>
<td>Mandated that telecommunications services and equipment be designed and fabricated to be accessible to people with disabilities.</td>
</tr>
<tr>
<td>1998</td>
<td>Assistive Technology Act</td>
<td>Provided states funding to develop and expand consumer-responsive technology programs for people with disabilities.</td>
</tr>
<tr>
<td>1998</td>
<td>Workforce Investment Act and Rehabilitation Act Amendments</td>
<td>“One-stop” shopping for employment services; emphasized consumer role in service selection and access to information.</td>
</tr>
<tr>
<td>1999</td>
<td>Ticket to Work Incentive Improvement Act</td>
<td>Provides SSI and SSDI beneficiaries a “ticket” to purchase VR services from employment network of their choosing.</td>
</tr>
<tr>
<td>2001</td>
<td>New Freedom Initiative and Executive Order 13217</td>
<td>Nationwide effort to eliminate barriers to community participation of people with disabilities; six federal agencies directed to review their policies in accordance with new emphasis on community access.</td>
</tr>
<tr>
<td>2008</td>
<td>Americans with Disabilities Amendments Act</td>
<td>Includes major life activities such as bending; major body functions are modified (e.g., digestive, bowel); and mitigating factors, not including glasses, will not be viewed when determining disability status.</td>
</tr>
</tbody>
</table>
consequently was frequently in jeopardy of being discontinued. Fortunately, the groundbreaking Social Security Act of 1935 included unemployment compensation, old age insurance, aid to dependent children, maternal and child health services, as well as other important programs. In addition, the VR program was made permanent so that an act of Congress would be required to dismantle the VR system. Similar to the veterans’ rehabilitation program, federal monitoring of civilian rehabilitation had become the duty of the Federal Board for Vocational Education. The Randolph–Sheppard Act of 1936 and Wagner-O’Day Act of 1938 provided opportunities for individuals with visual impairment to operate vending stands on federal property and required the federal government to purchase certain products from workshops for the blind, respectively. 

Between 1920 and 1943, VR provided services to only those with physical disabilities. The Barden-Lafollette Act of 1943 expanded services to individuals with mental retardation, mental illness, and blindness. World War II was monumental in changing the civilians’ and veteran rehabilitation systems. During the wartime industrial labor shortage, persons with disabilities demonstrated their ability to work. Furthermore, medical advances, such as the development of antibiotic medications, meant that many more military persons were able to survive. In fact, the board of physical medicine was established within the AMA at this time. The New York University medical school, under the leadership of Dr. Howard Rusk, established the first department of physical medicine. The Servicemen’s Readjustment Act (PL 73-346) of 1944, known as the “GI Bill of Rights,” guaranteed up to 4 years of tuition and a stipend for living expenses for returning veterans, whether disabled or not. Between 1943 and 1953, over 600,000 World War II veterans obtained VR services while another 8 million took advantage of the GI Bill.

The Vocational Rehabilitation Act Amendment of 1954 (PL 83-565) laid the groundwork for a tremendous expansion of rehabilitation programs. Important facets of this legislation included authorization for the use of federal funds to build and expand rehabilitation facilities, authorization of training grants to institutions for the education of new rehabilitation professionals, and extensive funding for research and demonstration projects to improve and disseminate knowledge of rehabilitation treatment. This legislation promoted the professionalization of VR by establishing graduate-level training programs throughout the United States. However, this increasing professionalism led to the alienation of many people who became part of the disability rights movement of the 1970s and 1980s. The ratio of federal to state matching funds changed from 50-50 to $3 for every $2 of state funds. The legislators also authorized 30 million dollars to the States for rehabilitation purposes and expanded annual funding. Research demonstration grants allowed state rehabilitation agencies or nonprofit agencies to conduct projects directed specifically at VR. Now, disability arenas such as the psychological, social, and behavioral components of disabilities were studied in a systematic manner. Results were then applied in training programs, policy and rehabilitation service mandates.

The Vocational Rehabilitation Act Amendments of 1965 expanded the federal state funding ratio to 75% to 25% and addressed extended evaluation for individuals with severe disabilities to determine if VR services would be beneficial. The Amendments also extended eligibility to include “behavior disorders,” which made it possible for those with substance abuse problems, public offenders, and those who were socially disadvantaged to obtain VR services. However, this provision was removed in the 1973 Rehabilitation Act due to financial as well as time constraints.

In 1961, legislation was introduced requesting that Vocational Rehabilitation Agencies provide IL services; however, it failed to pass. In 1972, a new bill (HR 8395) was written and passed to replace existing VR legislation by including comprehensive rehabilitation services and IL provisions. This legislation was vetoed by President Nixon, who believed that IL would dilute the resources of the VR program. The bill was resubmitted in 1973 and once again vetoed, as the President’s advisors felt that the rehabilitation of people without vocational potential was too expensive. In 1973, a compromise was reached and the Rehabilitation Act was made into law. Although the IL provisions were eliminated, an emphasis was placed on the delivery of VR services to individuals with severe disabilities.

The Rehabilitation Act of 1973 was a powerful piece of legislation and had a major impact on VR programs. Additional important features of the Act were the creation of the Individualized Written Rehabilitation Plans (now called the Individualized Plan for Employment or IPE) and consumer grievance procedures. These two innovative measures emphasized for the first time the notion of consumer empowerment, with simultaneous changes in language from “client” to “consumer.” Now, clients were seen as consumers, and this change acknowledged that the traditional “paternalistic” attitude of the service provider was a barrier to rehabilitation. Consumers became more assertive than in the past and were empowered to make autonomous decisions with the assistance of trained VR professionals. In this paradigm, people with disabilities work as a team with the rehabilitation counselor, occupational therapist, physical therapist, psychologist, and other medical professionals, signaling a change from the past, when they were expected to be compliant recipients of care.

Consumer advocacy was a powerful driving force on the tenets of the Rehabilitation Act of 1973. The act included a number of provisions to address discrimination and environmental barriers, as follows: (a) Section 501, mandating that the federal government itself practices nondiscrimination in its hiring practice; (b) Section 502, establishing the Architectural and Transportation Compliance Board to enforce accessibility standards for persons with disabilities; (c) Section 503, prohibiting discrimination in the hiring process based on disability status (which applied only to federal contract recipients or subcontractors); and (d) Section 504, in which participation from any federally sponsored program was prohibited for any qualified person with a disability. These programs included schools (elementary, secondary, postsecondary), hospitals, clinics, and welfare agencies. Accessibility of programs was
emphasized. This legislation mandated that employers or institutions receiving federal funds were required to make “reasonable accommodations” for otherwise qualified people with disabilities. For employers, this meant job restructuring, workplace modifications, provision of specialized training, or ongoing support (4).

In 1974, the Rehabilitation Act of 1973 was amended to include a broader definition of the term “handicapped individual.” The new definition emphasized limitations in major life activities rather than only vocational objectives. In 1978 (4,1,23), a number of amendments were made to the Rehabilitation Act of 1973, most importantly, Title VII entitled “Comprehensive Services for Independent Living.” Its purpose was to authorize grants to states to provide independently living services for individuals with disabilities without emphasizing the employment component. A vocational outcome was not necessary for these services. In addition, a bill of rights section was included for those with disabilities, and the National Council on Independent Living was established. This amendment made centers for IL partners with the traditional rehabilitation program. The Rehabilitation Act Amendments of 1986 authorized state VR agencies to provide supported employment services to individuals with severe disabilities who were not capable of competitive employment. Consequently, long-term placement in workshops was viewed as a less than favorable option. This Amendment also increased the use of rehabilitation engineering to assist with independence (4).

The Americans with Disabilities Act (ADA; PL 101-336), a hallmark civil rights legislation for persons with disabilities, was passed in 1990 (15). Its purposes were to end discrimination against people with disabilities and to promote both their social and economic integration (23). To be covered under the ADA, either a mental or physical impairment exists that must “substantially limit” one or more major life activities (e.g., walk, eat, self-care, work, etc.); or the individual has a record of the impairment (e.g., cancer in remission, substance abuse in remission, records of mental illness, educational record depicting learning disability); or is regarded as disabled (e.g., the employer does not hire an individual for a sales position due to a facial birthmark). Contagious diseases (HIV/AIDS) were included in the law, though not in work situations where they may pose a “direct threat” to others. Other conditions such as pedophilia, kleptomania, compulsive gambling, and transvestism were not protected by the ADA. Subsequent court decisions have determined that impairments that are mitigated by corrective measures (mitigating measures) are not covered under the ADA. For example, high blood pressure and diabetes do not substantially limit major life activities when medication regimes are followed. These court decisions have had the effect of creating ambiguity in how the ADA is interpreted, raising concerns that its original antidiscrimination protections are being gradually eroded (4). As a consequence, the ADA Amendments Act was signed into law in September, 2008. Taking effect in January, 2009, the ADA Amendments Act is designed to protect the original ADA of 1990. The Amendments Act retains the ADA’s definition of disability and maintains that the definition of disability should be viewed and interpreted in a “broad” manner. It offers clarification regarding what the law considers to be major life activities as well as mitigating measures. For example, it expands major life activities to include those that had not been specifically identified in the previous legislation, such as reading, bending, and communicating, and adds major bodily functions, such as the immune system and neurological functions. Under the Amendments Act, mitigating measures, not including contacts or eyeglasses, will not be taken into account when determining disability status. An impairment that is either in remission or episodic is considered a disability if a major life activity would be substantially limited when the impairment is in an active state. In addition, one “regarded as” having a disability is not entitled to a reasonable accommodation (24).

The ADA prohibits discrimination against people with disabilities in employment (Title I), public services (Title II), public transportation (Title III), places of public accommodation (Title IV), and telecommunications (Title V). The primary purpose of Title I is to provide equal employment opportunities to qualified individuals with disabilities. Under Title I, the ADA prohibits discrimination because of disability in the hiring, promotion, job training, and firing process (15). A qualified individual with disabilities is one who can perform the essential functions (e.g., word processing skills for a newspaper writer) of a job with or without a reasonable accommodation (e.g., voice-recognition computer software). Businesses with more than 15 employees are required to make reasonable accommodations for qualified candidates with disabilities unless such accommodations would impose “undue hardship.” Such accommodations might include improving worksite accessibility, equipment modification, work schedule modification, or provision of interpreters. Undue hardship is determined by the organization’s financial profile. Title II of the ADA prohibits discrimination against people with disabilities in state and local government programs such as the library, courtrooms, county museum, and public transportation must be accessible. Individuals with disabilities should have equal access to these services, and reasonable accommodations must be provided for these services (4).

Title III of the ADA, the public services provision, prohibits discrimination against persons with disabilities when using public establishments such as theatres, hotels, auditoriums, museums, and private schools. Both barrier removal and reasonable accommodations that are readily achievable (in terms of financial impact) are required. Title III underscores the fundamental right of Americans to enjoy all public accommodations in society (4). The presence of a disability should not preclude one from this freedom. Imagine not being able to use a hotel during a family vacation because there is no ramp to the front door. Title IV is the component of the ADA addressing telecommunication and requires that individuals with hearing and/or speech impediments have equal access to telephone services at a cost comparable to those without disabilities (15). Services include the Telecommunication Device
for the Deaf (a keyboard is used to communicate via telephone lines) or a Telephone Relay Service (in which a third-party operator relays messages from a nonspeaking patron to one who communicates through speech). Lastly, Title V prohibits retaliation or coercion toward individuals who use the ADA to protect their rights.

The Rehabilitation Act Amendments of 1992 emphasized consumer involvement in the policies and procedures of state VR agencies. These amendments emphasized employment outcomes for people with disabilities, commitment to IL services, informed choice, and consumer participation in VR process (4). In addition, Rehabilitation Advisory as well as Independent Living Councils were created and were required to be composed primarily of people with disabilities in order to provide guidance regarding agency policy and procedures. The order of selection process was also established for determining who would get services, with priority of services going to individuals whose disabilities were most severe. In addition, consumer autonomy was stressed. Each individualized rehabilitation plan must describe how the consumer was involved in plan development. Annual reviews of the plan are mandated, and the consumer or guardian must participate in the review and modifications to plans. Eligibility requirements were changed in that it was now assumed that the person could benefit from VR services and achieve an employment outcome, unless it could be shown otherwise (11). Terminology was changed; instead of “handicapped,” the term disabled was now used (15). Other amendment topics included support for transition from school-to-work, supported employment, on the job training, serving minorities, and personal care attendants.

In 1998, President Clinton signed the Workforce Investment Act (WIA; PL 105-220), which included the Rehabilitation Act Amendments (4,11,15). The WIA is based on a “one-stop” concept of locating job training, education and employment services available at a single location. The purpose of the one-stop career system was to address the “fragmentation” in service delivery that previously existed. Programs were streamlined and consolidated to avoid problems that occurred in the past with disjointed services. Now, all services are located in one arena (11). The one-stop shop provided a facility where an array of federally funded programs were available in one location. Programs include unemployment assistance resources, vocational training, placement counseling, vocational assessment, and daycare services for individuals while they utilized the one-stop services. This law emphasized increased consumer control of the vocational planning process. The Individualized Written Rehabilitation Plan was now called the IPE, and clients were expected to be active participants in plan development. Outreach to underserved minorities was also stressed. The law put more emphasis on the need for qualified VR counselors to provide services. A master’s degree in rehabilitation counseling and passing the certified rehabilitation counselor (CRC) examination were methods for establishing this qualified status (4).

In 1999, the Ticket to Work and Work Incentive Improvement Act (TWWIIA; PL 106-170) was passed, providing Supplemental Security Income (SSI) and Social Security Disability Insurance (SSDI) beneficiaries with a “ticket” they may use to obtain VR services from an employment network of their choice. This program is voluntary and consumers may choose their own rehabilitation service vendor, designated as an Employer Network, or EN. Vendors can be either private or public entities. For example, ILCs, state VR agencies, educational institutions, and employment agencies are examples of potential employment networks. The TWWIIA also provided for the removal of work disincentives such as medical coverage. Starting in October 2000, Medicare and Medicaid coverage was expanded to more people with disabilities who were employed. With SSDI, a 9-month work trial (with a 3-month grace period) is allowed, so the person’s benefits are not affected for 12 months. Extended medical coverage through Medicare is provided for up to 93 months following the trial work period for those who are working. Consequently, employment rates can increase if individuals with disabilities do not believe they will lose benefits due to employment. These supplemental medical as well as Social Security benefits help remove important disincentives to employment for individuals with disabilities (4).

On February 1, 2001 President Bush announced the New Freedom Initiative as part of a nationwide effort to remove barriers in community living for people with disabilities. The initiative’s goals are to: (a) increase access through technology; (b) expand educational opportunities for youths with disabilities; (c) promote home ownership; (d) integrate Americans with disabilities into the workforce; (e) expand transportation options for persons with disabilities; and (f) promote full access to community life. Thus, this comprehensive plan’s mission is to ensure that every American has the opportunity to learn and develop skills, engage in meaningful work, have autonomy in daily activities, and participate in community life. The president emphasized that individuals with disabilities, families, caregivers, and advocates were involved in a listening session with the purpose of improving community integration (25).

**THE CONSUMER MOVEMENT IN REHABILITATION**

As the consumer (instead of patient/client) movement grew in the late 1960s to early 1970s, with its emphasis on advocacy and self-determination, the time was ready for the IL and Disability Rights movements. The driving force of IL was that individuals with severe disabilities were able to both direct and manage their lives. Furthermore, the services and supports that people with disabilities need are best delivered by individuals who themselves have disabilities and whose knowledge about both disability and services is derived from firsthand experience. Empathy is a cornerstone of this model, with consumers serving and providing mentoring to other consumers. In this model, an individual with a disability may address the many concerns that a person with a newly acquired disability may encounter. Consumers may discuss adjustment to disability issues, adaptive equipment, and techniques for navigating
the many service systems (4). The model follows that of organizations such as Alcoholics Anonymous, in which substance use disorders are best addressed by those who have gone through a similar path.

**Empowerment of Consumers**

Disability legislation continues to emphasize the fundamental right of individuals with disabilities to make decisions about their own lives (4). Since 1973, rehabilitation practice and legislation have been driven by consumer-counselor partnerships and consumer empowerment, which have been the underpinning of all rehabilitation legislation, as exemplified by the following:

- The 1978 Rehabilitation Act Amendments provided that individuals with disabilities be guaranteed more involvement in their rehabilitation plan;
- The 1986 Rehabilitation Act Amendments included support for individual consumer rights and revised the IPE format to include consumers’ statements of their own rehabilitation goals;
- The 1990 ADA further strengthened consumer self-determination by ensuring rights in the areas of employment, transportation, public services, and public accommodations;
- The 1992 Amendments to the Rehabilitation Act also supported the movement toward self-determination.

Consumer empowerment is the driving force behind our current rehabilitation policy and practice paradigm. Since its inception, the IL movement has viewed consumer control as a necessary component of rehabilitation and direction, not only of the services needed by its individual constituents but also of the institutions and organizations that house its activities and administer its resources (4,15). Thus, the 1992 Amendments to the Rehabilitation Act required that a majority of ILC staff, management, and directors must be individuals with disabilities. This empowerment perspective sees the consumer as a competent change agent capable of identifying problems, finding solutions, and making independent decisions. Within this paradigm, the rehabilitation counselor serves as a guide or facilitator to problem solving and decision making. Knowledge is viewed as power, and the counselor encourages the consumer to gather information to enhance knowledge. The rehabilitation counselor practices informed consent and continually drives the point that the consumer is a competent agent in charge of his or her change (15).

**Consumerism, Accessibility, and Assistive Technology**

Social, environmental, and cultural barriers have long kept people with disabilities from having full access to work, education, and participation in their communities. Technological advances have been recognized as a means of expanding opportunities for people with disabilities, helping them to surmount environmental barriers, enhance functional capacities, and maximize independence (26,27). Use of appropriate technology can improve the mobility, communication, and IL skills of people with disabilities, thus dramatically enhancing their ability to become full participants in American society. To this end, the Technology-Related Assistance for Individuals with Disabilities Act was passed in 1988 to promote the use of technology services for people with disabilities (26). The Tech Act officially defined AT as “any item, piece of equipment or product system, whether acquired commercially off the shelf, modified or customized, that is used to increase, maintain, or improve functional capabilities of individuals with disabilities” (26). AT can range from “low-tech” devices such as walkers or canes to “high-tech” devices such as speech synthesizers or stair-climbing wheelchairs. This law was replaced by the Assistive Technology Act of 1998 (often called the "Tech Act"), which advocated statewide coordination and utilization of AT services and development of consumer-responsive technology programs (4). The 2004 amendments to the Tech Act promoted improved AT services through a standardized, statewide implementation process, which included plans for providing AT training to rehabilitation counselors (28).

Other laws affecting the lives of individuals with disabilities have also included technology as an essential component. For example, the Telecommunications Act of 1996 (PL 104-104) addressed the civil rights of people with disabilities by mandating that telecommunications services and equipment be “designed, developed, and fabricated to be accessible to and usable by individuals with disabilities, if readily achievable” (29,30). AT helps fulfill the legislative mandate of the Rehabilitation Act of 1973 and its subsequent 1986 and 1992 Amendments, which reinforce the need for active consumer participation in the VR process and emphasize consumer choice, self-determination, and consumer empowerment (31). Consistent with the emphasis on consumer involvement, the 1998 Amendments to the Rehabilitation Act mandate consideration of AT in provision of VR services (32). The importance of AT has also been increasingly recognized in provision of special education services to children with disabilities, as evidenced by the 1986 and 1990 amendments to the Individuals with Disabilities Education Act (IDEA), which mandated the inclusion of AT devices and services in education, and the 1997 Amendments to IDEA strengthened the role of AT in special education by requiring that AT services be considered for students who have an Individualized Education Program (IEP) (33). The ADA expanded public awareness of physical barriers and established requirements for uniform, national accessibility standards. Consistent with the mandate of the ADA, the Architectural and Transportation Barriers Compliance Board issued guidelines for accessible design in 1991, which were modified and adopted by the U.S. Department of Justice and became enforceable ADA Standards for Accessible Design (34).

While AT has significantly expanded job opportunities and improved the quality of life for many individuals with disabilities, the focus of AT services remains primarily on the individual, who must select and learn to use the assistive device. In contrast, the concept of "universal design" refers to
the design of products to be useable by all people, as much as possible, without the need for individual adaptation or specialized design. An example of universal design are pens and pencils with large rubber grips to minimize hand and wrist discomfort, products which are readily available in stores. Universal design and AT can be thought of as existing along a continuum, with universal design allowing access into the mainstream and AT meeting specific needs of the individual, with some overlap in products that can be either universal in design or AT (35).

AT’s impact on the lives of people with disabilities is not limited to technology itself; it has also empowered consumers by enabling them to become full participants in selecting and making decisions about AT services and devices. AT has also made a significant contribution to current VR philosophy, helping to expand its view from normalization (where the person with a disability must adapt to becoming as “nondisabled” as possible) to an empowerment model, where individuals make their own decisions and enjoy full integration within the communities where they live and work (27).

**VR SERVICE SYSTEMS**

Both VR and IL services foster consumer-driven services and are provided by systems of agencies and organizations that are often connected by common goals and principles, as well as funding mechanisms. The section below describes these systems in greater detail.

**VR Systems**

The state-federal VR program has been a major provider of VR services since the enactment of federal legislation almost 90 years ago. Private, for-profit rehabilitation firms also provide VR services, as do private, nonprofit community-based rehabilitation programs. Another important VR provider is the Veterans Administration, which offers VR services to veterans with service-connected disabilities (36). Increasingly, ILCs are providing VR services, many in coordination with state VR agencies (5).

**Rehabilitation Counselors**

Although many ILCs require that those providing services have firsthand experience with disability, rehabilitation counselors hired by state agencies or private rehabilitation firms are formally trained to provide counseling to people with disabilities. The rehabilitation counselor usually has a master’s degree in rehabilitation counseling, although graduate degrees in guidance and counseling, or social work, are also common. However, rehabilitation counseling is unique among other counseling professions and is a specialization within the rehabilitation field (37). A rehabilitation counselor may be certified through the Commission on Rehabilitation Counselor Certification (CRCC) or licensed through his or her state. Although licensure or certification is not a standard requirement, state agencies may require it. The CRCC has defined the scope of practice for rehabilitation counseling as being a process that helps individuals with diverse disabilities “to achieve their personal, career, and IL goals in the most integrated setting possible through the application of the counseling process” (37). The rehabilitation counselor may employ a broad range of strategies in this process, including assessment; diagnosis and treatment planning; career counseling; individual and/or group counseling, with an emphasis on helping individuals adjust to the impact of disability; coordination of services; methods to eliminate environmental and/or attitudinal barriers; job analysis, job development, and placement services; and assistance with job accommodations.

**State Vocational Rehabilitation Agencies**

State VR agencies operate under federal legislation, the Rehabilitation Act Amendments, and are available in every state to individuals with disabilities. State agencies work with multiple community partners to provide outreach to individuals, as well as to implement services. These partners can include school systems, ILCs, community mental health agencies, hospitals and health care clinics, substance abuse centers, local support groups, other state and county employment programs, and a host of social service agencies (5). State VR agencies may also contract with private rehabilitation firms or with community-based rehabilitation programs for an array of services, including vocational evaluation, job placement services, supported employment, and/or job coaching.

In the state VR agency, rehabilitation counselors and consumers work together in partnership to develop a rehabilitation plan, the IPE. Mandated by the Rehabilitation Act of 1973, the IPE includes (a) a specific vocational goal and timelines for its achievement; (b) services necessary to achieve the goal; (c) secondary goals needed for goal achievement; (d) services and providers required for the secondary goals; (e) criteria and procedures for evaluating progress; (f) counselor and consumer responsibilities; and (g) an annual review for as long as the case is open (4,5). Objectives in the plan might include provision of a service such as treatment, training (e.g., college, on-the-job training), job placement, and specialized adaptive equipment or transportation.

**Private Rehabilitation**

Private rehabilitation generally refers to services provided by rehabilitation practitioners employed by private firms that bill a third-party payer for their services (38). In the late 1970s and 1980s, there was a tremendous growth in the number of private, for-profit rehabilitation firms, largely due to state Workers Compensation legislation designed to foster rehabilitation of injured workers (39). As a consequence of this legislation, private rehabilitation firms emerged to meet the needs of insurance carriers and employers, who needed timely and efficient services (40,41). In addition, insurance companies were becoming increasingly concerned with controlling growing health care costs; they discovered that referring insured or ill workers early in the rehabilitation process could reduce their disability payments. Private rehabilitation firms tend to be
more efficient in their approach to helping individuals return to work than their state VR counterparts. Rehabilitation counselors in the private sector have smaller caseloads than counselors in state VR agencies, and they focus primarily on vocational guidance and placement. Counselors working in the private sector require basic business skills, knowledge of the insurance industry, an understanding of the worker’s compensation system, expertise in legal and medical case management, and the ability to provide vocational expert testimony (42). Many rehabilitation counselors in the private sector are self-employed or co-owners of small firms (43), although there has been a trend in recent years for smaller firms to merge with larger managed care organizations, insurance companies, and other health care entities. Private rehabilitation counselors often contract with insurance carriers to provide VR services to individuals who qualify for these services through workers compensation, auto no-fault or long-term disability policies, and their services are provided within the insurers’ parameters. Consequently, private rehabilitation firms are usually very responsive to the payer’s goals and objectives. Because worker’s compensation laws are governed by each state, rather than federal law, provision of VR services will vary among different states (5).

VR Services
Although services vary according to individual needs, the VR process consists primarily of four stages, which generally occur in sequence: (a) evaluation, (b) planning, (c) treatment (provision of services), and (d) termination, or achievement of the vocational goal, usually placement in a job. The rehabilitation counselor generally begins the evaluation phase by means of an initial interview with the consumer to obtain information, as well as to provide information to assist the consumer in making an informed choice. During the evaluation phase, additional information may be needed prior to rehabilitation planning and obtained through medical evaluation, psychological evaluation, vocational assessment, or related diagnostic services (4). Cultivating relationships with local employers to promote employment of people with disabilities also is an important component of the VR process. The following are services commonly offered by VR agencies during the VR process.

Vocational Assessment
VR begins with assessment of the consumer’s vocational interests, abilities, and vocational potential. By providing information about an individual’s strengths and weaknesses and identifying needed services, the vocational assessment helps the consumer and VR counselor set goals and establish a rehabilitation plan. Assessment may also be used to determine the consumer’s potential to benefit from VR. Vocational assessment can help the counselor and consumer answer the following questions: (a) can you return to your former occupation? (b) would you be able to return to your previous occupation with job accommodations and AT? (c) which of your skills may be transferable to another occupation? and (d) what training or other services would help you become successfully employed?

Assessment is a multidisciplinary process that initially involves gathering data from a variety of sources, such as work and educational history, as well as medical records. For individuals who have received physical rehabilitation services, these records may include a medical examination and reports from the rehabilitation team, such as occupational and physical therapy. The VR counselor may also refer the consumer for additional assessment services, such as a psychological evaluation or a vocational evaluation. The psychological evaluation can provide helpful information regarding the individual’s learning abilities, coping skills, and personality characteristics, while the vocational evaluation describes important work-related behaviors, capabilities, and interests.

Although the terms vocational evaluation and vocational assessment are often used interchangeably, vocational assessment is a general term that includes many different forms of evaluation. Vocational evaluation is defined specifically as a comprehensive assessment that utilizes a variety of tools, including paper-and-pencil tests, structured and unstructured interviews, and real or simulated work (44). With its focus on work-related abilities, a vocational evaluation may use work samples, situational assessments, and on-the-job evaluations. The work sample approach to measurement has been used most often in vocational evaluation, often through commercial work sample systems (e.g., VALPAR) designed to simulate specific jobs or a cluster of jobs; a vocational evaluation may also incorporate actual work samples from industry.

Using a variety of evaluation procedures can help verify other assessment data and thereby contribute to more accurate findings and recommendations (44). Cross-validating assessment data may be especially important for persons with disabilities, as concerns have been raised about the validity of paper-and-pencil tests for this population, particularly for individuals with severe disabilities (45). In addition, vocational evaluation may incorporate AT during the assessment process, for example, by modifying work samples or other test instruments, and by making recommendations that include consideration of specific AT devices that would maximize the consumer’s vocational potential (46). For example, the vocational evaluator may recommend speech recognition software to enable a consumer with limited hand function to perform a job requiring computer access.

Evaluation of an individual’s functional skills is an important part of the assessment process and can be applied to various domains, such as IL skills, interpersonal communication, sensory awareness, emotional stability, learning ability, and stamina. Many functional skills can be observed as part of a situational assessment or on-the-job evaluation, as well as through more formal methods, such as work samples and inventories. Sophisticated systems designed to measure a wide range of physical capacities such as lifting strength are also available, including the ERGOS Work Simulator and the BTE Simulator (47,48). These systems are more likely to be used as part of a Functional Capabilities Evaluation (FCE) conducted in a work conditioning program, hospital, or clinic, rather than during a vocational evaluation.
Job Placement

Job placement includes an array of services, such as job-seeking skills training, direct placement, job accommodation, supported employment, and job development. These services may be offered by rehabilitation counselors or by placement professionals (often called placement specialists or employment specialists). Job analysis is sometimes provided as a service in conjunction with job placement, though it may also occur as part of a vocational assessment. Employment readiness is an important consideration and may be affected by psychosocial factors, availability of accessible transportation, and medical status, including stability of the medical condition and capacities for stamina and endurance. Psychological readiness to work, including motivation, self-esteem, and coping resources also must be addressed.

Once the individual is ready to seek employment, he or she must develop or refine the employment skills that will be required for success in the job search, which will largely be determined by past experience. For example, the individual who acquired a disability later in life and who has a long work history will likely need to refine job-seeking skills that include addressing disability issues with potential employers. The individual with virtually no work history will require far more extensive job seeking skills training, which typically includes identifying and following up on job leads, resume writing, application completion, and interviewing skills (5).

Job placement activities can be viewed along a continuum, ranging from self-placement (usually called client-centered placement) to the counselor assuming all placement responsibility, often referred to as selective placement (49). The skills and personality traits of the job seeker, the nature of the disability, local labor market conditions and opportunities, and even luck influence the extent of counselor involvement that may be needed. The rehabilitation plan will target the consumer’s job goal, specify acceptable geographic and environmental criteria, consider the types of job accommodations needed, and specify follow-up and support services required. A primary role of the VR counselor or placement professional is to assist the consumer in developing job seeking skills, using such tools as coaching and role playing or videotaping a mock job interview. In addition to job seeking skills, he or she must also know how to respond to questions about disability, either on an application or in an interview. Knowledge of legal protections, such as those stipulated by the ADA, is critical for the individual with a disability who is seeking employment (5).

Consumers may also need help with identifying and learning how to request job accommodations. Although the individual must learn what types of accommodations are required, the VR counselor or placement specialist can often act as a consultant to the employer and help negotiate the accommodation process. As stipulated by the ADA, the employer is required to make reasonable accommodations, but this term is subject to various definitions. Some employers view job accommodations as prohibitively expensive, although the majority of accommodations cost less than $500 (50). In fact, job accommodations can reduce worker’s compensation and other costs. Some accommodations may be as simple as a rearrangement of equipment. For example, for individuals using wheelchairs, a height adjustable desk, a voice activated speakerphone, or moving office supplies to accessible drawers are low cost and simple accommodations. Other examples of accommodations include job restructuring (e.g., trading off job tasks with other workers), flexible schedules, large print, allowing use of personal care attendants or service animals, and large-button phones.

Supported employment is another model of job placement that is particularly effective for individuals with severe disabilities, such as TBI, developmental disabilities, or significant cognitive limitations. In contrast to the traditional model of job placement, which assumes that the individual is job-ready and has been sufficiently trained before obtaining a job, supported employment uses a “place-train” approach. That is, the individual is placed in a regular, competitive job setting and given needed support, such as assistance from a job coach and/or work mentor, to provide training on the job and ensure success (51).

Once the consumer obtains employment, follow-up services with both consumer and employer help to ensure a successful outcome and are an important step in the job placement process. The rehabilitation counselor or placement professional may need to intervene either with the consumer, employer, or both, to solve problems as they occur. Further job accommodations may be needed to help resolve difficulties with performance. However, if the consumer is ultimately unable to perform the job, the consumer and counselor may agree that it would be best to discontinue the job, as VR involves a successful outcome for both the consumer and the employer. Placement professionals must establish positive relationships with employers, and ultimate employment success may require provision of long-term follow-up services (52).

Job development is also an important component of job placement services. People with disabilities continue to encounter discrimination from employers, who may object to hiring individuals with disabilities due to common employer “myths” about individuals with disabilities, such as a belief that they will increase their insurance rates, will be subject to accidents, or have problems with job performance or stability. Thus, rehabilitation counselors and placement professionals need to dispel these misconceptions and to actively promote the benefits to employers of hiring individuals with disabilities. Rehabilitation professionals must be visible and active in business organizations and build credibility within the community. Services that can be marketed include recruitment and referral of qualified applicants, consultant services (e.g., ADA compliance, job accommodation, and disability awareness training programs), employee assistance programming, and support and follow-up services (52). Many placement programs also develop business advisory boards constituted of employers and other important business contacts in their community to share ideas and develop strategies for improving employment opportunities for people with disabilities (53).

Job analysis can be critical to the ultimate success of job placement efforts. A thorough analysis of a particular job
EMPLOYMENT OUTCOMES: A REVIEW OF FINDINGS

An overview of rehabilitation studies on employment, VR, and IL outcomes among people with physical disabilities suggests a number of variables that may assist rehabilitation professionals in identifying those individuals who can best benefit from their services. This research can help rehabilitation professionals better understand both environmental and personal characteristics that promote or hinder employment outcomes for individuals with disabilities. Equipped with these research findings, professionals can design programs and/or implement specialized techniques and knowledge to help persons with disabilities be successful in the rehabilitation arena (4).

Several studies have reported certain demographic factors such as race, gender, and disability severity, and disability characteristics as factors related to vocational outcomes (55–58). In general, those with successful vocational outcomes tend to be Caucasian, young adults, well-educated, and had a successful career prior to the acquired disability (59–61). However, because consumer variables are easier to identify and measure than environmental variables, these research findings may at least partly reflect a lack of clear-cut strategies for assessing the impact of environmental factors on VR outcomes (62). As more research is conducted based on the conceptual model of the International Classification of Functioning, Disability, and Health (ICF) discussed earlier, environmental facilitators and barriers can be specifically assessed for their impact on rehabilitation outcomes.

Beveridge and Fabian (55) determined that when consumers’ IPE is congruent with the employment outcome, they are more likely to earn higher wages. Although this finding was applicable to all consumers, it was particularly evident for those with cognitive disabilities. In addition, individuals with physical disabilities earned substantially more than persons with sensory and mental disabilities.

Age at the time of disability, premorbid employment status, work status, and psychological distress were also found to be significant predictors of successful employment for a group of people with TBI. In fact, learning as well as psychological disabilities coupled with TBI tended to have a negative impact on employment outcomes for this population (57,63). Individuals who are injured either at a very young age or over age 40 are less likely to return to work (58,64).

Psychosocial factors also appear to play an important role in employment outcomes for individuals with disabilities. For example, consumers’ self-awareness is a key factor in returning to employment (65). Poor self-awareness, as well as unrealistic goals, are primary reasons why consumers fail to secure employment following provision of VR services. Motivation appears to predict employment status among an array of disability groups (59–61). Internal locus of control is frequently discussed as an important predictor of job search and employment success. This proactive stance tends to produce more successful outcomes compared to consumers who perceive themselves as being more externally controlled (66,67).

In a landmark study of the outcomes and effectiveness of the state-federal VR system, the Longitudinal Study of the Vocational Rehabilitation Services Program (LSVRSP), used longitudinal design to track consumers’ progress over a period of 3 years (68). Of those participants who received services, 69% obtained an employment outcome, which is defined as either competitive or noncompetitive employment over a continuous 90-day period. Of those consumers with an employment outcome, 75% obtained competitive employment, while 25% received noncompetitive placements, which included supported employment, extended employment, homemaker, and unpaid family worker. The study identified a number of consumer characteristics, including type of disability, which had a significant impact on achievement of an employment outcome. For example, consumers with a hearing or orthopedic impairment were more likely to attain an employment outcome. Other characteristics associated with an employment outcome were working at time of application, higher number of dependents, and higher self-esteem. In contrast, individuals receiving disability benefits (e.g., SSI, SSDI) were less likely to obtain employment. The LSVRSP also examined the relationship of VR services to employment outcome and found that provision of job placement, on-the-job training, and business/vocational training were significant predictors of employment. That is, consumers who had received these services were more likely to be employed at follow-up 3 years later.

CONCLUSION

Historically, VR and IL services often seemed to have opposing goals, with the early developers of IL seeing VR professionals as
creating more barriers for individuals with disabilities through their paternalistic attitudes and origins in the medical model of disability. The consumer movement has been successful in effecting changes in policies, driving legislative mandates that have promoted consumer empowerment and facilitated the gradual merging of the goals of both VR and IL. Both VR and IL can be seen as serving essential and complementary roles in the rehabilitation process. VR and IL increasingly share a more holistic, ecological understanding of disability that stresses the role of the environment in either magnifying or minimizing the impact of impairment. In VR, focus has been shifting from efforts to change the person with the disability to removing or modifying barriers in the environment that hinder the person from being an active participant in work and society. These programs also complement each other in many ways. With their goal-oriented approach to services, VR programs assist consumers in obtaining employment or furthering their education to expand future career opportunities. IL programs foster greater sense of accomplishment and control over one’s own destiny. While promoting self-direction and self-sufficiency, they expand life opportunities and social identity for individuals with disabilities. By promoting independence, IL facilitates vocational options and thus can be seen to work in tandem with VR services.

In the past, the roles of VR and IL professionals also differed significantly but now are becoming more similar. Both rehabilitation and IL specialists are change agents who act as facilitators to help people with disabilities maximize their independence, participate as fully as possible in their communities, and advocate for their own needs and interests. However, VR and IL differ in the types of outcomes that each program emphasizes. VR programs define a successful outcome as employment, and preferably paid, competitive employment. Such an outcome is measurable, and outcome data from VR services can be gathered and statistically analyzed to assess its overall performance. At the same time, using employment as the sole criteria for success provides a narrow view of rehabilitation (69). In contrast, a successful outcome in IL is more individualized and broadly defined, as it is based on each consumer’s needs and desires, and usually pertains to progress in areas that are more difficult to pinpoint and measure, such as quality of life, coping skills, autonomy, and self-dignity. The ICF model offers a multidimensional perspective of health, in which persons with health conditions may or may not experience disability, depending on their interaction with the environment; this holds promise as a conceptual framework that could help unify VR and IL even further. For example, in the VR system, criteria based on the ICF could be used to measure indicators that reflect progress in other important aspects of life and more precisely assess program effectiveness, rather than relying solely on employment outcome (62). Such indicators could align closely with consumer goals of IL, such as improved quality of life. In addition, ICF-based criteria could include environmental factors that have long been recognized by the IL movement as being critical in the lives of individuals who experience or are likely to experience disability.

REFERENCES

INTRODUCTION

Mrs. X and Mrs. Y are both 62 years old. One year ago, both had right middle cerebral artery infarctions with resultant left upper- and lower-extremity weakness and left-sided visual inattention. The hospital course for both included admission to an acute care hospital followed by a 3-week stay at an acute rehabilitation hospital. Both were discharged home with home services. Mrs. X and Mrs. Y have supportive spouses who live with them and have children who live nearby. Following their strokes, Mrs. X and Mrs. Y retired. Both are able to walk unlimited distances in the home and community settings with a standard cane, though both indicate they walk slower and shorter distances than before the stroke. Both have discontinued driving. Both have gross grip strength of the left upper extremity but are unable to perform fine motor activities. Active shoulder range of motion is about 30% in the left upper extremity.

Mrs. X lives in a two-story townhouse in a highly populated urban area. A post office, grocery store, two community parks, the church she regularly attends, and a bus stop are located within five blocks of her home. There are ample level well-maintained sidewalks that are wide and well lit. Benches are scattered throughout the sidewalks, buses service the area frequently, and public transportation services for persons with disabilities are available. Mrs. X’s activities include visiting friends and family almost daily, volunteering to teach English classes at her church once a week, and participating in a book-of-the-month club that meets at her neighbor’s house, which she has just joined. When the weather is nice, she takes walks in the nearby park. Mrs. X has stopped grocery shopping since having her stroke and instead uses an online service that delivers the groceries to her home. Mrs. X generally gets around her community several times a week either by walking or by getting a ride with her friends and seems to have a fairly high degree of participation in her complex daily activities.

By contrast, Mrs. Y lives in a single level home in a suburban community. The town center is 2 miles away and contains a post office, library, and town hall with parking several hundred feet from the main entrances. Most of these buildings have stairs at the front entrance and a ramped access further from the stairs. There is a small grocery store a half mile from her home but they do not deliver groceries and a larger one 4 miles away. Before her stroke, she worked in the local school system and exercised at the local YMCA that was located 10 miles from her home. Most of her friends are scattered around town, particularly near the YMCA. The church Mrs. Y regularly attended prior to her stroke is 5 miles from her home.

Many of the sidewalks in Mrs. Y’s communities are uneven and are poorly maintained. A public transportation service does exist; however, the bus only comes twice a day. Many of the buildings have ramp access that is difficult to get to, stairs that have one railing, and limited parking. Mrs. Y has significantly curtailed her social activities since having her stroke, but she is still able to take care of the inside of her home. This includes vacuuming, preparing hot meals, and shopping for estate jewelry on e-bay. Mrs. Y was considering volunteering at a local bookstore in town; however, coordinating with her husband’s schedule proved to be too difficult. Mrs. Y is quite active in her home and often entertains friends and family in her home, but her community mobility is very limited and she has a rather low participation in her complex daily community activities (Fig. 18-1).

Despite the identical pathology, impairments, and functional limitations, the community participation levels of Mrs. X and Mrs. Y are very different. Mrs. X continued to participate in community activities; however, Mrs. Y did not. Medical providers are keenly aware of this paradox, but why does it exist? Is it even something that rehabilitation professionals should consider? If rehabilitation professionals should address this, how do we intervene to optimize community participation and the community-environment interaction? How do we assess the community-environment interaction?

This chapter reviews notions of how the environment impacts participation among adults with mobility limitations. We will discuss theoretical and conceptual frameworks linking the environment to participation, discuss methodological considerations regarding assessment of participation and the community environment, review the literature linking the community environment to participation among adults with mobility limitations, and, lastly, discuss several clinical implications.
implications in this area of rehabilitation science. Our focus of this review is on the physical aspects of the environment. Elements of the social, political, attitudinal, or technological aspects of the environment are not reviewed in this chapter.

THEORETICAL AND CONCEPTUAL CONSIDERATIONS

Community Participation

"Participation" is a term used in the World Health Organization's International Classification of Functioning, Disability, and Health (ICF) and is defined as "a person's involvement in a life situation." The concept of "life situation" pertains to areas in which people attain and define roles such as seeking and maintaining employment, seeking and securing an education, managing a household, and being a care provider. Nine domains are listed in the ICF that pertain to the activities and participations of people: learning and applying knowledge; general tasks and demands; communication; mobility; self-care; domestic life; interpersonal interactions and relationships; major life areas; and community, social, and civic life (1). Thus, the domain of participation addresses the extent to which and how people engage in complex day-to-day activities that comprise their daily lives. Fougeyrollas et al. (2) describe a similar concept "life habits" and define life habits as "daily activities..."
and social roles that ensure the survival and development of a person in society throughout his or her life."

The term "participation" is similar to the terms handicap, defined as "a disadvantage for an individual that limits or prevents the fulfillment of a role that is normal for that individual (3)," and disability, defined as "restrictions in personal and social role behaviors (4,5)." The concept that all these three terms are referring to encompasses the extent to which people engage in the complex daily activities that comprise the roles and activities of their daily lives. These broad role activities occur in the home and community setting. For the purposes of this chapter, however, we focus on community daily activities (community participation)—that is, activities that people do in their community when they leave their home—rather than on home activities.

Performance of complex daily activities is likely to be influenced by a multitude of factors. Managing one's health, for example, entails scheduling an appointment during a time that is compatible for the patient, which is dependent upon the schedule and ability of the patient or his or her caregivers. A doctor's visit entails securing transportation to the facility, which may entail public transportation, adaptive transportation, or the person driving himself or herself. Whether the patient takes public transportation or drives, he or she needs to navigate the physical premises of a parking garage and/or office building and arrive at the (correct) location of the appointment. Once at the correct location, the patient needs to navigate waiting rooms and small examining rooms, change clothing if necessary, may need to transfer onto equipment that is not adaptive, and pay for the service. Then, the patient needs to reverse all of these to get back to the origination point! "Going to the doctor's office" thus comprises a number of activities including transferring, walking or propelling a wheelchair, driving or securing transportation, dressing, managing money or billing payment, and coordinating a calendar. These factors in turn are likely to be influenced by a person's level of function, cognitive status, sensory systems and emotional status, the availability of social supports and other programs in the environment, and the presence of environmental obstacles or barriers that impede a patient's ability to get to the doctor's office.

The ICF model provides a conceptual framework to examine the complex processes that are involved in the above example. Conceptually, the model proposes that participation is influenced by an individual's body structure and function (defined as anatomical parts and physiological functions of organs and body systems) and his or her activity levels (defined as the execution of a task or action by the individual) (1). Other disablement theoretical models suggest similar relationships among body structures (impairments), functional activity (limitation in function), and performance of role behaviors (disability) (4,5). Conceptual frameworks help clinicians and researchers organize the complicated array of factors that influence health and health problems.

The ICF, however, also functions as a structured classification system of health and disability outcomes. The ICF provides a detailed description of relevant items in body systems, functions, activities, and participation. Thus, items from assessment instruments can be "linked" to the respective ICF identifiers in respective domains. This type of classification system and linking strategy can be useful to clinicians and researchers. For clinicians, it can help guide the assessment of clinically relevant factors in various patient populations (including the selection of measurement instruments). For researchers, it can aid in understanding the conceptual underpinnings of instruments to guide selection of the best instrument for the factor of interest to study. Indeed, an extensive amount of research on developing and validating the ICF classification structure is currently underway, with rules for linking instruments to ICF classification categories developed and validated (6).

The example described previously in this section, however, does not solely entail elements of body systems, functions, activities, or participation. Rather, this scenario is likely to be influenced by factors outside the individual (e.g., the environment) as well as internal behavioral or motivational factors. Environmental and intra-individual behavioral and personal factors in the ICF framework are proposed to be contextual factors that modify the relationships among body systems, functions, activities, and participation (1) (Fig. 18-2). These factors represent the social and physical circumstances in which the person lives—that is, characteristics of the person (e.g., attitudes) or elements of the environment (1,7–14). The role of the environment in disablement is also emphasized in the Institute of Medicine's 1997 report that states a person's involvement in life activities results from the interaction of the person within the environment in which the person exists. Limitations occur when the environment does not fit the individual (12,13). Exactly how, however, would the environment impact participation among persons with functional limitations and disabilities? What are the important elements of the environment that need to be considered from a clinical perspective? To understand these questions, we need to understand the conceptual underpinnings of environmental assessment.

**Conceptual Notions of the Environment-Participation Interaction**

The environment may be viewed as everything external to the individual and is believed to have an impact on people's behaviors and health outcomes. Researchers in the fields of

![Figure 18-2. International Classification of Functioning, Disability, and Health.](image-url)
transportation and public health have shown that housing density, land use diversity, and street connectivity and design are related to how people navigate their community (15–17). People who live in environments with greater residential to commercial mixes have higher physical activity levels compared to those persons who live in environments that are more residential (18,19–21). People who live closer to work establishments are more likely to walk to work, thereby having higher levels of physical activity (19). Generally, environments that are characterized as more “walkable” are associated with more physical activity patterns among their residents as compared to environments that are less “walkable” (20,22). Research linking environmental factors to transportation and physical activity sheds some interesting perspectives into environment-participation research. Will these “global” factors of land use mix and street connectivity have a role in determining participation levels of people with functional limitations and disabilities?

Participation in daily life activities is a much more complex phenomenon than engaging in physical activity or navigating one’s community. Which aspects of the environment impact participation? What are the characteristics of the environment-participation interaction? There are several conceptual frameworks in the field of rehabilitation that suggest answers to these questions. Fougeyrollas et al., in their early conceptual notions on how the environment was related to participation, defined “environment” as all facilitators and barriers external to the individual that influence participation (8). They suggested that the environment has different “levels” of influence: “micro personal,” “meso community,” and “macro societal” and identified four domains that were important for participation: (a) socioeconomic organization (e.g., family structure, political systems, and economic systems), (b) social roles (e.g., law, values, and attitudes), (c) nature (e.g., geography, climate, and time), and (d) development (e.g., architecture, land development, and technology) (8,23).

In the ICF, the “environment” is identified as a contextual factor that has positive and negative attributes that could either enhance or restrict ones’ level of involvement in life activities—that is, the environment is the “background” in which people’s life situations are embedded. The ICF organizes environmental factors into two levels: (a) individual and (b) societal. The individual level pertains to the immediate environment of the individual including settings such as home, workplace, and school. Physical and material features of the environment as well as direct contacts with people in the environment are included in this level. The societal level includes the formal and informal social structures, services, and overarching approaches or systems in the community or society that have an impact on individuals. Organizations and services related to work environment, community activities, government agencies, communication and transportation services, and informal social networks as well as laws, regulations, formal and informal rules, attitudes, and ideologies are included in this level. The ICF organizes environmental factors into five domains: (a) products and technology; (b) natural environment and human-made changes; (c) support and relationships; (d) attitudes; and (e) services, systems, and policies (1) (Fig. 18-3; Table 18-1).

Assessment of each of the ICF environmental domains is evolving. While some researchers have developed approaches to assess the broad array of environmental factors (24), others have elaborated on physical aspects of the environment. Shumway-Cook et al. (25) developed a community environment assessment that identifies eight dimensions of the physical environmental domain: (a) temporal, (b) physical load, (c) terrain, (d) postural transitions, (e) distance, (f) density, (g) attentional demands, and (h) ambient conditions. Others ascertain the presence of community mobility barriers, mobility devices and technologies, and transportation facilitators (26).

It is not entirely clear which elements of the environment are relevant for rehabilitation considerations. It is clear, though, that the environment in which people live is complex and entails a number of factors that have elements that could impact the degree to which people with mobility limitations could be involved in their life activities, and it is likely these factors range from the factors that individuals directly interact with to societal and cultural factors.

Measurement Considerations
There are several methodological challenges to conducting and interpreting research on the links between the environment and participation. First, measurement of participation and environment is paramount to assessing the relationships between these complex phenomena. Researchers need valid and reliable measures of participation and the environment—with the assessments for each conducted in a manner that minimizes bias in the statistical relationships. The complexity of this, however, is substantial. Both participation and environment are complex domains, and assessing participation in terms of the environment in which people live is a complicated phenomenon.
TABLE 18.1 Description of Environmental Domains in the International Classification of Functioning, Disability, and Health*

<table>
<thead>
<tr>
<th>Domain</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Products and technology</td>
<td>The natural or human-made products or systems of products, equipment, and technology in an individual’s immediate environment, that are gathered created, produced, or manufactured</td>
</tr>
<tr>
<td>Natural environment and</td>
<td>Animate and inanimate elements of the natural or physical environment, and components of that environment that have been modified by people as well as characteristics of the population in that environment</td>
</tr>
<tr>
<td>human-made changes to</td>
<td></td>
</tr>
<tr>
<td>environment</td>
<td></td>
</tr>
<tr>
<td>Supports and relationships</td>
<td>People or animals that provide practical physical or emotional support, nurturing, protection and assistance, and relationship with other persons in the home, place of work, school or at play or in other aspects of their daily lives</td>
</tr>
<tr>
<td>Attitudes</td>
<td>The attitudes that are the observable consequences of customs, practices, ideologies, values, norms, factual beliefs, and religious beliefs. The attitudes classified are those of people external to the person whose situation is being described, not of the person himself or herself.</td>
</tr>
<tr>
<td>Services, systems, and</td>
<td>Services comprise structured programs, operations, and services, public, private, or voluntary, established at a local, community, regional, national, or international level by employers, associations, organizations, agencies, or government in order to meet the needs of individuals and include the persons who provide these services. Systems and policies respectively comprise the administrative control and monitoring mechanisms, rules, regulations, and standards established by local, regional, national and international government, or other recognized authorities, which organize services, programs, and other infrastructural activities in various sectors of society.</td>
</tr>
<tr>
<td>policies</td>
<td></td>
</tr>
</tbody>
</table>


Participation
While there are numerous instruments available that assess activity (or function), there are fewer assessments of participation. Since the formulation of the ICF, however, several new instruments have been published that assess participation. One critical element of participation assessment is that the domains tapped should reflect the complex daily activities or life situations of people. Furthermore, if the aim of the instrument is to assess participation as a separate component of activity, then items pertaining to activity or function should not be included in the scale. Many measures in rehabilitation have items pertaining to activity and participation levels, which makes interpretation of the literature difficult. Below we describe a few published surveys assessing participation. Although we believe the conceptual and psychometric properties of these measures are good, there are a number of other valid and conceptually sound instruments ascertaining participation that could be used.

The Craig Handicap Assessment and Reporting Technique (CHART) is a 32-item self-report instrument ascertaining level of participation across six domains: (a) physical independence, (b) cognitive independence, (c) mobility, (d) occupation, (e) social integration, and (f) economic self-sufficiency. Each domain is scored on a 0 to 100 scale with higher scores indicating more participation. A score of 100 represents the participation level of persons without disabilities. A short 20-item version of the CHART is also available (CHART-20). Reliability and validity for the CHART and the CHART-20 are published (27).

The Late-Life Disability Instrument (LLDI) assesses two dimensions of participation: frequency and limitation (28). The frequency dimension pertains to the degree to which people participate in various complex daily life activities; the limitation dimension pertains to the degree to which people perceive they are limited in performing complex daily activities. The instrument includes 16 items for both the frequency and limitation scales and was developed for use in a general aging population. Thus there are a total of 32 items in the instrument, but the same items are assessed in terms of frequency and limitation—an important and unique distinction of this scale in comparison to other instruments. Factor analyses of the instrument identified two subscales in each domain. The frequency domain includes social and management subscales and the limitation domain includes instrumental and personal subscales. Reliability and validity of the LLDI are adequate (28).

The Assessment of Life Habits (LIFE-H 3.0) is another instrument that is designed to assess participation on a general population of people with disabilities (2,29). The LIFE-H 3.0 scale is a 69-item survey that ascertains degree of difficulty in carrying out daily activities and social roles and the type of assistance needed to carry out the habits in 12 categories: nutrition, fitness, personal care, communication, housing, mobility, responsibilities, interpersonal relationships, community life, education, employment, and recreation. The instrument can be scored in subdomains or overall summary scores. Validity and reliability are acceptable (30,31).

Gray et al. developed the participation survey/mobility (PARTS/M) to ascertain participation levels in the domains of activity and participation proposed in the ICF (32). The instrument assesses participation in 20 life activities that address six domains: self-care; mobility; domestic life; interpersonal interactions and relationships; major life areas; and community, social, and civic life. Temporal, health-related limitations, evaluative, and environmental support are assessed. Validity and reliability are adequate (32).
assessments, and (c) self-report approaches that inquire about the impact the environment has on one’s level of participation or characterizes the physical, political, and social aspects of the environment (Table 18-2). As mentioned previously, the environment is, viewed broadly, consisting of social, political, technological, and attitudinal elements. While there is an extensive scientific literature on the wide array of environmental factors and participation outcomes, we focus this review of instruments and the literature primarily on physical elements of the environment and their relation to participation level outcomes.

**Global Positioning Systems and Environment Layout**

Global positioning systems (GPSs) allow objective assessment of the structural layout of communities and neighborhoods including the transportation infrastructure consisting of roads and crosswalks, the location of buildings and houses, and even direct mapping at the street level. The GPS is a technology...
that can provide a relatively precise geographic location across the earth. GPS comprises of a number of satellites that orbit the earth and simultaneously emit their satellite location in a signal. Geographic locations on the earth can be calculated from GPS receivers using triangulation from at least four GPS satellite signals (35). GPS technology became operational in 1995, and since then, an increasing number of GPS-enabled products have become available for consumer use, such as cars, phones, and watches. Geographic information systems (GISs) (36) data represent real-world objects such as roads, buildings, and parks with digitized coordinates from GPSs. GIS incorporates different types of data, such as satellite images, elevation change, and street maps, into an integrated database organized around points on the earth allowing multiple attributes of any given geographic point to be queried. The presence of parks, walking trails, gyms, and other environmental factors that facilitate walking can be identified and counted from GIS data. The relative number of these facilitators within a 1-mile radius to a person’s home can then be used to create a walkability index (37,38).

GPS and GIS technologies have corresponded with the development of new web-based interface systems that have created user-friendly and cost-free methods to query map data even at the street level. The extent of these resources varies between different regions of the globe. Many of the United States’ major urban areas have detailed GPS and GIS mapping data at the street level, with GPS technology available throughout most of the United States. It is not, however, clear whether and how these sources can be used clinically or in clinical research to identify environment factors.

**Observational**

Shumway-Cook et al. (25) developed an observational approach in which specific factors of the environment comprising the physical domain are assessed by observing people navigate their environment and participate in complex daily activities. The administration of the measure involves a structured observational encounter between a researcher and a participant—that is, a researcher observes and videotapes community mobility activities such as going to a grocery store or seeing a health care provider. Though this approach provides the investigator with an independent assessment of environmental factors, the limitations are the protocol’s cost and the time it takes to administer it. While having good objective observational assessment elements, the Shumway-Cook protocol is not feasible in large-scale field studies where feasibility, study costs, and participant burden are paramount.

**Self-Report Assessment**

A feasible approach for large cohort field studies is the use of self-report assessment techniques, and there are several published instruments using this measurement technique. The CHIEF is a 24-item self-report instrument that assesses the frequency at which people encounter environmental barriers related to attitudes and support, services and assistance, physical and architectural, policies, and work and school, and the person’s perceived impact of each environmental factor on his or her daily life (24). Respondents are asked to indicate how often various barriers in the environment have been a problem over the past 12 months. If environmental factors have been a problem, persons are asked to indicate whether the factor has been a big or little problem. Reliability and validity for all domains are acceptable (24). The Measure of the Quality of Environment is a 72-item self-report instrument that assesses the extent to which various barriers and facilitators of the environment influence people’s participation in daily life activities (39). Instruments such as those reviewed above are important for identifying relevant environmental domains as well as important barriers and facilitators. However, measuring the impact the environment has on participation in a manner that assesses the environment and participation elements at the same time could produce a biased association with participation outcomes. For example, if people indicated that their environment contained a high number of elements that impacted their participation in life activities, the association between this measure and a measure of participation outcome could be statistically inflated. To examine how environmental factors affect people’s involvement in daily activities, a better approach may be to have people characterize the extent to which a person’s environment contains elements that could facilitate or restrict participation irrespective of his or her level of function or participation and then correlate this measure with level of disability.

There are a few self-report instruments that use this approach. As a complementary approach to the observational method developed by Shumway-Cook et al. (40), the investigators developed a self-report instrument to ascertain the extent to which people avoided and encountered barriers and facilitators in the physical environment. The self-report approach of Shumway-Cook et al. has been shown to be reliable (40); however, the limitation of this approach is that asking people to ascertain the extent to which they avoid or encounter barriers may not necessarily correspond to what factors are in their environment. For example, if someone with mobility limitations is unable to negotiate stairs, he or she may state stairs are avoided but their environment may not have stairs. In addition, whether people avoid aspects of their environment is likely to be confounded by function; persons will avoid stairs if they are unable to do stairs.

Another means to characterize the environment is to describe the actual elements of the environment. This measurement approach is taken by Keysor et al. in the home and community environment (HACE) (26). The HACE is a 36-item instrument that assesses barriers and facilitators in six conceptual domains: (a) home mobility, (b) community mobility, (c) basic mobility devices, (d) communication devices, (e) transportation factors, and (f) attitudes (26). The community mobility domain assesses the extent to which a person’s community has barriers such as uneven sidewalks and curbs without curb cuts. Basic mobility devices, communication devices, and transportation facilitators are assessed to indicate whether the respondent has the device or the type of
transportation irrespective of use. Reliability is acceptable for all subscales. This approach allows the person to characterize the environment irrespective of his or her level of participation or functional ability. The instrument, however, focuses on physical aspects of the environment and as such does not address the broad array of environmental factors articulated by the ICF. In part, this was due to finding poor reliability of items pertaining to policies and services that lead to the items being eliminated during pilot testing and development of the instrument. Fange and Iwarsson took a similar approach by ascertaining people’s perceptions of their housing environment such as satisfaction with housing, suitability, security and safety, importance, privacy, social contacts, flexibility, and accessibility (11). Reliability and content validity of the 31-item instrument are acceptable (11).

Applying linking rules between environmental assessments and participation outcomes is appealing; however, no valid approaches are currently identified in the literature. Researchers face several daunting tasks in developing linking rules in the area of participation and environment. First, the ICF currently does not differentiate between activity and participation level outcomes. This makes linking the environment to participation outcomes difficult. Furthermore, the environmental domain of the ICF is not clearly defined. Currently, the environmental domain in the ICF is described in terms of a broad-spectrum of elements ranging from proximal elements such as architectural barriers in one’s home, work establishment, or community to distal and diffuse elements such as governmental agencies, informal social networks, and attitudes and ideologies. It is not clear how to best measure aspects of the environment, particularly the diffuse elements such as governmental polices and programs and community attitudes. Third, it is not clear how to measure the environment in the context of how participation evolves. In other words, how does one measure the environment if the environment is a contextual factor that envelops the person and results in participation? Nonetheless, the potential for establishing valid linking rules for this area of the ICF is high and would be useful for clinical practice and research.

In summary, over the past several years, we have seen the development of a new generation of instruments to assess the environment. There are several approaches taken by researchers, with each approach having strengths and weaknesses.

**LINKS BETWEEN THE ENVIRONMENT AND PARTICIPATION: WHAT DOES THE EVIDENCE SHOW?**

The scientific literature linking the environment to participation among rehabilitation populations is in its infancy. Research over the past several years has advanced conceptual and methodological elements of this complex field of study (41), yet there are still a number of research questions that need to be answered. It is clear, however, that a number of people with mobility limitations who live in the United States identify barriers and facilitators in their environments (42–45). Some of these barriers are able to be overcome, whereas others are not. In a study of adult wheelchair users in Boston, MA, and Durham, NC, participants reported successful and unsuccessful traveling encounters to a number of community establishments including food stores, restaurants, drug stores, friends or relatives’ homes, banks and ATM machines, places for recreation, and doctor’s offices (45). A number of barriers were reported during these trips including bad weather or climate, no ramps or ramps that were too steep, no curb cuts or blocked cuts, no parking, no public transportation, small door width, busy traffic, unsafe neighborhoods, and other people’s lack of respect and negative attitudes toward wheelchair users (Fig. 18-4). A number of facilitators in the environment were also reported including assistance from other people, accessible transportation, assistive technologies, good weather, curb cuts, and handicap parking. Another population-based study of adults with physical or sensory disabilities showed that 25% of adults with physical or sensory disabilities reported needing home modifications but not having them and having difficulty accessing a health care provider’s office because of the physical layout or location of the property and that 13% reported unfair treatment at a provider’s office because of disability. Even higher percentages of all of these factors have been reported for people with more disability and lower socioeconomic status (46). These findings are generally supported by other investigators who show that people with mobility limitations report barriers in the built environment, social attitudes, social institutions, cultural norms, and technology domains (42–45) and facilitators in social attitudes, social support, and the availability of technology, devices and products, and accessible transportation (43–45).

A perception of environmental barriers and facilitators does not necessarily link these factors to restrictions or limitations in participation in the complex daily activities of people’s lives, nor does it provide definitive proof that these factors are present. To address these limitations, elements of the environment need to be assessed irrespective of functional impairment and the measure needs to be correlated with assessments of participation. There are a few studies that use these approaches though making comparisons is difficult because different assessment instruments are used, different disease adaptation time periods are addressed, and most studies are cross-sectional, limiting the ability to infer causality between the environment and the subsequent disability. Nonetheless, there are several noteworthy findings.

First, Shumway-Cook et al. (40), using an intriguing observational assessment of physical, ambient, and density characteristics of the environment, showed that people with mobility limitations were more likely to use the elevator and less likely to ascend two flights of stairs or walk on uneven surfaces as compared to age-matched persons without mobility limitations. However, there were no differences between groups pertaining to most of the elements of the environment-participation assessment such as average distance walked, street crossings with traffic lights, temperature, light conditions, and level of precipitation. This was an interesting finding because people...
with mobility limitations reported avoiding long-distance ambulation, crossing streets with traffic lights, crossing busy streets, using stairs and escalators, walking on curbs or uneven surfaces, and going out in icy conditions (40). Thus, in this instance, when actual behavior was observed, there were fewer differences between persons with mobility limitations and those without as compared to self-reported behaviors (25).

Another approach is to correlate self-report assessments of environments with self-report assessments of disability. There are a few studies using this approach, with most showing a small but statistically significant association of the environment with participation. The studies, however, cover an array of patient populations, and more definitive information is needed about the environment-participation relationship. Whiteneck et al. (42), in a large population-based cross-sectional study of 2,726 persons with spinal cord injury, showed that people’s perceptions of barriers in physical, social, and political environmental domains had a small but significant association with participation; however, the scales only accounted for 4% of the variance in participation. Demographics, injury-related factors, and activity limitations each accounted for approximately 20% of the variance in participation. Interestingly, though, environmental factors were more strongly associated with life satisfaction, accounting for approximately 10% of the variance in life satisfaction. Participants in this study, however, had on average at least 3 years of living with their spinal cord injury, and it is possible that participants adapted to their environment and had implemented some accommodations such that their levels of participation were not significantly impaired.

Similar findings are reported among people with traumatic brain injury. Whiteneck et al. (47) in a study of 73 persons with traumatic brain injury found that the environment was
an important determinant of participation. Transportation, the surroundings, government policies, attitudes, and the natural environment were perceived to have the most impact on people's lives. The total CHIEF score and the subscales of attitudes and services, representing higher environmental barrier impacts, were associated with more disability as measured by the overall CHART score. More physical and policy barriers, as measured by two subscales on the CHIEF, were associated with the mobility and occupation subscales of the CHART. More barriers in attitudes and services were associated with less cognitive independence. Rochette et al. (48) in a cross-sectional study of 51 stroke patients selected from a rehabilitation unit showed that environmental barriers had a small but significant relationship with participation, with environmental factors explaining 6% of the variance in participation. Keysor et al. (49) conducted a longitudinal follow-up study of 342 general rehabilitation patients. The investigators showed that a greater presence of barriers in the environment and lower levels of social support were associated with less participation 1 month after discharge from a hospital or rehabilitation center. However, at 6 months after the hospital or rehabilitation center discharge, the only factor in the environment that was associated with participation was social support.

Using neighborhood diversity and density measures, Clark and George (50) reported that neighborhood housing density and land use diversity were not directly associated with disability but did modify the relationship of functional limitation on participation among participants in the 1986 wave of the Duke University site of the Established Populations of Epidemiological Studies of the Elderly (EPESE) project. Specifically, level of dependency in ADL disability and IADL disability was increased among persons with severe functional limitations as the environment became more rural (i.e., less housing density), whereas elderly people with better functional activities reported less ADL disability as the environment became more rural. The authors noted that declining housing density in this population reflected rural versus semirural living (suburban and city) and hypothesized that rural areas would have less accessible pedestrian and public transportation environments and more substandard housing conditions in comparison to semirural areas. While use of population housing density could reflect the authors' hypotheses, these measures do not capture environmental factors at the "street level" (e.g., uneven sidewalks, curbs with curb cuts, or buildings without ramps at public entrances) that could be important for disability.

These studies, while providing some evidence that the environment is associated—at least a small amount—with participation, at least in the short term, do not provide an item analysis of which elements of the environment are associated with which elements of participation. In unpublished, preliminary analyses of a large cohort of people with knee and hip osteoarthritis, we found adequate handicapped parking, the ability to drive, and the presence of community parks were associated with greater frequency of participation, while the presence of public transportation was associated with the greater perceived limitation of participation. Most recently, Wilkie et al. (51) investigated environmental factors associated with restricted mobility outside of the home in 2,252 individuals over the age of 50 with knee pain. Participants were scored as restricted if they replied as getting out of the home as “some,” “a little,” or “none” in the past 4 weeks. After controlling for knee pain severity, limitations in walking, and demographic and socioeconomic characteristics, authors reported environmental factors including poor access to public transportation and no access to a car were associated with more restrictions in mobility outside of the home.

In sum, findings from the few published studies in the literature provide evidence that the environment in which people live has a significant but seemingly small association with participation. Most of the studies reviewed are cross-sectional, and stronger evidence could be derived from longitudinal and intervention studies. Nonetheless, the studies reviewed in this section use instruments that assess aspects of the community environment. It is unlikely that the degree of barriers or facilitators in the community is confounded by the functional status (or need) of the individual. Assessing characteristics of the community environment and statistically linking them to participation levels in cross-sectional study designs and analyses minimizes confounding by indication and adds validity to cross-sectional findings. Nonetheless, longitudinal studies would provide more definitive results on whether and how the environment impacts disability.

We would like to note, however, that our review is focused on global assessments of the environment as well as more focused assessment of the physical aspect of the environment. Our approach did not include social and social support elements of the community, technology and products, or services, policies, or programs. As identified in the ICF framework, these are important elements of the environment that are presumed to have an impact on participation. Review of them, however, was beyond the scope of this chapter.

CLINICAL IMPLICATIONS

The rehabilitation field has long underscored the importance of a person's environment on his or her recovery, health status, and participation in daily life activities. Likewise, US government policy has underscored the role of the environment in disability as is evidenced by the 1990 Americans with Disabilities Act (ADA), which protected the rights of people with disabilities by prohibiting discrimination and ensuring equal opportunity and access (52). Since the ADA was instituted, a number of advances regarding environmental accessibility have been achieved. Greater attention is paid to accessibility aspects of public buildings, airlines, subway and bus systems, as well as the technologies and adaptations that are needed for persons with disabilities to pursue employment, housing, and education opportunities.

Theoretical and conceptual frameworks of disability and participation suggest that the environment provides a context to their life situations—that is, participation in complex life
roles is embedded in the environment in which people live. In rehabilitation, this environment is believed to contain a wide array of factors that facilitate or hinder participation. Research supporting these notions in rehabilitation, however, is at its infancy. Nonetheless, the published literature does suggest a small but significant relationship, at least cross-sectionally, between some aspects of the physical environment (e.g., physical, architectural barriers, technologies, and neighborhood characteristics).

So, what clinical implications can we infer from the literature at this point in time? First, participation and environment are clearly complex, multifactorial, domains that are important yet challenging to assess. Nonetheless, involvement in community-level complex daily activities of people’s lives is an important element of rehabilitation, and achieving optimal participation outcomes for rehabilitation populations is critical. Rehabilitation science professionals, therefore, are encouraged to consider participation outcomes as part of routine care for rehabilitation patients.

Second, there is literature that suggests that, at least in the short term, the environment does have an impact on participation. Thus, it seems appropriate for rehabilitation science professionals to ascertain the extent to which environmental barriers and facilitators are present in the environment as a means to help patients optimize participation, at least during initial periods of rehabilitation. It is not, however, entirely clear which elements of the environment are related to participation.

There is some research that seems to indicate transportation facilitators are important; however, much more research needs to be done in this area. Nonetheless, the evidence seems to support the notion that the community environment does have an impact on participation. Health care providers, thus, are encouraged to consider the community environment in which people live as it might impact their levels of participation.

Lastly, it is clear that more research is greatly needed in this area of rehabilitation medicine. Stronger research linking the environment to participation and explicating which elements of the environment are influential on participation will be useful to guide clinical intervention and public policy.

REFERENCES

The practice of medicine in the United States, as in most nations, is profoundly influenced by government rules, regulations, legislation, and enforcement activity. Governments set standards of professional training, define the scope of practice, authorize individuals to practice, monitor practice performance, and enforce compliance with their standards.

Government also strongly influences the activities of health care providers through certificate of need controls, licensure, and public health requirements. Facilities are subject to many operating rules and requirements, are regularly reviewed, and face many obligations as to how they operate and are reimbursed for the services they provide.

The federal government is the single largest payer for inpatient rehabilitation services (through the Medicare program) and hence has a large impact on all providers and facilities in this segment of the health care industry. Medicare fee-for-service (FFS) beneficiaries account for approximately 70% of inpatient rehabilitation facility (IRF) cases (1), and Medicare spending for IRF services is projected to be $5.8 billion annually in 2008 and 2009 (2). Medicare actuaries projected $22.8 billion in spending in 2008 at skilled nursing facilities (SNFs) (3). As a result, the regulations Medicare imposes on providers generally influence practice for all patients, regardless of payer, and become universal standards of compliance.

State governments (in partnership with the federal government) are also very important purchasers of rehabilitation care and services through the Medicaid program. Another dual program of importance, the provision of vocational rehabilitation services, is federally administered by the Rehabilitation Services Administration (RSA) and in each state by its Vocational Rehabilitation Agency.

In addition, government has a central role in assuring protection of patient rights, safety, confidentiality, and non-discriminatory treatment. The legal system addresses quality of care, access to care, and equity.

Compliance with federal regulations is monitored by the Office of the Inspector General of the Department of Health and Human Services (HHS) and the Program Integrity Group of the Centers for Medicare and Medicaid Services (CMS). State Departments of Health typically provide surveillance for compliance with state laws and the Medicaid program.

In this chapter, we describe the various key legal and regulatory features that control and influence the field of physical medicine and rehabilitation (PM&R).

**MEANS OF GOVERNMENTAL INFLUENCE OVER PM&R**

As with any regulated field, government influences PM&R through the activities of both the legislative and executive branches. The legislative branch has the sole authority to authorize programs and to appropriate funds. The executive departments and agencies are charged with implementing legislation pursuant to statutory authority, including through agency rulemaking and the enforcement of specific regulatory agendas (4). The statutory basis for a regulation “can vary greatly in terms of its specificity, from very broad grants of authority that state only the general intent of the legislation and leave agencies with a great deal of discretion to very specific requirements delineating exactly what regulatory agencies should do and how they should take action” (5). The executive branch may also propose legislation for the legislative branch to consider.

Because the federal government is the single largest payer for rehabilitation services through the Medicare program, the statutes and regulations that set forth Medicare eligibility and reimbursement policies are of primary importance to the practice of PM&R. The following section provides examples from the Medicare reimbursement system to illustrate the means of governmental influence over PM&R.

**MEDICARE STATUTES AND THE RULEMAKING PROCESS**

The Balanced Budget Act of 1997 (BBA) (PL No. 105-133) authorized a number of major structural reforms to the Medicare program, including the transition to prospective payment systems (PPSs) for Inpatient Rehabilitation Facilities (IRFs), SNFs, and Home Health (HH) Care (6). Shortly after enactment of the BBA, the Medicare, Medicaid, and State Children’s Health Insurance Program
(SCHIP) Balanced Budget Refinement Act of 1999 (BBRA) (PL 106-113) and the Medicare, Medicaid, and SCHIP Benefits Improvement and Protection Act of 2000 (BIPA) (PL 106-554) authorized payment for long-term care hospitals (LTCH) based on prospectively set rates. Prior to this transition, PM&R providers were reimbursed retrospectively on a “reasonable cost basis” (7).

These statutes set forth the parameters for each PPS but direct the Secretary of the Department of HHS to implement these payment systems through the rulemaking process. The inpatient rehabilitation facility prospective payment system (IRF PPS), for example, requires providers to be paid predetermined per-discharge rates for “classes of patient discharges...by functional-related groups (each referred to as a “case mix group”), based on impairment, age, comorbidities, and functional capability of the patient and such other factors as the Secretary deems appropriate” (8). The statute also provides for adjustments for area wage variations; individual facility variations; individual patient outliers; lengths of stay; and the annual market basket inflationary update (9). In each case, the statute provides the Secretary with discretionary authority to establish the actual classification system for patients; the weighting factors for expected costliness; the payment rates; the outlier and special payment policies; the area wage adjustments; and the annual update (10).

Implementation in these areas involves important methodological considerations and significant determinations as to policy. Express policy objectives set forth in the IRF PPS include providing proper “incentives to furnish services as efficiently as possible without diminishing the quality of the care or limiting access to care” and creating “a payment system that is fair and equitable to facilities, beneficiaries, and the Medicare program” (11).

For each PPS, the Secretary also issues annually a Proposed and Final Rule to implement, refine, and update the reimbursement system for the upcoming fiscal year (FY) in question. The government publishes a notice of each Proposed Rule in the Federal Register and solicits stakeholder comments to which it then responds in the Final Rule (12). The FY 2009 Final Rule for the skilled nursing facility prospective payment system (SNF PPS), for example, provides for a 3.4% market basket update to the payment rates used under the SNF PPS and makes certain changes to the case-mix indexes (13).

The annual rulemaking process can lead to significant structural changes in the payment system. As added by the BBA, Section 1886(j) of the Social Security Act confers broad statutory authority upon the Secretary to propose refinements to the IRF PPS; parallel provisions exist with respect to the other PPS. Under this authority, for example, the FY 2006 IRF PPS Final Rule made a number of substantial refinements to the IRF PPS case-mix classification system (the case-mix groups and the corresponding relative weights) and the case-level and facility-level adjustments (14). Further refinements were made in FY 2007 and FY 2008 “to ensure that IRF PPS payments continue to reflect as accurately as possible the costs of care” (15).

**STATUTORY AND REGULATORY DEVELOPMENTS**

Despite the basic relationship between the legislative and the rulemaking processes, policy development is almost never a linear process. In the PM&R context, legislation is often a means to alter, suspend, or otherwise supersede regulatory determinations. Legislation is also often a means to temporarily “fix” or “patch” previous laws, or to seek lasting changes in the payment systems. New statutory provisions, moreover, can establish additional regulatory requirements, which themselves must be implemented.

The list of enactments that have significantly impacted the Medicare program, and PM&R in particular, since the passage of the BBA in 1997, include the Medicare, Medicaid, and SCHIP BBRA of 1999 (PL 106-113); Medicare, Medicaid, and SCHIP BIPA of 2001 (PL 106-554); Medicare Prescription Drug, Improvement, and Modernization Act of 2003 (PL 108-173); Deficit Reduction Act of 2005 (DRA) (PL 109-171); the Tax Relief and Health Care Act of 2006 (TRHCA) (PL 109-432); the Medicare, Medicaid, and SCHIP Extension Act of 2007 (MMSEA) (PL 110-173); and the Medicare Improvements for Patients and Providers Act of 2008 (MIPPA) (PL 110-275). Examples of key statutory and regulatory developments follow below.

**The 75 Percent Rule**

It has been stated that, “No governmental requirement related to the furnishing of inpatient rehabilitation services has generated more controversy than the criteria for being classified as an inpatient rehabilitation facility—the so-called ‘75 Percent Rule’” (16). In the Final Rule implementing the Tax Equity and Fiscal Responsibility Act of 1982 (TEFRA) in January 1984, the Secretary determined that, to qualify as an inpatient rehabilitation hospital or rehabilitation unit and therefore be distinguished from an acute care hospital, a facility must have “served an inpatient population of whom at least 75 percent required intensive rehabilitation services for one or more of 10 conditions specified in the regulations” within the most recent 12 month cost reporting period (17). The August 7, 2001 Final Rule that implemented the IRF PPS did not change the procedures for classification as an IRF (18). SNFs, by contrast, have no special requirements for offering rehabilitation services. There are also no specific program regulations or requirements for rehabilitation services that may be provided in an LTCH.

On June 7, 2002, the CMS declared a moratorium on enforcement of the 75 Percent Rule because it found the rule was being unevenly enforced by the various state Departments of Health and the Medicare Fiscal Intermediaries (FIs) (19). The moratorium ended in FY 2004 when a new Final Rule for IRF classification was published, which expanded the original 10 conditions to 13 and provided that the 75% compliance threshold was to be phased-in over several years, becoming fully effective for cost reporting periods beginning on or after July 1, 2007. The Final Rule also phased out the use of specific
patient comorbidities as counting toward the required 75% compliance threshold (20).

Implementation of the revised 75 Percent Rule created a serious barrier to access to PM&R services for patients in inpatient rehabilitation hospitals and units as seen in Table 19-1. This table makes it clear that there was a dramatic drop in the number of cases admitted to rehabilitation hospitals and units, and an increase in length of stay and payment per case which represents the more severely medically and functionally impaired patients who were admitted.

It spawned intensive legislative advocacy. In the DRA of 2006, Congress responded and delayed implementation of the 75 Percent Rule by 1 year (21). After that moratorium expired, Congress subsequently passed the MMSEA of 2007, which indefinitely set the compliance threshold at 60% (22). MMSEA also provided that patient comorbidities be permanently included in the calculations used to determine whether an IRF meets the compliance percentage. With this legislative provision in place, the focus then turned to implementation of the new 60 Percent Rule (23). As a trade-off for achieving permanent relief on the 75 Percent Rule by establishing a statutory 60% compliance threshold, the field of medical rehabilitation experienced an adverse economic consequence when the MMSEA rolled back the IRF update factor to 0% in FY 2008 and held it a 0% for FY 2009. Freezing or reducing updates has become an increasingly prominent means of seeking Medicare cost savings (24).

### Payment Updates and Other Payment System Changes

Under the various PPS, and under the Physician Fee Schedule (PFS) for professional services provided under Part B of the Medicare program, reimbursement rates are to be updated annually. The update calculation is made during the rule-making process based on the index or process provided for in the underlying statute. Commonly, however, legislation is proposed and enacted as a means to override the update formula or amount. Section 131 of the MIPPA of 2008, for example, substituted a positive update of 1.1% to payment rates under the Medicare PFS for the negative update that would have resulted from the application of the statutory formula.

### Outpatient Therapy Caps

As a cost control measure, Section 4541 of the BBA of 1997 required the CMS to impose a cap on Medicare outpatient rehabilitation therapy services for outpatient physical, speech-language and occupational therapy services by all providers other than hospital outpatient departments. Because of their complexity and direct harm to Medicare beneficiaries, the therapy caps have been subject to a “series of moratoria” enacted by Congress to avoid the strict limit placed on access to outpatient rehabilitative care (25). Most recently, Section 141 of MIPPA extended the exceptions process for therapy caps at least through December 31, 2009.

### Medicare Recovery Audit Contractors

The Medicare Modernization Act of 2003 (MMA) established the Recovery Audit Contractors (RAC) program as a demonstration program to identify improper Medicare payments. The TRHCA of 2006 made the RAC program permanent and directed the CMS to expand the program to all 50 states by 2010. The RAC demonstration program as implemented in California by the contractor, PRG Shultz, resulted in numerous coverage denials based upon medical necessity reasons. Subsequently, the affected California providers appealed these denials and a large percentage (90%) of these denied coverage decisions were reversed when they were appealed, based on anecdotal reports from California IRF providers. Specifically, 6% of the overpayments collected under the demonstration program through March 27, 2008 were from IRFs (26). Of those cases, 94% were deemed medically unnecessary services or settings (27). The number of claims

### TABLE 19.1 The Number of IRF Cases has Declined Since 2004, While Payments per Case have Increased

<table>
<thead>
<tr>
<th>TEFRA</th>
<th>PPS</th>
<th>Average Annual Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medicare spending (in billions)</td>
<td>$4.51</td>
<td>$5.65</td>
</tr>
<tr>
<td>Unique beneficiaries</td>
<td>N/A</td>
<td>398,000</td>
</tr>
<tr>
<td>IRF patients per 10,000</td>
<td>N/A</td>
<td>114</td>
</tr>
<tr>
<td>FFS beneficiaries</td>
<td>Cases</td>
<td>415,579</td>
</tr>
<tr>
<td>Payment per case</td>
<td>$9,982</td>
<td>$11,152</td>
</tr>
<tr>
<td>ALOS (in days)</td>
<td>14.0</td>
<td>13.3</td>
</tr>
</tbody>
</table>

Note: IRF (inpatient rehabilitation facility), TEFRA (Tax Equity and Fiscal Responsibility Act of 1982), PPS (prospective payment system), N/A (not available), FFS (fee-for-service, ALOS (average length of stay).

Source: MedPAC analysis of MedPAR data from CMS and data on aggregate Medicare spending for IRF services from CMS Office of the Actuary. Adapted from Medicare Payment Advisory Commission March 2009; MedPAC Report to Congress.
in California with overpayments less cases overturned on appeal as of March 27, 2008 was over 4,480 (28). However, upon a subsequent review by a different CMS contractor, nearly 40% of the cases denied by the California RAC demonstration contractor were found to be incorrect (29). In implementing the program nationwide, CMS regulations relating to medical necessity denials and appeals remain a key issue (30). Legislation was introduced in the 110th Congress to impose a moratorium on the nationwide rollout of the RAC program but did not advance (31). Legislation requiring enhanced oversight and accountability of RACs is anticipated in the 111th Congress.

Quality of Care
As part of ongoing initiatives to refocus payment incentives toward quality, CMS has already implemented a “pay for reporting” system for physicians where payment rates are tied to the reporting of quality measures. MIPPA made the Physician Quality Reporting Initiative permanent and authorized incentive payments through 2010. A potential next step is “pay for quality” where “providers would not only be rewarded for reporting quality activities but their payment would also be increased or decreased depending on how well they perform on these quality measures” (32). Given the pervasive influence of the Medicare program, any such reimbursement mechanism can be expected to influence practice for all patients, regardless of payer. Payment system reform will likely include enhanced performance outcome measurements applicable to rehabilitation and LTCH.

THE PERVASIVE INFLUENCE OF MEDICARE AND MEDICAID

Given their scope of coverage, the Medicare and Medicaid programs have a profound influence on the practice of PM&R. This section sets forth the structure and features of each program.

The Medicare Program
Prior to the passage of the Medicare program, physicians practiced primarily in their offices and some in the hospitals and units providing medical rehabilitation services. Payment came from some insurance companies, private pay, and some state programs (33).

The original Medicare and Medicaid programs were enacted in 1965 through Titles XVIII and XIX, respectively, of the Social Security Act. The Medicare system was originally administered by the Social Security Administration, but in 1977 management was transferred to the Health Care Financing Administration, which was renamed the CMS in 2001. CMS is part of the U.S. Department of Health and Human Services and maintains a Web site at www.cms.hhs.gov.

Beneficiaries become eligible for Medicare by virtue of age (65 or older), disability, or if they have end-stage renal disease (ESRD) (34). Medicare has four parts which describe and cover various aspects of the program. Medicare Part A and Part B are referred to as the original Medicare program, or the FFSs program. Medicare Part C is known as Medicare Advantage (MA) and gives beneficiaries the option to receive their Medicare benefits through managed care and private health insurance plans. Medicare Part D provides a prescription drug benefit (35).

The programs have grown exponentially since inception. Short-term Medicare growth is reflected below (Table 19-2).

Medicare payments for federal FY 2009 were estimated in the President’s FY 2009 budget at $485.9 billion. The chart below displays the Medicare payments, known as outlays, estimated for FY 2009 (Fig. 19-1).

Medicare Part A
Medicare Part A, in general, covers inpatient hospital services, critical access hospitals, SNFs (not including custodial or long-term care), hospice care, and the Part A portion of HH care and services. Part A thus covers the services provided in inpatient medical rehabilitation hospitals and units, long-term acute care hospitals, psychiatric hospitals, cancer hospitals, and children’s hospitals (36).

People with over 40 quarters of covered employment are entitled to Part A without paying a premium, but most covered services require a beneficiary co-payment or coinsurance (37).
In 2009, beneficiaries paid a $1,068 deductible for a hospital stay of 1 to 60 days and $133.50 daily coinsurance for days 21 to 100 in an SNF (38). Part A is financed through a payroll tax.

The Social Security Act also provides for the enrollment, subject to payment of a monthly premium, of individuals who do not meet the requirement of 40 quarters of covered employment and are not otherwise eligible (39). CMS estimates that approximately 588,000 enrollees voluntarily enrolled in Medicare Part A by paying a monthly premium of $433 for CY 2009 (40).

Under Part A, providers are paid on the basis of a PPS specific to that provider. For example, acute care hospitals are paid per discharge under the inpatient prospective payment system (IPPS) using a classification system called Medicare Severity Diagnosis Related Groups (MS-DRGs) based on data from a series of ICD-9-CM codes (41). Notably, on January 16, 2009, the Department of HHS published the Final Rule on the adoption of ICD-10-CM, the new diagnosis coding system that is being developed as a replacement for ICD-9-CM with a compliance date of October 1, 2013 (42).

Inpatient rehabilitation hospitals and units are paid per discharge through the IRF PPS using a classification system known as Case Mix Groups (CMGs) which are derived from information collected on the inpatient rehabilitation facilities patient assessment instrument (43). SNFs are paid on a per diem basis under the SNF PPS, using a classification system known as the Resource Utilization Groups which are based on information taken from the data collection instrument known as the Minimum Data Set (44). Home health agencies (HHA) are paid per episode on the basis of the home health prospective payment system which uses a classification system known as the home health resource groups which are derived from information on the data collection known as the Home Health Outcome and Assessment Information Set (45).

LTCHs are paid also on a per-discharge basis through the long-term care prospective payment system using a case mix classification system known as the Medicare Severity Long-Term Care Diagnosis Related Groups. However, like the acute inpatient hospitals, there is no separate patient assessment instrument. The information entered on the billing—specifically the ICD-9-CMs—determines the MS LTC DRG which specifies the payment rate.

In policy arenas, each payment system and provider is referred to as a “silo of care.” There is growing interest in consolidating these into one PPS. The Post Acute Care Reform Plan developed by CMS sets an “ultimate goal of site neutral payment for PAC [postacute care] services” (46).

In addition, in December 2008, the Congressional Budget Office (CBO) released Budget Options Volume I: Health Care. This document outlines a number of proposals for reforming Medicare. Option 30 proposes to bundle payments for hospital and postacute care. Under this option, Medicare would pay a single, bundled rate to an inpatient hospital that includes the cost of postacute care based on the average cost for a specific MS-DRG across all current postacute settings (47).

This option is projected to save the Medicare program nearly $19 billion over the 10-year budget window. This proposal was echoed in the President’s FY 2010 Budget Outline, A New Era of Responsibility: Renewing America’s Promise (48). The significance of these developments is discussed further below.

**Medicare Part B**

Part B of Medicare is known as Supplementary Medical Insurance. It covers physician services including all specialty areas (49). Physician services are paid on the basis of the PFS (50). Physical therapists, occupational therapists, and speech-language pathologist services are also covered under Part B and paid under the PFS (51). Part B additionally covers outpatient services such as the therapy services mentioned above and other outpatient services including certain laboratory services, preventive exams and screening, shots, durable medical equipment, prosthetics and orthotics and supplies, comprehensive outpatient rehabilitation facilities, and certain HH care services (52). Hospital outpatient services are paid on a separate outpatient prospective payment system using a classification system known as Ambulatory Patient Classifications, which combines services that are clinically similar and similarly resource-intensive (53). Durable medical equipment is paid according to a fee schedule separate from the PFS (54).

Part B coverage is voluntary and about 93% of Medicare beneficiaries are enrolled in it (55). Approximately 25% of Part B costs are financed by beneficiary premiums, with the balance being covered by the government’s general revenue funds (56). The standard monthly premium was $96.40 for CY 2009, but premiums increase at higher levels of adjusted gross income (56).

**Medicare Part C**

Medicare Part C is known as the MA Program. It is the third in a series of iterations on Medicare managed care plans and programs. The BBA of 1997 first established a new “Part C” in the Medicare statute which provided for a Medicare + Choice program (57). Title II of the Medicare Prescription Drug Improvement and Modernization Act of 2003 made revisions to the existing provisions in Medicare Part C, establishing what is now called the MA program (58). The MA program allows Medicare beneficiaries to receive their Medicare benefits from private plans rather than from the traditional FFS program (59). Plans may offer additional services and benefits, which may be reflected in the plan’s monthly premium (59).

Only people who have both Medicare Part A and Medicare Part B are eligible to join an Advantage Plan. Under § 1851(a)(1) of the Social Security Act, as modified by § 201 of the MMA (“Implementation of Medicare Advantage Program”), every individual entitled to Medicare Part A and enrolled under Medicare Part B, except for most individuals with ESRD, may elect to receive benefits through either the original Medicare program or an MA plan. In 2008, about 22% of beneficiaries were enrolled in an advantage plan and there was a continuing trend upward in enrollment (59). Technically, Medicare beneficiaries participating in an MA plan are to be able to receive
all regularly covered Medicare services, rights, and protections (59). However, one constant concern in the medical rehabilitation community is that these plans do not provide all of the covered services, especially inpatient hospital-level rehabilitation services. Specifically, MA plans have been prone to sending many traditional inpatient rehabilitation patients such as stroke patients and orthopedic patients to SNFs for care.

There are several types of plans. They include traditional Medicare Health Maintenance Organizations (HMOs), Medicare Preferred Provider Organizations (PPOs), Medicare Private Fee-for-Service (PFFS), Medicare Special Needs Plans (SNP), and Medicare Medical Savings Accounts (MSAs). Additionally, there are Medicare Cost Plans that are similar to HMOs, with the exception that services received outside the plan are covered under the Original Medicare Plan. There are also demonstration and pilot programs and the Program of All Inclusive Care for the Elderly (60).

**Medicare Part D**

Medicare Part D is a voluntary, outpatient prescription drug benefit. It was created by the Medicare Prescription Drug Improvement and Modernization Act of 2003 and represents the largest expansion of the Medicare program since its original enactment.

There are two ways of receiving benefits under Part D: (a) through a stand-alone Medicare Prescription Drug Plan which adds drug coverage to Original Medicare, some Medicare Cost Plans, some Medicare PFFS Plans, and Medicare MSA Plans; or (b) through Medicare Advantage-Prescription Drug plans that offer Medicare prescription drug coverage through the MA plan along with Part A and Part B coverage (61).

Under the 2009 standard drug benefit, beneficiaries pay a monthly premium plus a $295 deductible, after which Medicare covers 75% of allowable drug expenses up to a benefit limit of $2,700. At this point, there is a “coverage gap” where the beneficiary pays out-of-pocket until the catastrophic limit of $4,350 is reached, at which point Medicare covers subsequent expenditures with a 5% beneficiary coinsurance (62). Private plans can make different provisions for the coverage gap (62). This so-called “donut hole” in coverage is a controversial aspect of the Part D program.

**Medicaid**

The Medicaid Program provides medical benefits to groups of low-income people, some who may have no medical insurance or inadequate medical insurance. Although the federal government establishes eligibility and general guidelines, the Medicaid program requirements are actually established by each state. The range of benefits a person is eligible for under Medicaid depends on the state in which he or she lives. States are required to include certain types of individuals or eligibility groups under their Medicaid plans. They may include other groups as well. Eligibility groups are referred to as categorically needy, who are required to be covered; medically needy, who may be covered at the state’s discretion; and special groups (Table 19–3) (63).

Categorically needy people are
- Families who meet states’ Aid to Families with Dependent Children eligibility requirements in effect on July 16, 1996
- Pregnant women and children under age 6 whose family income is at or below 133% of the Federal poverty line
- Children aged 6 to 19 with family income up to 100% of the Federal poverty level
- Caretakers (relatives or legal guardians who take care of children under age 18 [or 19 if still in high school])
- Supplemental Security Income (SSI) recipients (or, in certain states, aged, blind, and disabled people who meet requirements that are more restrictive than those of the SSI program)
- Individuals and couples who are living in medical institutions and who have monthly income up to 300% of the SSI income standard (Federal benefit rate) (63)

Medically needy people fall above the income and resource threshold to be eligible as categorically needy. If a state has a medically needy program, it must include pregnant women through a 60-day postpartum period, children under age 18, certain newborns for 1 year, and certain protected blind persons (64).

States may also, at the state’s option, provide Medicaid to a number of other groups based on medical necessity, including
- Blind persons (blindness is determined using the SSI program standards or state standards)
- Disabled persons (disability is determined using the SSI program standards or state standards) (65)

The Medicaid program may also include various special groups such as
- Qualified Medicare Beneficiaries
- Qualified Working Disabled (66)

States are required to offer certain services, such as inpatient hospital services, and may offer others. However, since states have a large degree of discretion, coverage of medical rehabilitation services varies widely, including coverage of inpatient rehabilitation hospital services. Medicaid coverage of rehabilitation services has been the subject of considerable regulatory and legislative activity in recent years (67). In August 2007, CMS issued a proposed rehabilitation services rule to

<table>
<thead>
<tr>
<th>Table 19.3 Medicaid enrollment</th>
<th>Medicaid Enrollment (enrollees in millions)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2007</td>
</tr>
<tr>
<td>Aged 65 and over</td>
<td>5.0</td>
</tr>
<tr>
<td>Blind and Disabled</td>
<td>8.5</td>
</tr>
<tr>
<td>Adults</td>
<td>11.1</td>
</tr>
<tr>
<td>Children</td>
<td>23.5</td>
</tr>
<tr>
<td>Territories</td>
<td>1.0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>49.1</strong></td>
</tr>
</tbody>
</table>
more clearly define the scope of the rehabilitation benefit; identify services that could be claimed as rehabilitation under Medicaid; and ultimately reduce federal Medicaid spending for rehabilitative care (68). Implementation of this Final Rule has been delayed by several legislative moratoria, including a provision contained in the landmark economic stimulus legislation passed in 2009.

State Children’s Health Insurance Program
In addition to a state’s Medicaid program, states have a health insurance program for children up to age 19, known as the SCHIP. SCHIP was created under the BBA of 1997 as Title XXI of the Social Security Act (69). In some states, SCHIP is part of the state’s Medicaid program; in other states it is separate; and in some states it is a combination of both types of programs. These programs are for children whose parents’ income is too high to be eligible for Medicaid but not high enough to buy private insurance. Not all the insurance programs provide the same benefits, but they all include immunizations and care for healthy babies and children at no cost. Families may have to pay a premium or a small co-payment for other services depending on their income (70).

OTHER GOVERNMENTAL PROGRAMS
Other federal and state programs also pay for various levels of medical rehabilitation services. It is important that a practitioner be aware of these for reference. They are described briefly here.

Vocational Rehabilitation
The Rehabilitation Act of 1973 (PL 93-112), as amended, is the major legislative source for programs and initiatives administered by the RSA (71). Title I establishes the state vocational rehabilitation programs and agencies. These programs are paid for by a combination of state and federal funds (71).

State vocational rehabilitation agencies must, at a minimum, provide evaluations of rehabilitation potential, including diagnostic and related services; counseling, guidance, referral, and placement services; and vocational and other training services to enhance the employability of eligible individuals (72). They also must provide a variety of other services, including physical and mental restoration services, if comparable services and benefits are not available under any other program (72). Restoration has historically included a broad range of medical and healthcare services, including medical rehabilitation services (72). Success for a counselor working for a state agency is usually determined by the number of successful case closures for individual clients, which may mean gainful employment, entering a training program, participating in special programs, or returning to living in the community (73).

Workers Compensation
Workers’ Compensation (WC) laws focus on ensuring that employees who are injured or disabled in the course of their employment are provided with medical care, rehabilitative services, and fixed monetary awards, eliminating the need for prolonged litigation. Some laws also protect employers and fellow workers by limiting the amount an injured employee can recover from an employer and by eliminating the liability of co-workers in most accidents. Contrary to general medical insurance, the focus here is on returning employees to work, and there is a strong incentive to provide rehabilitation services.

The first workers’ compensation program was established by the federal government for sectors of its civilian workforce in 1908, which was then expanded to cover essentially all civilian Federal employees through the Federal Employee’s Compensation Act (FECA) of 1916 (74). There are now multiple separate nonfederal workers’ compensation programs, one in each state, plus the District of Columbia, Puerto Rico, and the Virgin Islands. The Office of Workers’ Compensation Programs (OWCPs) within the U.S. Department of Labor provides a survey state workers’ compensation laws (75). There are also federal workers’ compensation programs which are mentioned below.

The basic concept underlying the workers’ compensation laws is to provide a mechanism for alternative dispute resolution at the workplace, thereby relieving employers of liability for negligence and relieving employees of the need to initiate lawsuits in order to receive any compensation for injuries. Considerable variation exists among programs based upon the state or territorial law.

The definition of disability used in workers’ compensation programs is generally less restrictive than that employed in the federal disability programs (SSDI/SSI). Generally, individuals claiming workers’ compensation benefits must show only that they are unable to perform their former employment or to obtain employment suitable to their qualifications and training. Although workers’ compensation benefits were originally intended largely for victims of industrial accidents, in recent years, workers’ compensation laws have been broadened to cover individuals who have occupational diseases as well (76).

Although all workers’ compensation programs provide coverage for medical services and medical rehabilitation services, the use of such services can be problematic where the employer and the employee contest a benefit claim. In such cases, the employee may perceive a strong incentive to maintain his or her “disabled” status by not actively participating in rehabilitation activities. Similarly, in certain circumstances, the relatively high weekly wage replacement benefits may also act as a disincentive to their acceptance of the rehabilitation services that might assist them to return to work.

The Department of Labor’s OWCP administers four major disability compensation programs which provide wage replacement benefits, medical treatment, vocational rehabilitation, and other benefits to federal workers or their dependents who are injured at work or acquire an occupational disease (77). The Federal Employers’ Liability Act (FELA) and Merchant Marine Act (which essentially extended FELA to sailors) protect railroad workers and merchant marines and are discussed below.
The Federal Employees’ Compensation Act (FECA) provides workers’ compensation for nonmilitary federal employees (78). Many of its provisions are typical of most state worker compensation laws. Awards are limited to “disability or death” sustained while in the performance of the employee’s duties due to injury or disease, but not caused willfully by the employee or by intoxication. The Act covers medical expenses due to the disability and may require the employee to undergo rehabilitation and/or job retraining. A disabled employee receives two thirds of his or her normal monthly salary during the disability and may receive more for permanent physical injuries, or if he or she has dependents. The Act provides for compensation for the survivors of an employee who is killed. The Act is administered by the OWCPs.

The OWCPs also administer the following disability benefit laws. Congress enacted the Longshore and Harbor Workers’ Compensation Act (79) to provide workers’ compensation to specified employees of private maritime employers. The Black Lung Benefits Act (79) provides for workers’ compensation for miners suffering from “black lung” (pneumoconiosis). The Act requires liable mine operators to pay disability payments and establishes a fund administered by the Secretary of Labor to provide disability payments to miners when the mine operator is unknown or unable to pay. The Energy Employees Occupational Illness Compensation Program Act (80) provides benefits to employees and qualified survivors of the Department of Energy and its contractors who developed radiation-related cancer, chronic beryllium disease, or chronic silicosis. Monetary and medical care compensation is only available to people and their survivors who worked or were exposed at certain facilities during specific time periods.

The Federal Employers’ Liability Act (FELA) (81), while not a workers’ compensation statute, provides that railroads engaged in interstate commerce are liable for injuries to their employees if they have been negligent. The Merchant Marine Act (the Jones Act) (82) provides seamen with the same protection from employer negligence as FELA provides railroad workers.

Veterans Administration for Discharged Veterans

The mission of the Veterans Healthcare System is to serve the needs of America’s veterans by providing primary care, specialized care, and related medical and social support services, as defined by the Veterans Administration (83). To accomplish this mission, the Veterans Health Administration (VHA) created an integrated healthcare system. The VHA has multiple offices, including the Office of Patient Care Services and the Office of the Assistant Deputy Undersecretary of Health (84).

The Office of Patient Care Services houses VHA’s clinically related programs that support the actual delivery of patient care services in the field (85). The office “integrates professional knowledge and practice skills into policy, planning, and systemwide development of patient care guidelines, critical pathways, and practice parameters” (85). The office is organized “around 13 substantive, cross-functional strategic health care groups,” that replaced a “traditional ‘stovepipe’ structure organized around discrete professions and disciplines” (85).

The Office of the Assistant Deputy Under Secretary for Health acts as “an advisor for development, implementation, and impact of VHA policy, strategic planning, and knowledge/data management (“Turning Information Into Insight”)” in support of the offices of the Under Secretary for Health and Deputy Under Secretary for Health (85). The Office contains four central offices and two field-based groups that “coordinate various aspects of VHA’s national policy development and strategic planning initiatives” (85).

Under the VA system, virtually all types of inpatient and outpatient care, including medical rehabilitation services, are covered for eligible veterans. Services are usually administered at a VA health care facility. If a particular facility is fully occupied, priority is given to veterans according to their VA priority group, irrespective of severity of illness or disability (except in the case of emergencies). Occasionally, when warranted by special circumstances, the VA authorizes payment for hospital care at a non-VA facility.

In 2004, VA established four special Polytrauma Rehabilitation Centers that focus on the injuries active duty personnel experienced in the Iraq and Afghanistan wars (86). A fifth center is scheduled to open in San Antonio, Texas, in 2011 (86). VA recognizes polytrauma as “one of several ‘signature injuries’” for veterans of Iraq and Afghanistan. Polytrauma care includes “injuries to more than one physical region or organ system, one of which may be life threatening, and which results in physical, cognitive, psychological, or psychosocial impairments and functional disability,” including traumatic brain injury, hearing loss, amputations, fractures, burns, and visual impairment (86).

TRICARE and the Military Health Care System for Active Duty, Retired, and Discharged Service Personnel

TRICARE is the Department of Defense’s health care program for “serving active duty service members, National Guard and Reserve members, retirees, their families, survivors, and certain former spouses worldwide” (87). It pays for medically necessary services and supplies, including rehabilitation therapy to improve, restore, or maintain function, or minimize deterioration of function. The “treatment must be medically necessary and appropriate medical care,” and must be “rendered by an authorized provider, necessary to the establishment of a safe and effective maintenance program in connection with a specific medical condition, provided at a skilled level, and must not be custodial care or otherwise excluded from coverage” (88). Providers can contract with the regional TRICARE contractor which then negotiates rates with providers. During the course of treatment, a single patient may be involved with TRICARE, the VA, and Medicare which can cause strain and confusion among providers and patients, and their families who must navigate among the various systems. To address such issues, the Bush Administration established a Commission on Care for America’s Returning Wounded Warriors in 2007 which recommended restructuring of these systems (89).
**Public Health Service**

The Office of Public Health and Science (OPHS) within the Department of HHS is composed of “12 core public health offices and the Commissioned Corps, a uniformed service of more than 6,000 health professionals who serve at HHS and other federal agencies” (90). The Assistant Secretary for Health oversees the OPHS and is the primary advisor to the Secretary of HHS on public health matters (90). The Office of the Surgeon General, under the direction of the Surgeon General, oversees the operations of the 6,000-member Commissioned Corps of the U.S. Public Health Service” (91). The Surgeon General “serves as America’s chief health educator by providing Americans the best scientific information available on how to improve their health and reduce the risk of illness and injury” (91).

The Indian Health Service (IHS) is part of the U.S. Department of Health and Human Services and is responsible for providing federal health services to American Indians and Alaska Natives. The IHS currently provides health services to approximately 1.9 million of the 3.3 million American Indians and Alaska Natives who belong to more than 562 federally recognized tribes in 35 states (92). The agency FY 2008 budget was $3.35 billion (93).

The IHS provides health care to Native Americans and Alaska Natives at 31 hospitals, 50 health centers, and 31 health stations (92). Thirty-four urban Indian health projects supplement these facilities with a variety of health and referral services (92). The IHS employs approximately 2,500 nurses, 850 physicians, 40 engineers, 500 pharmacists, and 300 dentists, as well as other health professionals totaling more than 15,000 in all (92).

**PRIVATE PAYERS OF MEDICAL REHABILITATION SERVICES**

Although the private sector finances fewer medical rehabilitation services than does the public sector, it plays an important role in the financing of medical rehabilitation services. The private sector payers of rehabilitation include health insurers, disability insurers, prepaid managed care providers, self-insured employers, and casualty insurers.

There are two general types of insurance in the private sector, each of which has a different orientation to medical rehabilitation services. They are:

- Life, health, and disability insurance, which provide coverage for the insured individual and typically the individual’s family
- Property and casualty insurance (including workers’ compensation, automobile liability, automobile no-fault, personal liability, and general liability), which provides coverage for injuries to individuals other than the insured

With the exceptions of life insurance and property insurance (which typically do not cover the costs of disability), private sector insurance also plays a role in paying for medical rehabilitation services. Based on known governmental payments through the Medicare and Medicaid programs, private sector insurers (including Blue Cross/Blue Shield, HMOs, PPOs, PPAs, workers’ compensation plans as mentioned above and other "commercial insurance") in the aggregate account for approximately 30% to 40% of all inpatient rehabilitation hospital revenues.

Unfortunately, many private sector payers have not embraced the full scope of medical rehabilitation services, both inpatient and outpatient. Thus, they offer a broad variety of medical rehabilitation benefits; at one end of the spectrum from zero medical rehabilitation coverage to full coverage of all medical rehabilitation services with comprehensive case management. Each type of payer has different obligations and different incentives to cover medical rehabilitation services.

**Health and Disability Insurance**

Of the two general types of private sector insurance discussed above, health and disability insurance pays the largest percentage of medical rehabilitation costs (94). Health insurance pays for medical rehabilitation services to the extent that such services are covered in the insurance contract between the insurer and the insured, usually an employer and the insurance company. Disability insurance compensates insured employees and their families for lost income resulting from a disability that prevents the employee from engaging in the occupation held at the onset of the disability (94). Like health insurance, coverage under disability insurance is dictated by the provisions of the contract between the insurer and the insured. Unlike health insurance, disability insurance contracts typically do not formally cover the costs of medical care, such as medical rehabilitation, since it is primarily an income replacement program. Despite this, some long-term disability insurance plans have paid for medical rehabilitation services on an ad hoc, extracartual basis in order to return the enrollee to work as soon as possible (95).

The two types of health insurance most relevant to the financing of medical rehabilitation are (a) primary hospitalization and (b) major medical insurance. Primary hospitalization pays the cost of covered hospital bed and ancillary services up to prescribed limits. In cases that require extended acute hospital stays, such as spinal cord injury and head injury, hospitalization insurance is often exhausted by the time the individual is medically stable enough to start rehabilitation. Major medical insurance, which provides broad and substantial coverage for large, unpredictable medical expenses, then pays for the remainder of covered services (up to prescribed ceiling, if any).

Due to the high costs of trauma care, posttrauma acute care, and medical rehabilitative care, a health insurance policy that does not include a substantial major medical component for catastrophic illness or injury is not likely to provide adequate coverage for comprehensive medical rehabilitation services or extended rehabilitation services. Therefore, the prevalence of major medical insurance, to the extent that such policies cover medical rehabilitation, is particularly relevant in considering the financing of rehabilitation services.
STATE LEGISLATIVE AND REGULATORY ASPECTS OF PM&R PRACTICE

Professional Licensing
In contrast to the national programs of voluntary certification by specialty boards (such as the American Board of Physical Medicine and Rehabilitation), medical licensure is administered by each state. State laws define the scope of practice for physicians, as well as all other licensed independent practitioners (such as psychologists, physician assistants, nurses, advanced practice nurses and nurse practitioners, podiatrists, pharmacists, physical therapists, occupational therapists, chiropractors, and many other professionals). Practice laws may be amended from time to time, and many professions seek to change the limits and scope of their practices by lobbying for changes to these laws (96).

State licensing authorities generally are organized to have a governing body of appointed professionals or other stakeholders, and a full-time staff (97). These agencies administer the regulations around licensure and handle tasks such as verifying credentials and training, assessing practitioner competency (through administration of tests or acceptance of other standardized measures), and monitoring ongoing practice. Many states use their licensing authorities to capture data about practitioners and to enforce training in professional areas deemed to be important, such as risk management, cultural sensitivity, and other topics.

State licensing authorities also oversee programs for practitioners who run afoul of proper practice, violating professional standards of practice, or laws (97). Licensing authorities enjoy the authority to restrict a practitioner's activities or to revoke a license in order to protect the citizens of the state.

Facility Definitions and Licenses
Organizations that seek to deliver health care services generally need to be licensed to do so by the state. Many states also control the health care market by maintaining Certificate of Need or Determination of Need laws (and agencies to administer them) that regulate or restrict the number of beds or facilities that may operate, and also control expansion and significant renovation expenditures (98). Rehabilitation hospitals and outpatient units are subject to scheduled and unscheduled surveys by the state. Unscheduled surveys may be random or as a result of a complaint or allegation of deficiency. In general, when a hospital is found to have operating deficiencies, it is given an opportunity to appeal the findings and to remediate the problems. In the extreme, the state has the authority to prevent the entity from admitting new patients, or even operating, and may impose monetary fines on the owners for certain violations.

Corporate Practice of Medicine Laws
A number of states have laws controlling or preventing the employment of physicians and other professionals by corporations, under the theory that their independent practice judgment could be compromised by obligations imposed by a corporate employer. These laws generally do not apply to not-for-profit corporations or professional practice corporations. In states with strong corporate practice of medicine laws, there may be few to none facilities operating that are structured as for-profit enterprises (99).

Professional Liability Insurance
Malpractice insurance is usually desired by practitioners and rehabilitation hospitals to protect them from adverse judgments in liability suits. Premium rates are influenced by the specialty, local market, individual state’s Insurance Commissioner or agency, and the history of paid claims in the state. In many states, strong economic incentives appeal to plaintiff attorneys (who may be working under a contingency fee arrangement and could receive one third of any award) to encourage filing malpractice suits. States that have passed laws to cap awards in some manner generally have more modest premium rates than those with unconstrained systems (100).

The Federal Budget Process
The annual federal budget process dictates the approach to the appropriated and the entitlement programs. Annual appropriations are necessary to fund several key PM&R programs, including funding for health research and related demonstration projects. The federal budget process distinguishes between “discretionary” or “appropriated” programs and “mandatory or entitlement programs” (101). Mandatory or entitlement spending has its legal basis in previously enacted statutes that provide spending automatically, as for eligible beneficiaries of the Medicare and Medicaid programs. Congress can affect spending on entitlement programs by changing eligibility requirements or the structure of programs through the authorizing legislation. Such changes to “direct” spending are under the jurisdiction of the authorizing, not the appropriating, committees.

Hence, Medicare and Medicaid, as entitlement programs, are not controlled by annual appropriations but can be altered through changes to the authorizing legislation (101). For example, the President’s FY 2009 Proposed Budget included Medicare legislative proposals with estimated savings of $12.2 billion in FY 2009 and $178 billion over the 5-year budget window, achieving those savings through reductions in many of the Medicare payment updates (102). The above referenced Budget Outline for FY 2010 references Medicare and Medicaid proposals with an estimated savings of $1.8 billion for FY 2010 and of $316.005 billion over 10 years.

The federal budget process is controlled by a series of procedures set forth in Titles I through IX of the Congressional Budget and Impoundment Control Act of 1974 (“Congressional Budget Act”), as amended (103). The process begins with the submission in early February of the President’s Proposed Budget request to the Congress. The President always presents the recommended annual budget for the next upcoming FY which begins on October 1 of any year (e.g., FY 2011 begins on October 1, 2010 and ends on September 30 of the subsequent year). The President’s FY 2010 Budget released in
February 2009 proposed spending for departments and agencies beginning on October 1, 2009. The House and Senate Budget Committees use the President’s budget request as a starting point to draft what is known as the “Budget Resolution.” Unlike an ordinary bill, the Budget Resolution is a “concurent” Congressional resolution, and therefore does not require a Presidential signature or veto. It also requires only a majority vote to pass and cannot be filibustered in the Senate (104).

The Congressional Budget Resolution sets spending targets in each of 19 broad spending categories (known as budget “functions”). Those amounts are then allocated to the corresponding Congressional Committees and Subcommittees through two processes—known as 302(a) and 302(b), respectively, and serve as the target for discretionary programs under their jurisdiction. With respect to the annual appropriations process, the Appropriations Committee receives a single 302(a) allocation and decides upon the 302(b) allocations to the Subcommittees with jurisdiction over each of the annual appropriations bills (105).

The Budget Resolution is often little more than a “target” for the appropriators and can readily be exceeded or bypassed. The House and Senate each pass separate Budget Resolutions and attempt to reconcile differences and reach a final Conference agreement which then is passed by both Houses typically by late March. A number of times Congress has actually failed to reach agreement on a Budget Resolution and the legislative process typically proceeds. For example, a final Budget Resolution was never adopted or enacted by Congress for FY 2009.

Appropriators and authorizing Committees enjoy the authority to make precise determinations of levels of funding and/or budgetary cuts to departments and agencies within their respective jurisdictions. The aggregate spending allocations for discretionary spending serve only as an internal control, enforceable through “points of order” and other procedural mechanisms. The Budget Resolution has proved particularly dispensable in the House because the Rules Committee can provide for it to be waived by a simple majority vote (106).

There are 12 Appropriations Subcommittees in both the House and Senate. The individual appropriations bills receive consideration through the normal Congressional Appropriations Committee process as each Subcommittee’s bill is drafted; marked up (e.g., amended within the Committee); and reported to the chamber floor for a vote. Many appropriations bills contain certain accounts where the Appropriations Committee can direct funds to individual projects or programs that meet the criteria of that account. This process, known as earmarking, is a subject of some controversy but is an important mechanism to secure Congressional funding for specific programs and priorities, including those advancing the practice of PM&R.

Despite the timeline delineated in the Budget Act, Congress often fails to pass the appropriations bills before the end of the FY. To continue funding for federal programs, Congress must then enact a Continuing Resolution (or series to Continuing Resolutions) to keep the federal government operating. Typically, these bills extend appropriations at the prior FY level, or even at a reduced level below that of the prior FY, until the next appropriations bill is passed. Congress often achieves passage of the appropriations bills by combining some number of them into a single, large omnibus bill that is easier to pass as it becomes an “all or nothing” proposition (107).

CRITICAL CURRENT REGULATORY AND LEGISLATIVE ISSUES

At the time of drafting this chapter, a number of major legislative and regulatory issues were under active debate.

Health Care Reform and Medicare Overhaul

The Obama Administration has announced that health care reform is a high priority and the leadership of the relevant key House and Senate Committees has been immersed in the legislative process of developing a bill. President Obama convened a health reform summit at the White House; Congressional Committee leaders have organized various “stakeholder” meetings to elicit support and educate disparate interest groups; and there is daily press attention on this topic.

As discussed above, the CBO has outlined a number of proposals for reforming the Medicare program, including options to bundle payments for hospital and postacute care and to reduce the update factor for hospitals’ inpatient operating payments under Medicare by 1% point (108). This proposal by CBO is extremely significant. Members of Congress and their staff look to the CBO, and to this volume in particular, to find available savings to use for other legislative initiatives. This is especially true in light of the desire by President Barack Obama and key Members of Congress to engage in extensive and costly health care reform. The President’s Budget Outline for FY10 calls for “reserve fund of more than $630 billion over 10 years that will be dedicated toward financing reforms to our health care system” (109). Given the severe current and foreseeable fiscal strains on the government, President Obama has also called for major overhaul of the nation’s retiree benefit programs as part of a deficit reduction program (110).

Definitions of Rehabilitation Providers

CMS continues to conduct various studies designed to review the entire postacute care field, with the stated intention of creating a new measurement system for patients that would be common to all settings, and using that tool to establish a new reimbursement methodology for IRFs, LTCHs, SNFs, and HHA s (111). Medicare is conducting the Post Acute Care Payment Reform Demonstration project (PAC PRD) authorized by the DRA of 2005. Part of that project required the Secretary of the Department of HHS to develop a common patient assessment instrument. The Continuity Assessment Record and Evaluation (CARE) tool (112), developed by CMS’ contractor, Research Triangle Institute International (RTI), is, at the time of this writing, the instrument developed
pursuant to the PAC PRD and which may be utilized to establish this “site neutral” payment methodology. The definitions of the various entities that provide rehabilitation services and programs will become even more meaningful under this new proposed reimbursement methodology. In the outpatient venue, CMS is conducting another project, Developing Outpatient Therapy Payment Alternatives, which is also seeking to develop an assessment tool that may be used for a future change in outpatient therapy payments in all settings. RTI is also the CMS contractor on this project. Both of these efforts are comprehensive and represent initial efforts in developing comprehensive data collection tools across the Medicare Part A providers such as IRFs, LTCHs, SNFs and HH and, separately, the vast and more disparate array of outpatient therapy providers paid under Part B of Medicare.

**Medical Necessity Review**
A key issue under discussion is the determination of medical necessity by a third party to justify payment for any service provided to a patient. Many commercial insurers, as well as the Medicare program, conduct retrospective audits and reviews of bills and deny payment coverage on the basis of their judgment that the service delivered was not medically necessary and/or could have been provided in a less intensive setting. These reviews by Medicare are conducted by either (a) the FIs (now known as Medicare Administrative Contractors; (b) RAC; or (c) Zone Program Integrity Contractors. CMS has implemented a nationwide version of the RAC Program. Legislation was previously introduced to impose a moratorium on the RAC program, and further legislative action related to RAC contractors and/or medical necessity can be anticipated in the 111th Congress.

**Payment Updates**
For the last several years, the President has proposed reducing the IRF market basket update in his annual budget (113). For example, in February 2008, President Bush proposed to freeze IRF’s payment update from 2010 to 2011, with a market basket update minus 0.65% after that (113). The budget also set out five procedures for which IRFs and SNFs would receive the same reimbursement (113). The Obama Administration has, as noted above, proposed major reductions in the Medicare and Medicaid programs for FY 2010 in order to fulfill the promises made during the campaign to also propose major health care reforms, but many details have not yet been addressed.

**Methods of Advocacy and Representation**
The field of PM&R impacts many stakeholders, including patients, physicians, therapists and other professionals and staff, providers, and families. Given the scope of the issues on the legislative and regulatory horizon, effective advocacy and representation are of more importance than ever before. At the core of the advocacy process is the notion that accurate and informed stakeholder input produces educated and committed Members of Congress, better legislation, and more effective regulation. In parallel, stakeholders benefit from representation offering expertise in the complexities of the political and legislative processes, as well as the subject area, to function as a necessary intermediary.

Within the field of PM&R, as is common, a number of associations have formed around various concerns and interests within the field. These associations educate and mobilize their members, pool organizational resources, and provide a focal point for advocacy efforts in state capitals and in Washington, DC. The American Medical Rehabilitation Providers Association is the leading national trade association representing over 350 inpatient rehabilitation hospitals, rehabilitation units, outpatient rehabilitation service providers, and several SNFs providing medical rehabilitation services to well over 400,000 people a year. The American Academy of Physical and Rehabilitative Medicine represents physicians dedicated to rehabilitative medical care. National associations representing therapy professionals include the American Physical Therapy Association, American Occupational Therapy Association, and the American Speech-Language-Hearing Association. Patient groups include the American Association of People with Disabilities, the largest national nonprofit cross-disability member organization in the United States.

Given the complexities of the legislative process, professional representation has become an increasingly indispensable element of effective advocacy. Lobbyists working on behalf of PM&R stakeholders, in turn, must often pursue both direct and indirect legislative strategies to succeed in today’s advocacy environment.

In executing a direct strategy, the advocate works within the legislative and regulatory policy processes to educate and persuade elected officials and regulators. This requires accurate and credible information about the impact of a policy or proposed policy, presented in a way that the decision maker understands and that is meaningful to him or her both substantively and politically (114). It also requires a detailed knowledge of “how the various institutions work, internally and with each other,” and more generally having “a clear fix on how the government actually works, how the pieces fit together, and how things get done” (114). Direct advocacy involves careful analysis and drafting of legislative proposals; tracking their progress in Congress; arranging meetings on Capitol Hill; working with key Congressional staffers; working with key administrative personnel; and cultivating relationships and alignments.

The most effective advocacy genuinely engages and activates grassroots constituents in the task of educating and communicating with Members of Congress and federal departmental and agency officials. When a Member of Congress understands the impact of a policy—or lack of federal action—on patients, families, and providers within a given community, that Member is more frequently motivated to intervene or undertake Congressional action to be responsive to the constituent and community needs.

An indirect or external communications strategy is often important to demonstrate the extent of grassroots support or
opposition on a given issue. Successful advocacy thus requires building coalitions around common interests and issues and using media and communications techniques to mobilize support and to make it known to key elected officials. National organizations working together with a common agenda and message can enhance effectiveness. Likewise, mixed messages, discord, competing messages, or disparate policy positions within a given field can serve as the quickest way to harm or kill prospects for effective advocacy.

Internet-based technologies have now become an important means to connect people to one another and to their elected representatives and often supplant print and broadcast communications. The challenge for the advocate is to effectively leverage these resources, and that is a challenge that requires lobbyists to be both substantive policy experts and sophisticated communications strategists.

**CONCLUSION**

The rule of law is one of the cornerstones of a civilized society. The laws and regulations that impact the care of patients in need of medical rehabilitation are especially complex and important, since rehabilitation functions strongly in the realms of activity and participation, as well as disease and impairment. To successfully practice any discipline within the field of medical rehabilitation, an understanding of the opportunities and restrictions provided by these laws and regulations is necessary. In addition, one part of the responsibility of professionals in the field is advocacy for advancing the cause of persons with a disability, and awareness of rules and regulations that may require amendment or change is essential to that end.

**REFERENCES**


15. Inpatient Rehabilitation Facility Prospective Payment System for Federal FY 2007; Final Rule (71 FR 48354); Inpatient Rehabilitation Facility Prospective Payment System for Federal FY 2008; Final Rule (72 FR 44283).
17. Prospective Payment for Medicare Inpatient Hospital Services; Final Rule (49 FR 234).
18. Prospective Payment System for Inpatient Rehabilitation Facilities; Final Rule (66 FR 41316).
27. The Medicare Recovery Audit Contractor (RAC) Program: An Evaluation of the 3-Year Demonstration, Centers for Medicare and Medicaid Services, June 2008:37, Appendix F.
32. Senator Max Baucus (D-Mont.), *Call to Action – Health Reform 2009*. November 12, 2008. (Senator Baucus is the Chairman of the Senate Finance Committee, which has exclusive jurisdiction over the Medicare and Medicaid Programs.)
34. Social Security Act § 1811 (42 U.S.C. § 1395).


44. Prospective Payment System and Consolidated Billing for Skilled Nursing Facilities for FY2009; Final Rule (73 FR 46415), August 8, 2008.


53. Changes to the Hospital Outpatient Prospective Payment System and CY 2009 Payment Rates for the CY 2009 OPPS; 73 FR 68501–69380, November 18, 2008.

54. The data files for the Medicare Durable Medical Equipment, Prosthetics, Orthotics, and Supplies (DMEPOS) Fee Schedule are available at http://www.cms.hhs.gov/DMEPOSFeeSched/01_overview.asp.

55. Figure based on analysis using the most recent Medicare enrollment data from CMS’s website. See http://www.cms.hhs.gov/MedicareEnRpts/Downloads/06All.pdf.


57. Social Security Act § 1851, Part C-Medicare–Choice Program.


61. CMS. Medicare and You 2009, 63.


72. Regulations pursuant to the Rehabilitation Act are found at 34 CFR Part 361. See especially Section 361.41 through 361.90 on the “Provision and Scope of Services.”

73. CFR Part 361.82 through 361.89 (discussing performance, evaluation, and reporting issues).


87. TRICARE Management Activity. What is TRICARE? Available at http://www.tricare.mil/mybenefit/home/overview/WhatIsTRICARE.


92. Indian Health Service. Fact Sheet: Year 2008 Profile. Available at http://info.ihs.gov/Profile08.asp.


97. Federation of State Medical Boards, Directory of State Medical Boards. Available at http://www1.va.gov/health/orgs.asp.


A 59-year-old man suffered a left hemispheric cerebrovascular accident. Two weeks later, he was admitted to a rehabilitation unit for treatment of deficits resulting from expressive language impairment and right-sided weakness. He did not agree with the self-care, mobility, and speech goals proposed by the rehabilitation team. Rather, he was anxious to return to his home and work. His wife of 30 years was apprehensive about caring for him in light of his level of disability, and urged the team to continue therapy.

Patients, family members, and practitioners often disagree about the goals, processes, or utility of rehabilitation therapy. Health care practitioners face difficult dilemmas as they attempt to set a course that all parties will endorse. In this patient’s case, practitioners were torn between honoring the patient’s desire for discharge and respecting the concerns and wishes of his wife. They did not want to treat a patient who failed to provide informed consent, yet they realized that the patient’s wife was not prepared to care for him at home. They knew that training would improve his functional abilities and enhance his eventual return home.

Rehabilitation practitioners confront moral quandaries often during the course of practice. A single moral principle—such as that of autonomy or beneficence—may fail to outweigh other principles, yet choices must be made. Practitioners must attempt to reconcile and assign priority to conflicting moral obligations (1). Decisions of a moral nature are distinguishable from those governed by law, technology, religion, or politics. They focus on what is proper rather than on what is possible or legally permissible. Considerations of etiquette, cost, and convenience play an insignificant role in moral decision making.

The terms moral and ethical are closely related; both stress manners, customs, and character. Cicero apparently used the Latin word moralis to translate the Greek ethikos (2). Contemporary usage reflects a divergence of meaning, however. Ethics refers to theoretical and contemplative descriptions of values. Morality describes conduct—that is, whether behaviors are right or wrong. This chapter traces the development of ethical systems in medicine and discusses their application to rehabilitation practice. It describes moral conflicts inherent to issues frequently encountered in clinical rehabilitation practice and recommends approaches to resolve these dilemmas. Policy issues including resource allocation, health system reforms and costs, and responsibilities of rehabilitation professionals are discussed.

HISTORICAL DEVELOPMENT

Religious Influences
In the 5th century B.C., Hippocrates described scientific, technological, and ethical aspects of medical care (3). The Hippocratic Oath derived from the traditions of a religious sect known as the Pythagoreans. In swearing allegiance to the Oath, a small group of physicians on the Isle of Cos vowed secrecy and loyalty to their teachers and promised to seek the virtues of purity and holiness. Compelled to help patients and forbidden from harming them, physicians alone were qualified to determine how to treat sick patients.

Religious traditions influenced the development of medical ethics through the Middle Ages, when monks dominated medical practice, and beyond (4). Thereafter, Catholics incorporated principles of medical decision making into their moral theology (3). Protestants, too, have examined specific ethical topics in detail and integrated concepts of medical ethics into a larger, systematic theology (3). Orthodox Jews have linked Talmudic and rabbinical teachings to the practice of medicine, with emphasis on the preservation and sanctity of life (3).

Secular Influences
During the Age of Enlightenment, the influence of religion on medical morality diminished as secular theories of philosophical reasoning arose. Controversy flourished as scholars studied, debated, and published a range of theories of medical care (3).

Codes of medical practice developed over time in response to medical challenges. In 1789, an epidemic of typhoid created chaos in a British hospital as staff members assumed onerous and unfamiliar responsibilities. As tension heightened and some staff resigned, a retired physician, Thomas Percival, sought to restore calm by designing a code of professional conduct. He instructed physicians to address the needs of individual patients in preference to those of the larger society and to show “tenderness with steadiness and condescension...
Recent Developments

Medical ethics evolved relatively little over the ensuing years. But more recently, particularly during the past 40 years, the practice of medicine has evolved dramatically. So, too, has medical ethics. Technological advances have established a scientific basis for medical treatment. Developments in health care and the biological sciences have produced complex and difficult moral dilemmas that transcend discrete professional boundaries.

Callahan describes five factors that give rise to ethical tensions in health care (6). First, technologies such as renal dialysis, organ transplantation, genetic engineering, and embryonal transplantation have expanded our ability to intervene in nature. A strong social commitment to health care coupled with a compelling tendency to apply available technology has made it difficult to restrict the use of technology.

Second, medical resources are costly. When medicine could do little to help people, care tended to be cheap. But now that improved neonatal, emergency, and acute care interventions save many lives, the chronicity and cost of disease have skyrocketed. For more than 40 years, health care costs have risen substantially faster than personal income or the overall economy (7). At $2 trillion annually, health care expenditures now exceed one sixth of the American economy (8).

Callahan cites an expanded role of the public as a third factor prompting recognition of ethical issues (6). The solitary and secretive aura of Hippocratic medicine has been supplanted by health care that is conducted in an increasingly public arena. More than 80% of Americans die in hospitals. Taxpayers fund health care entitlement programs and support medical research. Research on human subjects is regulated by federally mandated Institutional Review Boards. Legislation is in place to protect patients and assure their privacy.

The language of rights has evolved as well. Support for individual rights in American society and enhanced respect for the rights of racial minorities, women, and disabled persons have directed attention to patients’ rights. In the medical context, respect for self-determination and personal dignity requires recognition that patients have a right to make their own decisions (3).

Finally, Callahan cites increasing concern about quality of life, since many lives are now preserved and extended by medical interventions. We may wonder, however, what kind of life some survivors will lead (9). At times, the burdens of cure appear to outweigh its benefits (10).

Dramatic change has spread throughout the health care insurance industry in recent years. Mechanisms of reimbursement for health care have shifted due to the emergence of managed care and other for-profit systems. Expectations concerning the relationships and roles of payers, practitioners, and consumers have been profoundly altered.

What About Rehabilitation?

Until the recent past, little formal attention was directed to ethical aspects of rehabilitative care. A number of explanations exist (11). Still a relatively young field, rehabilitation medicine has sought to acquire recognition and acceptance by the medical community (12). Its chronic care dilemmas may seem to lack the drama of life-and-death decisions. Often patients are treated over an extended period of time by a broad range of practitioners, none of whom possesses clear responsibility for addressing ethical issues.

Recognition of moral dilemmas inherent to chronic care has increased interest in ethical aspects of rehabilitation. Questions have been raised about duties of professionals, dynamics of professional-patient relationships, roles and expectations of family members, and goals of care (11,13–29). A discussion of fundamental ethical principles furnishes a conceptual framework to study issues relevant to rehabilitation medicine.

Ethical Principles: Beneficence, Autonomy, and Justice

The term beneficence connotes kindness, charity, and the doing of good; it refers to a moral obligation to help other people, refrain from harming them, and attempt to balance benefits with harms. In the health care setting, beneficence entails an obligation to promote the health and well-being of patients and to prevent disease, injury, pain, and suffering (2).

Whether beneficence can be assured is uncertain when patients’ values conflict with traditional medical values of healing and care. Differences of opinion among patients, family members, and professionals about the best interests of patients or criteria for a good quality of life are not easily resolved. Discrepant interests may be difficult to balance within a moral framework. Beauchamp and Mcullough suggest that “beneficence includes the obligation to balance benefits against harms, benefits against alternative benefits, and harms against alternative harms” (4).

The principle of autonomy is grounded in the notion of respect for the values and beliefs of other people. Americans are entitled to privacy and to make decisions about their lives. They are perceived as possessing a right to self-determination that ensures freedom to make personal choices and resist the intervention of others. Within the context of health care, autonomy underlies the medical doctrine of informed consent. There is an obligation to inform patients accurately about their diagnoses and treatment alternatives, as well as to seek their permission before instituting treatment. Decisions are respected, even if they appear to be unwise (2,4).

Many authors describe tension between the principle of beneficence, which requires acting in a patient’s best interests, and that of autonomy, which entails respecting patient choices (4,30). Balancing the two principles is a perpetual struggle for health care givers who consider certain decisions of patients to be harmful. It is not clear how to view patient behavior when patient choices seem at odds with information presented to them. To Englehardt, “the moral obligation to
respect persons will often constrain physicians to acquiesce in patients’ choices—choices that most likely will lead to the loss of important goods” (30). In everyday practice, health care givers may be tempted to act paternalistically by restricting patient freedom to make autonomous choices if those choices appear to compromise the patient’s best interests.

The principle of justice concerns questions of what is due to whom and how to distribute the burdens and benefits of living in a society. An egalitarian model of distribution obliges society to provide all its members with a fair share of health care resources and to treat people equitably. However, scarcity of resources or competition for them can create conflict (2). Is it fairer for people to share social goods equally, or for those who have less to be given more?

Americans have yet to define the basic medical services that will be provided for all citizens. For example, despite the fact that sexually transmitted diseases are relatively common and may present public health hazards, some people receive neither prevention nor treatment efforts. But other treatments, such as organ transplants or elective cosmetic surgery, appear to be readily available to some patients. Entitlement programs in certain states pay for procedures not funded in others. Emergency treatments may be generally available, but aftercare and rehabilitation to improve the lives saved may not be funded adequately. Millions of Americans, many of whom work, do not qualify for publicly funded insurance programs, yet cannot afford to buy private health care insurance. Even for those who have insurance, reimbursement plans may provide vastly differing levels of coverage.

Concern for individual dignity and choice has overshadowed a holistic view of society’s needs. Callahan ponders what kind of medicine is needed by a society which must address other needs, including education, housing, welfare, and culture (9). Determining how much health care to support spurs us to reflect on how we want to live within our larger society. Fundamental questions of meaning are not easy to answer. In contrast to other developed nations which may possess a strong orientation to community, Americans have lagged in developing and implementing a sound health care delivery system based on a just social policy.

**CLINICAL PRACTICE ISSUES**

Acute care physicians attempt to save lives, relieve symptoms, reverse the course of pathologic processes, and discharge medically stable patients. Unlike acute care medicine, rehabilitation does not center on a sick patient whom treatment is expected to cure. Rather, rehabilitation practitioners treat dysfunctions that are chronic, often irreversible, and rarely curable. Residual disability may well persist throughout a person’s life.

Medical rehabilitative care addresses impairment caused by pathologic processes that include disease, accident, and congenital abnormality. Disabled persons experience restricted ability to perform certain functions. When unable to execute activities important to role fulfillment, a person is said to be handicapped (31). Rehabilitation therapy attempts to ameliorate handicap by restoring skills and capabilities through functional retraining and environmental adaptation. Many professionals—including, but not limited to, physicians, nurses, psychologists, social workers, and educators, as well as physical, speech, occupational, recreational, and vocational therapists—contribute to this effort. They must reach beyond pathology and physiologic dysfunction to understand the unique familial, social, vocational, psychological, and financial dimensions of patients. Successful rehabilitation requires attention not only to medical but to noomdical factors ranging from skin tolerance to household layout and availability of caregivers.

The need for a team of rehabilitation professionals and family members to engage in each patient’s treatment leads to complex relationships that may be portrayed by three points of a triangle representing patient; family, generally including several people; and health care team, including practitioners of multiple disciplines. Those at each point share concerns with others but harbor unique considerations as well. Competing rights and obligations of patients, family members, and practitioners, and blurred responsibilities and loyalties can cause confusion or conflict.

Following the seminal description of ethical issues in rehabilitation medicine in the Hastings Center Report in 1987, Kirscher studied ethical issues of greatest concern to rehabilitation clinicians (17,32). Practicing clinicians were asked to describe ethical conflicts encountered in daily practice. A broad range of clinicians listed four primary issues: reimbursement and allocation of scarce resources; determination of rehabilitation goals; compromised decision-making capacity in patients; and concerns about confidentiality (32). They considered conflicts among the treatment team especially troubling. Practitioners expressed unease about balancing their obligations to payers with their roles as patient advocates. How to ensure justice in treatment decisions worried them.

In another study, registered nurses, many of whom held management positions, reported concern about misallocation of medical resources—including both overtreatment and suboptimal treatment of patients (33). They questioned how to assure an optimal balance of providing good and avoiding harm while respecting patient autonomy. Ethical issues identified in similar surveys (Table 20-1) are discussed in this chapter within the context of clinical and policy dilemmas inherent to rehabilitation care.

**Selection of Patients**

Provision of medical care in the United States does not necessarily reflect need for it; in 2007, more than 47 million Americans lacked health insurance, and over one third of the population under the age of 65 was uninsured at some point in that year or in 2006 (34). Even for patients who are insured by Medicare, studies have shown that race, income level, and geography affect the care received (35). Many patients may not be entitled to reimbursement for rehabilitation, nor is rehabilitative care guaranteed for those whose lives have been saved by acute care interventions.
Rehabilitation practitioners screen potential patients before selecting those who will receive treatment. Providers recognize that not all patients will benefit from rehabilitation therapy, since some are too ill to participate in therapy or have impairments that are difficult to rehabilitate, while others have relatively insignificant functional impairment (36). Assessment of prospective patients requires practitioners to review information derived from hospital records and consult with referring and other treating professionals. They may also examine patients or interview family members.

Because treatment requires patients to solve problems by applying new approaches to their functional needs, their ability to learn and retain information is considered crucial to successful rehabilitation. A person’s age and predicted course of recovery also influence a decision to initiate care. Those patients who are anticipated to make significant progress are usually viewed as good candidates for rehabilitation despite severe dysfunction.

Practitioners explore nonmedical parameters in admission decisions. Particular attention is directed to whether family members are physically and emotionally available to help patients, because strong social support systems correlate with positive patient outcomes (37). Financial capacity is a powerful determinant of access to services, since gains made during treatment are more likely to be retained if financial resources are available for further services or equipment as needed.

Features of the rehabilitation unit itself influence selection of patients. Some units specialize in treatment of specific disorders or impairments. Others emphasize training patients for work or independent living. Fluctuating bed availability or staffing patterns may affect selection such that surplus program capacity at a given time may prompt admission of patients who might otherwise be rejected (16).

**Values**

Although practitioners are guided in the selection process by a patient’s potential to benefit, some professionals may neglect evidence-based information regarding rehabilitation prognosis. Ability to pay and the burden of care that staff are likely to experience also influence decisions, yet the fact that admission criteria are not publicly disclosed not only confers significant flexibility on decision makers but creates a potential for injustice as a result of bias or subjectivity. A lack of training with respect to moral dilemmas may cause practitioners to make judgments that reflect their personal belief systems and values (16). Decisions may also be influenced by perceptions that society desires to save money in its care for disabled persons, or by practitioner bias about the likelihood of good rehabilitation outcomes in older patients (17). Decisions may be made without sufficient understanding of the body of evidence-based information regarding rehabilitation prognosis.

Dilemmas in selection abound. Who should be selected: a patient with considerable need but a relatively poor prognosis, or one with lesser need and disability, but the likelihood of a better outcome? A young person who will apply his or her training throughout a lengthy life span, or an older person with few remaining years despite his contributions to society over a lifetime? How much treatment should be available to patients who bear responsibility for their disability, or for those who have been noncompliant with past treatment?

Our current approach to patient selection seems to favor those already well off. Englehardt describes the impact of lotteries: a “natural lottery” of one’s talents and abilities, diseases and illnesses; and a “social lottery” that includes educational and work status, financial and insurance profiles, and social attractiveness (38). “Winners” can navigate complicated medical systems to gain assistance coping with disease, but “losers” may be uninformed about even the availability of services and resources. Patients disadvantaged by socioeconomic factors may experience restricted access to rehabilitation, but absent an explanation for this rejection, they are unlikely to challenge selection decisions effectively.

**Recommendations**

Potential for injustice inheres to a system that fails to provide clear criteria or standards for those who make difficult selection determinations. Although screening is necessary to assure that patients are medically stable and to demonstrate that they have remediable functional disabilities, the screening process has significant shortcomings. Practitioners should standardize selection guidelines and formulate them in writing. Reasons for rejecting patients, and a commitment to reevaluating those patients in the future, should be assured. Otherwise, a single decision made relatively early in the course of disease may preclude rehabilitative care entirely. A mechanism for patient appeals could provide valuable checks and balances to the selection process.

**Goal Setting for Individual Patients**

In order to develop patient treatment plans, staff members discuss the requirements of the postdischarge environment with patients and their family members. Initial treatment goals are outlined, reviewed, and adjusted during the course of therapy to ensure the relevance and feasibility of patient goals over time.

---

**TABLE 20.1 Clinical Practice to Rehabilitation Care**

<table>
<thead>
<tr>
<th>Clinical Practice Issues</th>
<th>Policy Issues</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Selection of patients</td>
<td>1. Allocation of resources</td>
</tr>
<tr>
<td>2. Goal setting for patients</td>
<td>2. Medical insurance and implications for rehabilitation</td>
</tr>
<tr>
<td>3. Patient-practitioner relationships</td>
<td>3. Professional responsibilities</td>
</tr>
<tr>
<td>4. Professional and team issues</td>
<td></td>
</tr>
<tr>
<td>5. Duties and rights of family members</td>
<td></td>
</tr>
<tr>
<td>6. Quality of life and termination of treatment</td>
<td></td>
</tr>
</tbody>
</table>
Authors who have addressed goal setting during patient care include Trieschman, who emphasizes the importance of consulting with patients and family members to cast and refine goals (22, 24, 25, 39–42). She cautions caregivers to refrain from imposing goals on patients who may reject them in the long run anyway and stresses the importance of assuring transfer of skills from the rehabilitation setting to the home environment (42). Becker et al. emphasize the need to promote close interaction among family and staff members and to assure that the full team resolves discrepancies among treatment goals (39).

Problems of Current Practice

Although rehabilitation practitioners encourage patients to participate in designing their treatment program, this may be difficult for some patients. On admission to rehabilitation, many patients are exhausted by pain, weakness, fatigue, depression, or anxiety. Few come to terms with new or exacerbated disability and the demands of their condition absent real-life and home experiences that could illuminate their post-discharge needs (25, 43). They may know little about what can be achieved from rehabilitation. The rehabilitation unit itself may be unsettling. Expectation to socialize with strangers who may have visible scars and dysfunction at the very time that visits from friends and family are restricted by a demanding therapeutic schedule may cause patients to feel cut off from all that is familiar.

Values

People have different interpretations of data about probabilities and outcomes. One's approach to information regarding risks and benefits, pain, cost, health, and disability is influenced by his or her personal values. It is not uncommon for patients, family members, health care teams, and insurers to advocate for discrepant or even mutually exclusive goals. Patients are likely to want to make their own decisions; they know best what is personally meaningful and how exhausted or disheartened they feel. At the same time, they may feel indebted to family members who want to pursue different goals. Relatives may believe that their opinions should take priority since they will be serving as caregivers. Practitioners’ experiences with disability may incline them to usurp decision-making power rather than to accept patient or family choices that appear likely to expend time, money, or effort unnecessarily. Tensions abound (25–27).

Recommendations

The principle of autonomy holds self-determination paramount regardless of whether patients make ill-advised choices. We know that patient autonomy can be compromised by limited knowledge of medicine. But the technical expertise possessed by practitioners does not imply moral authority. Professionals should refrain from imposing their personal values on patients who generally know best what makes sense for themselves. Experience demonstrates that in an unrestricted environment, former patients may refuse tasks that require extensive time, patience, or concentration. They may discard unattractive or cumbersome equipment and ignore recommendations for self-care and home exercise programs. Transportation, social networks, or finances may be insufficient to support interventions made in the rehabilitation setting (25).

Practitioners may not appreciate the potency of their recommendations for patients who have experienced profound, often abrupt, loss of physical ability. Patients require education about the costs, risks, and effectiveness of treatment options, and benefit from an opportunity to explore fully the medical and functional ramifications of alternative treatment approaches (44). Clinicians should communicate the values, ethical norms, and institutional priorities that underlie their recommendations. Tauber believes that only sustained, deliberate efforts to elucidate values held by all interested parties and to delineate major conflicts among those values help to resolve the moral tensions that impact care of disabled persons (45).

Patient-Practitioner Relationships

Relationships between rehabilitation patients and care providers are likely to be of long duration, in contrast to some acute care relationships (15). The nature of the moral rules and principles that determine exchange of information and provision of services in extended health care relationships bears exploration.

Contractual Model

American laws and customs have evolved over time with respect to patient rights. An increasingly egalitarian relationship between patients and practitioners has displaced paternalism (17). A “contractual” model of care requires practitioners to tell patients the truth about their circumstances and to present options in an accurate and balanced manner. A professional duty to act beneficently toward patients is constrained by respect for patient autonomy. Physicians supply the medical care desired by autonomous people who make informed decisions about that care. Respect for confidentiality and patient privacy is deemed essential to the development of trusting and egalitarian relationships between patients and physicians.

Caplan describes factors that compromise the relevance of a contractual model to rehabilitation (15). Relationships in rehabilitation are multifaceted rather than restricted to a single practitioner and patient. Patients have multiple health care providers, and family members, too, often assume an integral role in treatment. People who have experienced profound impairment need time to adjust to the reality of disability. Many are anguished by their functional losses. Facing an uncertain future and possessing little knowledge of ways that disability will impact their life choices, patients may feel unprepared to make important decisions. Similarly, family members are likely to be unfamiliar with issues that arise in the context of their loved ones’ impairments.

Competency and Recommendations

Caplan describes the complex and evolving nature of relationships among patients and their rehabilitation providers.
He argues that an effective and proper role of physicians in medical decision making may entail more guidance on the part of professionals than a model that specifies the physician as a provider of information to an autonomous patient (15). He notes that the competence of patients during the earliest phases of rehabilitation treatment may be questionable, and contends that rehabilitation professionals are justified in overriding the autonomous wishes of patients at the outset of treatment (17,22,26,27). If patients have not yet adapted to impairment or come to understand their future possibilities, Caplan suggests that practitioners are justified in using persuasion to encourage even reluctant patients to participate in early rehabilitation. He recognizes that this approach tolerates more beneficence on the part of providers than customary in contemporary medicine, but suggests that over the long run, it restores patient identity, capacity to cope, and autonomy (15). Caplan characterizes this as an “educational model” of care. He underscores the importance of earning the understanding and cooperation of patients at the same time that practitioners seek to understand patients’ values and preferences as fully as possible.

Periodic assessment of patient capacities to make autonomous choices is essential. Practitioners must focus on restoring autonomy quickly so that patients will resume decision making as they adjust to the consequences of impairment. An independent committee could be charged with assuring that patient autonomy is enhanced as soon as possible during treatment. Paternalism would be considered appropriate only as a bridge to restoring autonomy to patients, with practitioners recognizing that beneficence, however well intended, may compromise a patient’s best interests over time (17).

Assessment of competency in patients with neuropsychological dysfunction requires presenting information in modalities and formats that patients are able to process. Simple tests such as the mini-mental status exam may be inadequate to discern patient capacity. More comprehensive evaluation, including determining specific competencies in appropriate contexts over time, may be needed to understand patient decision-making capacity. In the event that patients and caregivers disagree on a course of treatment, the values underlying those disagreements should be discerned and discussed thoughtfully, rather than providers assuming that a patient is incompetent. Competent patients understand potential benefits, risks, and consequences of treatment options and can communicate their decisions (44,46,47).

**Professional and Team Issues**

Rehabilitation treatment is delivered by a multidisciplinary group of professionals who work together as a team. Amelioration of patients’ functional deficits, psychosocial challenges, and vocational needs requires rehabilitation teams to provide coordinated and comprehensive treatment generally not offered by individual caregivers who work independently. Experienced teams have demonstrated the ability to provide efficient, organized services (19,22).

**Moral Problems of Teams**

Patients and family members may be relatively unaccustomed to working intimately with so many practitioners in the healthcare setting. Distinguishing the responsibilities and lines of authority among team members may be challenging. Clarifying what each professional expects and needs to know is not easy for many patients. Acting consistently while working with a wide variety of professionals can be daunting for them as well.

Members of a rehabilitation team should recognize the vulnerability of patients to team pressure, no matter how subtle. Patients who are exhausted, frightened, or confused may be intimidated by professionals. Purtilo suggests that patients may feel compelled to follow recommendations with which they do not agree when outnumbered by the team (19). Sometimes team members need to decide how to treat patient “secrets” or how to reconcile conflicting information gathered from patients and relatives.

Practitioners serve as teachers and guides to enhance patient function and assist adjustment to disability. Addressing the needs of many patients simultaneously requires the team to balance the conflicting interests of multiple patients. Institutional policies about patient schedules, smoking, or electronic devices, for example, may clash with the wishes of particular patients. The team must balance the interests of individuals with those of the collective and the guidelines of the institution (17).

Controversy about authority and responsibility can arise among team members as they recommend diverse and perhaps conflicting goals for patients or disagree about setting priorities among consensus goals. Certain team members may wish to work with patients in ways that conflict with those of their colleagues (48). A team in which authority is distributed may not have agreed on a method to resolve dissension. In addition, the fact that team members share lengthy and strenuous work hours may impose a sense of loyalty that makes some members uncomfortable questioning others.

**Recommendations**

Purtilo suggests that teams explore and clarify personal and shared values in order to develop a “common moral language” to frame ethical decisions (19). Team members should be trained about and practice effective team dynamics, and should work within a framework of administrative guidelines that promote airing and resolution of conflicts (17,32). Accountable to the entire team for their actions, each member should be encouraged to raise questions about professional behaviors that seem to compromise patient interests. Team members must assure their accessibility to patients and families and strive to avoid intimidating them inadvertently. Even patients or families who are difficult to manage must be treated respectfully.

Professionals should explain to patients and their relatives how teams share responsibility and designate authority (17). Lines of communication should be clarified to alleviate patient uncertainty. Families should know what to expect in terms of decision making. Practitioners should protect patient privacy...
and confidentiality to the extent possible, while emphasizing to patients and relatives the need to share relevant information among those involved in a patient’s care. Teams should strive to resolve the conflicts that inevitably arise among patients, family members, and professionals. They should seek advice and support from other professionals when such insight can advance rational decision making (47). Ethics committees can be asked to provide a neutral forum for discussion and mediation.

**Duties and Rights of Family Members**

Family members frequently assume a vitally important role in the care of disabled relatives. The interest and commitment of relatives may determine whether a patient is admitted for rehabilitation. During treatment, family members are expected to meet with the team to learn caregiving skills, revise treatment goals, and establish postdischarge arrangements.

**Obligations**

There is an expectation on the part of society that relatives will assist one another when needs arise. Family or family-like relationships are considered uniquely extensive and interconnected (17). Family members often undertake caretaking duties with the understanding that they are able to provide special emotional support and affection as well as the physical care that patients require (14).

The need for family caretaking has increased as the extent and cost of professional care have risen (49). Some rehabilitative care has shifted to home settings. Early discharge to home is thought to enhance patient autonomy while enabling patients to test their skills and the feasibility of their goals in a real-life setting. Outpatient or home settings are preferred by many patients and are less expensive than inpatient facilities.

A randomized, controlled trial of stroke patients with moderately severe disability contrasted routine hospital rehabilitation and early discharge with rehabilitation care provided at home. Patients in the latter group showed better recovery of activities of daily living, ambulation, motor capacity, manual dexterity, socialization, and satisfaction. They required half the resources of the hospitalized group and experienced no difficulties in the use of home help. A negative impact on family caregivers was not noted (50).

Callahan has observed that many families discover that providing care to their disabled members is mutually satisfying and rewarding (14). Caretakers develop and hone skills and resources to adapt effectively to shifting demands. They may take pride in their ability to identify patient needs and to provide care sensitively and with compassion. Some family members are exceptionally responsive to a relative’s situation and thus offer optimal care.

Other families experience difficulties, however. Unresolved tensions among patients and caregivers may interfere with satisfactory relationships. The demands of caring for a disabled person may exceed the capacities of some family members. Strain may result from limited financial resources or inadequate physical supports. Relatives may feel angry, sad, or depressed about the condition of their loved one. They may question whether they are prepared to cope over long years with the circumstances of a severely disabled person or one who has an uncertain prognosis. Facing an unanticipated situation that they have not chosen, their personal happiness or autonomy may feel threatened.

**Limits of Duty**

There is no simple formula to determine what family members ought to be expected to give patients, or to identify the limits of duty. Some relatives may gladly dedicate the remainder of their days to caregiving, whereas others may view this as unjustified self-sacrifice (14). Family commitments are complicated by the fact that families today are smaller and more dispersed than in the past, making it difficult to share caregiving tasks among several relatives. Women may feel a special responsibility to become caretakers, but employment often restricts their availability to do so. Some families are not physically able to provide adequate care despite wishing to do so, whereas the financial or emotional resources of others are too meager. Some relatives are unwilling to relinquish personal plans, hopes, or dreams even if they are best qualified to address patient needs and vulnerabilities (14). Health care providers rarely have sufficient understanding about the lives and intimate relationships of specific patients and family members to know what to advise in difficult situations.

**Recommendations**

Practitioners may be uncertain about the amount of persuasion that is justified in an attempt to encourage potential caregivers to commit themselves to patients. When patient needs are minimal and family members poised to sacrifice relatively little, the patient’s best interests may well entail encouraging relatives to fulfill family obligations. In circumstances of severe disability in which significant sacrifice will be required, however, strong persuasion does not appear justified. Callahan points out that our society neither rewards nor honors people who transcend their own needs to care for others. Such people are, in fact, more likely to meet with social isolation than with commendation and should not be expected to act as heroes (14).

Society has yet to develop mechanisms to reimburse the financial and psychological services that could minimize the burden on caregivers. It is critical that family members are furnished with tools to sustain them, including daycare centers, respite care, counseling, self-help groups, and adequate physical facilities (49). Only then can society expect any but the most extraordinary people to embrace an opportunity to care for a seriously disabled relative.

**Quality of Life and Termination of Treatment**

Many factors contribute to decisions to terminate treatment in the rehabilitation setting (17). Patients undergoing rehabilitation care are expected to make steady and measurable progress toward attaining their goals. When progress slows significantly or patients appear to have reached a plateau in
degree of improvement, members of the treatment team may doubt whether continued therapy is justified. Such questions about efficacy of treatment are typically raised first by professionals rather than family. At times, patients and their relatives seek goals that professionals no longer consider realistic or important. At other times, clinicians may be concerned that patients who are scheduled for discharge have received insufficient treatment (33).

Whose Values?
Moral values of team members influence decisions to terminate treatment. Practitioners assess somewhat nebulous concepts, such as “benefit,” “productivity,” “functional improvement,” and “integration into society” to delineate end points of treatment (39–41,51). Their subjective judgments about the capacity of patients to cope with impairment outside the rehabilitation setting and the validity of family members’ requests for services and equipment affect their appraisals. Practitioner perceptions about acceptable levels of function do likewise, even though their personal values may differ substantially from those of patients (17). Not uncommonly, insurance providers determine the length and timing of treatment, which may not reflect a time frame that patients or practitioners consider reasonable. It is time-consuming for practitioners to advocate for public policies to assure adequate reimbursement by insurance plans for evidence-based rehabilitation therapy.

Our society lacks a common context that relates ambiguous concepts of health, function, and quality of life to cherished personal values such as autonomy and independence. Medical practitioners do not have the final word on quality of life, nor are their theoretical views—particularly if unexamined—necessarily more insightful than those of others.

Research about the quality of life of persons with disability indicates that adversity does not necessarily cause a person to appraise his quality of life negatively (52). In fact, persons who adapt to disability may be wholly satisfied with the meaningfulness of life (53). However, gaps between one's activities, aspirations, and accomplishments may impair quality of life to the extent that significance is attached to those unattained wishes (54). To assure thoughtful decisions about terminating treatment, practitioners must understand the critical importance of patients’ values in impacting patient assessments of elements that make their lives satisfying.

The progress of some patients during rehabilitation may be examined sooner than that of others. Patients who are considered noncompliant, uncooperative, or poorly motivated may be so difficult to manage that the team discusses discharge relatively early in the course of treatment. Other patients have limited insurance coverage. Professionals may feel pressured to use scarce institutional resources for new patients, and thus decide to wean care from longer-term patients. A patient’s home setting and his anticipated need for assistance at discharge also affect the timing of a decision to curtail care.

In recent years, it has become increasingly common for patients to begin rehabilitation care before they are able to participate fully in training. Because some patients may not be eligible for readmission following a temporary discharge, practitioners may be reluctant to interrupt treatment early. At other times, onset of medical complications results in transfer to acute care, regardless of progress made in rehabilitation. Whether to readmit patients for additional treatment can be a source of disagreement among both patients and professionals.

Recommendations
Even wise and morally sensitive caregivers feel challenged to allocate care wisely among patients. Sometimes, rehabilitation teams fail to explain the criteria that influence their decision to end treatment. At other times, patients and relatives may not appreciate the significance of the factors assessed. Surely, patients and their families have a right to know the parameters used to evaluate patient progress and the guidelines that determine whether treatment will be continued. Practitioners have a duty to document patient milestones in order to assure that objective data underlie decisions to stop care. Engaging patients and relatives in team discussions about termination of care is essential if their desires are to be honored (17).

POLICY ISSUES
Allocation of Resources
Over 125 million Americans live with chronic illness, disability, or functional limitations (55). Babies who would have died from complications of prematurity or congenital abnormality only a few years ago survive, often with significant disability. Many severely injured people overcome life-threatening conditions. Rapid medical triage and substantial technological advance enable victims of severe war wounds to survive. Americans today live considerably longer than their ancestors; by 2040, 23% of the population will be older than 65 years (56). Population growth, extended life span, and successful acute care interventions have rendered chronic disease increasingly prevalent. Persons with chronic disease frequently benefit from rehabilitation services to enhance their functional skills at work, school, or in the home (49).

Costs of Health Care and Limitations of Access
Costs of health care have increased annually since the 1980s. In 1950, approximately 5% of gross national product (GNP) was spent on health care; by 1998, health expenditures were 15% of the GNP, by far the highest level in the world and nearly double that of the nearest competitor nation (57,58). Today’s health spending represents 16% of GNP (56). Even middle-income Americans worry that they may be excluded from necessary health care services because of inability to pay (59).

Our national appetite for medical care increases. Costs are driven by hospital-based services, complex technology, fee-for-service reimbursement plans, litigation concerns, and treatment of patients near the end of life (56). Treatment of severe injury is costly as well; some who survive trauma to the brain or spinal cord require care costing hundreds of thousands of dollars.
High costs of care and generalized economic tightening have restricted access to care. The Census Bureau’s March 2000 Population Survey indicated that more than a third of poor families and a quarter of near-poor families (those with incomes between 100% and 199% of the poverty level) lacked health care coverage. Categorical restrictions and ceilings on income eligibility for Medicaid exclude many employed low-income people; 72% of uninsured persons live in families with at least one full-time worker. Impoverished adults are more likely than children to be uninsured (43% and 26%, respectively), despite the fact that 71% of those who are uninsured have completed high school, and more than a third have had some college education (60).

Studies show that although less healthy people are more likely than those who are healthy to qualify for public coverage, they are less likely to be insured. Minorities are uninsured in higher numbers than the white population. People who live in inner city or rural areas often have restricted access to care. Those who lack care have worse health than those who receive services, even controlling for factors such as increased stress and poor hygiene (61).

Churchill comments that, despite a sense of moral repugnance, Americans do ration health care according to the ability to pay (56). Many uninsured people forego preventive care and basic services such as eyeglasses, hearing aids, and routine dental care. Even desperately needed care is typically sporadic for those without financial resources. The Commonwealth Fund conducted a health survey of 12,000 adults in seven developed countries (Australia, Canada, Germany, the Netherlands, New Zealand, the United Kingdom, and the United States). Americans were most likely to complain of restricted access to care; 37% of those surveyed reported skipping doctor visits, tests, and prescriptions in the preceding year due to cost (62).

Interventions designed to extend life often overshadow those that enhance its quality. The heroism and drama of rescue medicine prevail over seemingly mundane preventive and primary care interventions. Our fragmented, uncoordinated system of private health insurance and publicly funded entitlement programs fails to address the medical needs of many, especially those who have chronic disease. Allocation decisions are made absent a larger context that would help to identify and weigh priorities (9).

Principle of Justice
A just society does not allocate health care to favor those who are insured, wealthy, and white. A morally acceptable health care system must honor principles of equity and justice. We respect a community that offers mutual and reciprocal assistance to its members and honors its social obligations to care for those who are sick without discriminating among individuals. After all, no one is immune to unexpected disease or calamity—misfortune can, in fact, strengthen human bonds of mutual affection (56).

But what services should be offered in a society that cannot afford everything? Principles of utility would direct us to select services that provide the greatest good for the greatest number of people (2). Principles of justice imply that services should be based on need. Yet, assessment of need may be subjective, biased by highly individualistic desires, hopes, and preferences (9).

In an environment of finite resources, limits on health care should derive from generic guidelines that apply to common conditions. Personalized appeals for specialized services should be discouraged (56). Access to basic and primary care improves health outcome; use of more specialized care should depend on evidence regarding its effectiveness and an understanding of the efficiency with which it can be provided. Costly care of marginal benefit should be replaced by treatments that improve quality of life.

Some health care planners have concluded that poorly distributed medical services may compromise not only the health of those persons who have inadequate access to care but also those who receive excessive services. A 2004 federal study showed no differences in health outcomes across America, regardless of substantial disparities in health care spending among regions. For example, per capita spending in Utah was just 59% of that in Massachusetts, yet population health ratings were similar (63).

In a comparative study of patients conducted in Minneapolis, Miami, Portland, Oregon, and Orange County, California, neither life expectancy nor quality of life in the last 6 months of life was improved by spending more on health care (64). Expenditures in Miami far exceeded those in Minneapolis, yet medical outcomes were similar. During the 6 months prior to death, Medicare costs in Miami were double that of Minneapolis. Miami patients visited medical specialists six times more frequently than those in Minneapolis, spent twice as much time in the hospital, and were admitted to intensive care units more than twice as often. Lifetime spending for a typical 65-year-old in Miami was $50,000 more than in Minneapolis. But no difference in outcomes was noted. The Congressional Budget Office estimates that less than half of all medical care in the United States is supported by sound evidence of its effectiveness (7). In fact, hospitalized patients may actually do worse than those who receive less treatment and experience fewer infections and other complications of treatment.

Recent Evolution of Medical Insurance and Implications for Rehabilitation
Despite benefiting from the most advanced medical technology in the world, Americans worry that such care may one day exceed their economic reach. The last major change in federal health care policy was the 1965 Medicare Act. In 1971, President Richard Nixon tried to forestall the ensuing national interest in single payer health insurance by proposing federal mandates for employer coverage and a Medicaid-like program available to those who were otherwise not insured with income-based sliding scale premiums. Although federal legislation with these provisions was never passed, increases in health costs over the next two decades drove states including Massachusetts, Vermont, Oregon, Minnesota, and Tennessee
to implement mandates for coverage. Such programs, however, proved costly and troublesome to implement as legislators repeatedly failed to enforce the mandates or to finance new coverage for the poor. Even those persons who faced fines were unable to purchase health insurance if they did not receive subsidies. The number of uninsured Americans continued to rise (65).

By the early 1990s, concern about inadequate health care coverage prompted President Bill Clinton to appoint a task force to draft reform measures ensuring universal and affordable health care. Although those reform efforts failed, the focus on health care costs propelled changes in systems of reimbursement, including managed care, that were intended to control medical costs.

Rates and duration of hospitalization decreased thereafter, but overall medical costs continue to grow. No reform effort included real cost controls (56,61).

Reliance in America on private health insurance continues to make universal coverage unaffordable, as a relentless rise in costs drives up insurance premiums. A 2008 report of the Commonwealth Fund indicated that although the United States spends more than twice as much per capita as other industrialized nations on health care, it ranks last in preventable mortality. Many developed countries, by comparison, have more centralized health systems and public insurance. Even in countries with private insurance such as Germany and Switzerland, more efficient insurance administration reduces costs substantially (66). An estimated $350 billion, or nearly 30% of health care expenditures, is spent annually on American medical bureaucracy (66).

The 2008 Commonwealth Fund report indicates that 75 million Americans are believed to lack any or adequate health insurance. The economic expansion of 2000 to 2007 created no increase in real household income. As premiums have increased faster than wages and inflation, more people have sought insurance from public programs such as Medicaid and the State Children’s Health Insurance Program. The number of people covered by private insurance or employers continues to decline (66,67). A national survey of employers revealed skepticism regarding employer mandates to fund insurance for uninsured persons through employer health care plans and payroll taxes. A small business lobbying group has announced opposition to Medicare-style single payer financing (68). And, despite substantial public pressure during the fall of 2007, Congress was unable to override President Bush’s veto of legislation to expand the State Children’s Health Insurance Program from 6.6 million children to more than 10 million uninsured children.

One notable exception to the failure to provide health insurance to all citizens is a program currently under way in Massachusetts. In 2006, the Republican governor and Democratic-controlled legislature enacted far-reaching health insurance reform. All residents must subscribe to health insurance or suffer a tax penalty; employers must offer coverage or pay a small assessment if they do not do so. An expanded state/federal Medicaid program is available to low-income individuals, who otherwise receive subsidies to help purchase private insurance offered at substantially lower rates than in the past. By mid-2008, two thirds of those who lacked health insurance at the outset of the program were covered, including 40% of whom now had private policies without using government subsides. The cost of the program per person, shared between the state and redirected federal funds, has been less than anticipated, although unexpectedly rapid enrollment has increased the total budget for subsidized care (69).

Concerns about access, cost, and quality of health care are once again receiving national political attention. Throughout the 2008 presidential campaign, the public voiced strong support for guaranteed health care after presidential hopeful John Edwards proposed a plan for universal health insurance in February 2007. Both major political parties responded to public interest.

Some politicians recommended modeling health care reform on new public policies, and others advocated expanding the Federal Employees’ Health Benefits Program, which currently provides health care to 8 million federal employees through an array of private and public insurance plans.

Both nominees for president in 2008 proposed health insurance plans. Democratic nominee Barack Obama emphasized mandates for employers to provide health coverage for employees or pay a tax and proposed subsidies and tax credits for small businesses and individuals, regulating the insurance industry and enhancing competition among payers. Republican nominee John McCain proposed a market-based plan predicated on individuals purchasing insurance with universally available tax credits or refunds. The existing employer tax advantage for offering health insurance would be cancelled, calling into question whether employers would continue to offer such plans (70,71).

Election-year interest in health care access, quality, and costs continues. A panel of experts on medicine, academia, business, insurance, and politics was convened by the Massachusetts Medical Society and the New England Journal of Medicine in May 2008 to address physician dissatisfaction with primary care. The experts noted that the current reimbursement system rewards physicians and institutions for sophisticated technological procedures rather than for cognitive work and time spent with patients. They discussed the rewards that other countries have realized from investments in electronic systems and other health-related information technology. The panel underscored the need for the academic health community to focus on addressing the complex economic research and policy questions concerning improved access and quality of health care (72,73). At this writing, it is unclear whether public demand for postelection health care reform will succeed in countering entrenched industry interests.

**Professional Responsibilities**

**Professional Duties**

Professionals who act as “gatekeepers” to managed-care services are torn between containing medical costs and advocating for
resources likely to benefit individual patients. Seventy-five percent of physicians surveyed were concerned about the ethics of financial incentives intended to restrain testing, treatment, and medical referrals. Most of those surveyed—87%—believed that physicians should be empowered to discuss limitations of treatment coverage with patients (74).

Physicians worry that balancing the care of individual patients with stewardship of collective health care resources impairs their duty of loyalty to their patients. Many feel that their patients have lost trust in them over time (74). Caplan agrees that gatekeeping at the bedside undermines patient ability to trust caregivers. He advocates instituting public policies to assure ethical solutions to limitations on resources for patient care (75). Standardized treatment guidelines should be coupled with provisions that enable patients who find their treatment alternatives unreasonably limited to appeal those decisions (76).

Severely disabled patients may be particularly vulnerable to the constraints of managed-care systems. Patients with multisystem deficits who lack information about the potential benefits of rehabilitation care may be unable to overcome a bias to limit their treatment (76). Restricted access to care may fail to remedy characteristics of disease and disability that lead to social disadvantage. In a society that fosters individualism above mutual obligations to others, persons may be at particular risk if they are deemed incapable of contributing fully to society. A just health care system would not compromise quality and length of treatment for those who have chronic impairment (77).

Rehabilitation professionals have a duty to maintain personal and professional standards of competence in their field. Appropriate behaviors for professionals in many clinical fields are defined by discipline-specific codes of ethics. The commercial interests of health care providers who have complex business relationships with pharmaceutical, assistive device, and medical equipment companies invite increasing concern. Weber et al. describe the risks posed by pharmaceutical and biomedical businesses that provide gifts, free samples, equipment, medical devices, educational and informational programs, dinners, and retreats for medical professionals (78). Research shows that such relationships impact physicians’ prescribing behavior (79). Commercial sponsorship of research, equity ownership in biomedical companies, and consulting relationships can create a culture of entitlement among providers. Restrictions on industry relationships and requirements for disclosure are becoming common in medical schools and teaching hospitals (80). Professional codes of ethics and business conduct must be enforced to protect the interests of patients and the integrity of professionals in a landscape of commercial interests (81, 82).

Reports indicate that treatment provided in certain “joint ventures” is more costly than that delivered at similar facilities not owned by doctors. A study of doctor-owned physical therapy facilities in Florida noted a lower quality of care, fewer licensed therapists, and shorter patient treatment sessions than occurred in nonphysician-owned facilities (83).

Other professional interfaces may raise concerns (17). Dynamics within the treatment team can hamper the effectiveness of practitioners. Appropriate safeguards enable providers to question the recommendations or conduct of their co-workers or that of colleagues in other institutions. Individual team members may need support in disputing team norms or expectations. Practitioners should be permitted to decline roles with which they are personally uncomfortable, whether for reasons of professional expertise or personal values. They also have a duty to inform patients of treatments that have been recommended for potential benefit without scientific affirmation of efficacy.

Rehabilitation institutions should strive to serve disabled patients within the broader community. Institutions should widen and deepen their community engagement to reinforce integration of disabled persons and society. A virtuous institution conveys morally sound values by its responsiveness to families and broader social networks at the same time that it assures skillful care of patients in a fiscally responsible environment (84).

**Prevention Efforts**

At times, disability results from preventable accidents or unwise lifestyle choices. For example, alcohol use and excessive speed may precede traffic accidents, and road injuries are more serious when seat belts or car seats are not used. Absence of helmets worsens head injuries in motorcyclists and bicyclists. Firearms produce severely disabling injuries.

The knowledge that so many of their patients have sustained disability that was preventable should mobilize rehabilitation practitioners to advocate public policies to enhance injury prevention, including gun control, helmet laws, and severe penalties against drunk driving and excessive speed. Although balancing individual rights and responsibilities with needs of the larger society is always challenging, it would seem that practitioners have an obligation to advance public policies that prevent needless disability. Certainly, rehabilitation practitioners are able to vividly describe the devastating long-term ramifications of spinal cord injury and brain injury.

**PEDAGOGIC ISSUES**

Medical institutions and professional societies now include ethics topics in their educational programs. If ethics education is to be considered as important for rehabilitation students, as it is for students of other disciplines, faculty members need resources and time to familiarize themselves with the study of ethics. Certain professions and specialties (such as nursing, family practice, and internal medicine) have introduced an ethics requirement into their certification requirements. Rehabilitation accrediting agencies could do likewise (17). As the number of knowledgeable and committed instructors increases, rehabilitation practitioners might add formal certification requirements in ethics for students in specialty training programs.
Continuing education in rehabilitation should emphasize the study of ethics. Some rehabilitation institutions have initiated ethics grand rounds, and others have developed ethics committees similar to those in acute care hospitals to explore issues and advance understanding through discussion and educational workshops. A Scholars Program at the Rehabilitation Institute of Chicago and University of Illinois trains selected clinicians over the course of a year on traditional medical ethics, disability studies, and practical clinical topics for a certificate in disability ethics. Monthly interdisciplinary disability ethics seminars, quarterly grand rounds, and in-house ethics newsletters are part of this work (32,85). Similar programs will increase the ethical literacy of professionals.

Journal editors can encourage examination of clinical case studies and scholarly writing about policy aspects of rehabilitation. Organizers of medical ethics conferences could sponsor symposia and panel discussions on intriguing rehabilitation topics. Rehabilitation professionals could collaborate with community organizations to enhance discussion about ethical care. Education about the ethical challenges that confront patients, families, and practitioners has been noted to be invaluable for advocacy groups, institutional trustees, staff, patients, and family members (17).

CONCLUSION

Rehabilitation practitioners will continue to confront significant moral challenges in the coming years. They must strive to ensure excellence of patient care predicated on sound scientific research and must assure that patients are treated with compassion and respect in an era dominated by financial strain and rapid technological innovation. The extent to which providers accept differences among people and the manner in which they attend to and understand patients and families are critical in today's complex and intimidating health care environment. Providing reassurance and comfort in conjunction with competent care has never been more important.

Health care practitioners have a duty to recognize and address inequities of our current medical system. As rehabilitation practitioners examine the quality and availability of medical resources, they must respond to society's failure to provide millions of Americans with access to basic medical resources. Rehabilitation clinicians should identify conflicts of interest in their practices and set exacting standards for professional conduct in the face of business opportunities and other temptations. Practitioners have a responsibility to ponder the role of medical rehabilitation in an era of limited resources. Conscious of the fact that some medical needs will remain unmet in a society that has other important needs, they must recognize and limit care of marginal benefit or excessive cost. They must assist in the development of sound public policy as American society attempts to balance the needs of vulnerable individuals with those of the larger society.

REFERENCES

82. Pear R. Study says fees are often higher when doctor has stake in clinic. New York Times, August 9, 1991:1.
84. Kirschner K. Personal communication (Telephone Conversation, December 10, 2001).
Interactions with the Medicolegal System

Most clinicians in physical medicine and rehabilitation (PM&R) will eventually interact with the legal system. Many medical professionals carry negative feelings about being forced to deal with any type of medicolegal proceeding. Some even avoid clinical work that will more likely lead to depositions. Others feel there is no such thing as a truly “impartial” or “independent” medical evaluation. Much of this emotion arises from unfamiliarity and misunderstanding of the process involved. Although few practitioners relish this aspect of clinical practice, it need not be terribly onerous. The experience should be no worse than neutral if one is fully versed in the details. An overarching goal of this chapter is to remove some negative feelings about this type of work, not by artificially casting it in a favorable light but by desensationalizing the process through familiarization and preparation.

The most common types of medicolegal interactions include independent medical evaluations (IME), depositions, expert witness reviews, malpractice complaints, and issues relating to the structure within which physiatrists conduct their practices (medical staffs, employment and partnership relationships, and managed care provider contracts). Physiatrists often perform IMEs, and even though attorneys may or may not be involved with IME cases, being available to perform them often leads to subsequent deposition, during which attorneys will necessarily be involved. Giving a deposition is quite commonplace and is a medicolegal interaction most of us will encounter. Because this experience is far removed from anything taught (or usually even observed) during training years for most physicians, a detailed review of this process is presented. Expert witness reviews are less common, and a brief discussion will suffice. Optimistically, malpractice complaints will not be encountered often, if ever, and are addressed only generally in this chapter. Practice issues are numerous and complex, and are therefore outlined only briefly.

STANDARDS OF CLINICAL PRACTICE FOR MEDICOLEGAL EVALUATIONS

Medicolegal issues should be considered during the development of sound standards of clinical practice, including principles of review of records, taking the history, performing the physical examination, and generation of the report. Adherence to some commonsense guidelines will prove infinitely valuable if a chart is later retrieved for medicolegal reasons. These issues should not, however, become the driving force behind how a practice is arranged. That is, one should not prepare reports as if their primary purpose is an eventual deposition. Above all, your routine clinical decisions should never be shaped solely by legal considerations.

Review of Outside Records

A few narrative reports can be read at the time of the examination, but a large file should be reviewed earlier and then quickly scanned just before the history and physical (H&P). Radiographs should be studied if available. It is only necessary to review documents that are pertinent to your examination. For example, although you may note underlying cardiovascular disease, there is no need to detail the corresponding laboratory results, diagnostic test results, and the like. Your opinions in unrelated areas will not be legally admissible anyway, other than how they might directly affect rehabilitation issues. Do not discard records sent to you, however, even if they are peripheral to your assessment. An attorney may want to review every note ever sent to you. If some files are missing, suspicions may be raised that they were discarded for some ignoble reason. The records sent to you for purposes of an IME, that is, not for a treatment patient, should be retained for a minimum of 2 years. There is no universally accepted guideline for the time frame of record retention beyond that, and this may vary by jurisdiction.

Taking the History

As always, it is imperative to be thorough while not becoming immersed in superfluous detail. The proper degree of completeness allows for full reproduction of the history months or years later, when direct recall of the patient is impossible. A chronological sequencing of events will facilitate later review. The line of questioning should be orderly: onset of the problem, ongoing and current symptoms and how they are changing, diagnostic testing and the patient’s understanding of test results, past and present treatment efforts, history of similar problems, and pertinent, unrelated medical history. Your report might be the only documentation of an occupational or functional history, including the relationship of symptoms or job status to intercurrent events. Separate the effects of work-related and non–work-related injuries as much as possible. Include details of the effect of the patient’s problem as it relates to physical, psychological, social, and vocational issues, where applicable.
The Physical Examination

Third-party payers, defense attorneys, case managers, and others often balk at a “diagnosis” based solely on subjective patient reports without objective abnormalities. Where possible, generate a set of observations that are as objective as possible to counter future attempts by an attorney to dismiss all subjective reporting as unreliable. A common example is the distribution of trigger or tender points in a patient with a myofascial condition. It is helpful, years later, to have supportive documentation that the tender spots were at characteristic, reproducible locations described in textbooks (e.g., upper trapezius, levator scapula insertion, lateral epicondyle). On the other hand, if the patient reports a nonphysiologic pattern (e.g., severe tenderness with minimal pressure over every area examined), your description should be clear on this point. This type of reporting justifiably lends an objective air to what is inherently a subjective topic, that is, the reporting of pain.

Similarly, it will be more helpful to document joint or spine ranges of motion in terms of percentage or degrees rather than with modifiers such as mildly or moderately reduced. However, do not think that endless goniometric joint motion or computerized strength measurements or other data will by themselves lend additional credence to your report. A medicolegal report is useful because of the clear observation and objective reasoning used in reaching conclusions, not because of an overabundance of data.

When a physical examination finding is described as positive, specify for what symptom or sign it is positive. To illustrate, a straight-leg-raise test might cause pain. Was the pain shooting from the low back to the posterior thigh and then to the ankle, at 45 degrees of leg elevation, with an electric quality, or present at 80 degrees for only a mild pulling sensation in the hamstring area? The same pertains to Spurling’s maneuver, Tinel’s sign, and many others.

The Report

In writing the IME report, after the introduction add a passage such as, “Mr. Smith was told that no treatment would be provided, and that the report would be sent only to you. He understood these points and proceeded.” This affirms the inherently different nature of this visit, in that a doctor-patient relationship has not been established.

The next section should note the origin and content of outside records. The pertinent findings can be cataloged either by date or content (e.g., summarize all clinic notes in order, then scanning studies, then electrodiagnostic findings). It is unnecessary to routinely provide details about each note as to date, content, or conclusions. It is acceptable with acknowledgment to refer to an outside record as a summary of prior test- and/or treatment. For imaging studies, your report should mention whether you reviewed the films and the radiologist’s report, the films only, or the report only. Keep in mind that a cross-examining attorney will be quick to point out that you are not formally qualified to interpret such studies, that is, that you are not a board-certified radiologist. If the actual films were not available, though, the same attorney would certainly attempt to have you admit that it is useful to see the films yourself, since radiologists are not clinicians. There is no need to vary your own standard of practice in such cases, although as a rule, it is preferable to have viewed the images yourself.

All initial and any supplemental reports must be executed with word processing. Handwritten notes are unacceptable other than for brief items such as telephone messages, and perhaps prescription renewals and the like. This allows for legibility after photocopying, faxing, and recopying, as often happens with these documents in the medicolegal realm. A sound practice pattern is to dictate any note regarding direct patient communication. You will often be asked at the time of deposition if your habit is to dictate clinic notes after each patient, at the end of the office day, on the drive home after hours, or even at a later date. It is of course preferable to allow dictation time as part of each office visit to ensure highest recall, possibly eliminating the need for handwritten notes taken during the history.

To the extent possible, the history, physical examination, and your conclusions should be summarized logically and in an objective a fashion as feasible. Avoid judgmental or pejorative terms, such as exaggeration of symptoms, or embellishment of pain reports. These are impossible to support legally, as is the claim that the patient does not appear legitimate. In cases where there are no objective abnormalities, it is preferable to provide a dispassionate discussion about the lack of correlation between the symptoms and disability. For example, “The only impairment on examination or testing is minimal tightness of the left trapezius muscle. This does not explain the symptoms of chronic neck and low back pain, bilateral arm numbness, or the 9 months off work.” Similarly, it is important to outline those impairments that do account for signs and symptoms, such as “C7 nerve root abnormality on the right is confirmed by the reduced triceps reflex and strength and by the positive Spurling’s maneuver. This corresponds to the MRI scan finding of a right, low cervical disc herniation and to the EMG abnormalities.” This type of thorough recording of the thought process will simplify the chart review and deposition at a later date.

INDEPENDENT MEDICAL EVALUATION

Definition

The usual IME is arranged when one party desires to obtain an expert medical opinion from a practitioner who has never previously examined the patient. Your opinion is then used to help resolve a conflict between the patient and a third party, often an employer or insurance company. The IME request may originate from a company physician or worker’s compensation representative, an insurance company providing worker’s compensation coverage, a case manager for an insurance company, or an attorney representing either side in this conflict.

Although your medical expertise is requested, there is no doctor-patient relationship established under traditional law. Because of that, the IME report cannot be introduced into the legal record unless a deposition occurs. This means that an IME will more likely lead to a deposition than will treatment
evaluations. Another consequence of the lack of a doctor-patient relationship is that the IME physician is effectively shielded from malpractice suits arising from this activity. However, there has been recent discussion to reverse that “immunity.” A recent ruling in Michigan allowed a professional liability case to proceed against a physician who had performed an IME and who allegedly caused harm during the physical examination. Several courts have begun to require physicians who conduct examinations related to employment to notify patients of life-threatening conditions even if they are unrelated to employment.

Ethics of IME

The underlying conflict between two parties mentioned earlier is the genesis of much of the controversy surrounding the medical ethics of IMEs. Take the example of a case manager for an insurance company that provides worker’s compensation coverage for a large manufacturer. It could be inferred that the case manager will remain in good standing only if she attempts to reduce the monetary consequences of the claims she oversees. The logical next step might be that she develops a list of IME physicians on whose opinions she can rely (i.e., who routinely determine that there is no impairment present, or if there is, that it is less severe and of less importance than the patient’s own physician believes, and/or that it is unrelated to employment factors). The converse applies to plaintiff attorneys, who know which practitioners and IME doctors usually find severe and long-lasting disability, even in the absence of objective impairment. The “independent” or “impartial” in IME is thus potentially compromised.

Some physicians have built a particular reputation for favoring either the defense side (companies, company physicians, case managers, worker’s compensation insurance companies, defense attorneys) or the plaintiff side (usually plaintiff attorneys). You could be asked during deposition what proportion of your legal work arises from which side. A physician who performs IMEs exclusively for one side may appear less objective to outsiders. Having said that, practice realities mean that it is virtually impossible to see anything approaching 50% of IMEs from each side. Your reputation will be most well served if it remains crystal clear to all concerned that no one can reliably count on your opinion in advance; your conclusions in a given case will depend only on the particulars of that case.

Perhaps it is now more apparent why some believe this medicolegal world of IMEs and depositions should be avoided. Fortunately, judges, administrative law clerks, and mediators increasingly recognize a “boilerplate” IME and its authors, and will discount or even ignore its conclusions. Even strong-willed participants from both sides eventually understand that such one-sided opinions will not survive scrutiny and may cost more in the long run, as challenges and appeals drag out.

The straightforward, underlying principle for an IME, as with all of medical practice, is to always keep in mind what is in the best interest of the person you are examining. This would seem to sway physicians toward the plaintiff side of matters, since after all we should be patient advocates. In reality, an ethical physician will not be labeled as residing in either the plaintiff or defense camp, attesting to a high level of objectivity and impartiality. Some examples will illustrate this principle.

Case I

A young adult sustains a C6 complete spinal cord injury with tetraplegia from a motor vehicle accident. Months later, the insurance company wishes to settle its financial responsibility with a lump-sum payment. The patient’s attorney does not feel the proposed amount is nearly adequate to cover predictable and potential expenses for a lifetime. As a specialist in spinal injury medicine, your IME input has little to do with the impairment, which is not in question. You agree that long-term expenses are more likely to be met by rejecting a lump sum in favor of ongoing responsibility on the part of the insurance company. This “covers” the patient for unforeseen expenses even decades later and is in his best interest.

Case II

A middle-aged machine operator sustains a back injury at work. A physician recommended by a friend determines that the patient is totally disabled indefinitely, perhaps permanently. Treatment includes three-times-weekly hot packs and ultrasound in the physician’s office. Diagnostic testing is normal, and your IME finds abnormal back mechanics from asymmetric muscle tightness but no additional underlying impairment. It is clearly not in the best interest of this patient (or of his family, of society, etc.) to allow indefinite disability, when a return to full work status would be a simple matter to achieve. Your recommendation is for a directed course of physical therapy to include manual techniques, restricted return to work while therapy is under way, with anticipated unrestricted work duty thereafter.

Impartiality is not difficult to conceptualize or attain. Each case is judged on its merits, with emphasis on objectivity. Reliable information can include your physical examination findings, diagnostic test results, laboratory studies, or even a surveillance videotape. If you are successful in achieving impartiality in all cases, you will develop a reputation that you do not automatically favor one side or another. Some third parties might not refer IMEs to you because of that, but that is business you are wise to do without.

IME Without Treatment

IME without treatment is the most common type of IME referral. You will see the patient one time only, leading to an expert opinion about one or several questions. The question may be as deceivingly simple, such as, “Was there a mild closed head injury?” In the musculoskeletal arena, several questions are usually posed:

• What is the impairment (the physiologic abnormality)?
• If an impairment is present, was it caused by the accident/injury in question?
• Will the impairment be temporary? Permanent?
• What disability arises from the impairment? In other words, what is the clinical relevance of the impairment?
• Have the diagnostic tests so far performed been justified?
• Are any additional diagnostic studies required?
• Has treatment to date been effective and warranted?
• What additional treatment is needed, if any?
• Has the patient achieved maximal medical improvement (MMI)? That is, would any further, formal medical input improve the impairment or the functional deficit arising from the impairment?
• Is the patient able to return to work with or without restrictions? How long will restrictions last, and will they be affected by your recommended treatment?
• In some states, what is the permanent impairment rating?
• For some cases, for example, after auto accident, is assistance with household activities needed? Is attendant care needed, and for how long?

IME with Option to Treat
This is a less common type of referral. The evaluating and treating physicians hold quite different roles in the medicolegal world. That aside, there are times when the referring source arranges an IME soon after an injury because of uncertainty about whether an impairment exists. If you then identify a treatable cause of the symptoms, you might be asked, and it might make sense, to accept the treatment role. This occurs mostly when there has been no treatment so far, or at least no beneficial treatment. IME with treatment option usually implies that a less than typical adversarial component is present. Problems arise when the patient does not respond to your treatment as expected; in that case, the referring source might need to arrange yet another IME.

Serial IME
An initial IME does not typically allow the physician to oversee testing or treatment. It can be frustrating to witness inadequate or improper care without the opportunity to rectify the situation. Some third parties will ask for follow-up IME(s), an intermediate step between opinion only and treatment. For example, you might determine with the first examination that only restricted work is possible for 6 weeks while the treatment you outlined is undertaken. The company could then request reevaluation at 6 weeks so that you can decide if unrestricted return to work is now feasible. Some patients are returned for numerous IMEs, especially if the recommended treatment has not been provided. Remember that your recommendations are not binding on the examinee or the treating physician. If a pattern of ineffective treatment persists, it is then incumbent upon the third-party payer to deny funding for any further such interventions.

DEPOSITION
Definition
A deposition is the process of taking your testimony, under oath, as part of a formal legal proceeding. It is usually a substitute for your appearance in the courtroom, and as such could be thought of as a favor to you to keep you from having to disrupt your clinic schedule for court testimony. However, providing deposition testimony does not preclude a later court appearance.

Whatever you say in a deposition carries the full weight and responsibility of being in the courtroom itself, and you should handle it with equal respect. Although your examination results appear in a dictated report, the deposition allows for expansion and clarification of the content of your reports. As a physician, you will usually offer deposition testimony as an expert witness. A critical point to remember is that a medical expert witness may offer opinions, whereas a lay witness (nonexpert) may only testify as to facts, without lending interpretation.

A deposition is “called” (i.e., arranged) when a case has progressed (or regressed) to the point of legal action. These are typically civil cases involving insurance benefits—worker’s compensation, disability, automobile—or other payback carrying a perceived value to both parties. The fact that a deposition has been requested means that the differences of opinion have not been rectified during initial attempts at negotiation, compromise, and/or mediation.

If your testimony is needed for the courts, you might wonder why you are not simply served with a subpoena to appear there, as would any other witness. The answer, based on years of tradition, is that the attorney who wants your opinion does not want to inconvenience you. A courtroom date is subject to repeated cancellations and delays, a pattern that does not mesh well with running a medical practice. A physician could even turn into a “hostile witness,” whose answers are a surprise to the attorney. Depositions are therefore typically scheduled in your office, at your convenience, although standard practices vary by state and region.

Scheduling the Deposition
Physicians are generally permitted by court rules to be compensated for their time in preparing for and attending depositions. Once you have experience with depositions, you might realize that it is wise to outline specific scheduling criteria. In the author’s practice, over 50% of scheduled depositions never take place and are canceled without regard to the havoc wrought on your clinic schedule. They are often canceled because attorneys use an upcoming deposition to “encourage” the opposing party to settle the case out of court.

Take the time to draft a form letter that is sent to an attorney’s office whenever deposition time is requested. The letter should outline your office protocol, including at least

1. The fee. A 1-hour minimum fee is common, perhaps with 15-minute prorated charges thereafter. If your practice leads to many brief depositions, for example, to review an EMG study only, you could consider a minimum fee based on a shorter time. The time for which you bill will include record review, conferences with the attorney, and actual testimony.

2. Prepayment. If your practice experiences frequent non-payment of deposition fees after the fact, consider the

Prepayment
following: your office will not schedule the deposition until prepayment of 1 hour’s fee has been received and has cleared at the bank. This guideline may seem harsh and could be needed only temporarily with certain legal offices, if at all.  

3. **Time allotment.** It is impossible to predict with any accuracy the duration of a given deposition. Perhaps the most reliable indicator is the personalities of the opposing attorneys, which of course you typically do not know in advance. Generally, 45 to 90 minutes suffices, and it makes sense to schedule for the end of the day. Make sure your office confirms the deposition that morning, just in case it has been canceled or postponed without you having been notified.

4. **Cancellation policy.** The rationale for a cancellation fee based on the amount of notice is the likelihood of filling that slot with patient appointments. If a deposition is canceled, the only effective consequence is financial. The fee should be larger when short notice is given. If five or more business days’ notice is given, for example, there may be no fee.

You can easily defend your fee and cancellation policies. The large block of time set aside for a deposition displaces revenue-generating activity. Although it is possible for an attorney to become so upset with these guidelines that a subpoena is issued for your court appearance in person, this is highly unlikely.

---

**The Day of the Deposition**

On the morning of the day of the deposition your office should verify that the deposition is still taking place, as noted earlier. Allow about 15 minutes to review a small file and to meet with the attorney who requested your testimony. If the file is large, set aside additional time for its review. You should be familiar in detail with the chart material you generated; any of it is subject to exhaustive questioning. Take the time to highlight, underline, or flag pertinent passages, and keep these reports immediately available during the deposition. It is unnecessary to review in tedious detail any chart content beyond what you are directly responsible for (office notes, electrodiagnostic testing, outside test results that you ordered, physical therapy progress notes). Attorneys will typically not question you about reports from other clinicians, which are legally hearsay to you.

Your meeting with the requesting attorney before the deposition is to review the process and the general tone of your responses to typical queries. It is not to plot strategy or to engender an “us versus them” mentality. You should be informed about test results or evaluations concluded after your interactions with the patient, in case these are raised during testimony.

A small number of depositions are canceled after this “preconference” if it is apparent that your responses will not support the requesting attorney’s case. Assuming the deposition proceeds, the lawyer who was not at the preconference might ask you about details you discussed, or if you arranged what your answers would be. It is wise to understand that you should not reach that level of specificity in a deposition preconference.

The typical deposition involves four people: the clinician, two attorneys, and a court reporter either typing along or speaking into a soundproof device. There could be three or more attorneys, each representing a different interest. For example, a patient may claim two work injuries and the employer might have changed compensation insurance carriers between the events. The patient is rarely present.

Some depositions are video-recorded. You will sit across from the camera, which will focus only on you regardless of who is speaking. The video technician becomes an additional participant, although sometimes replacing the court reporter.

To begin the deposition, you will be sworn in by the court reporter. The attorney who scheduled your testimony will make introductions and then likely briefly review your curriculum vitae and training/certification to establish you as an expert. You might be asked to provide an overview of the field of PM&R—be brief and use lay terms. Learn to slow your rate of speech and enunciate.

The main point of the deposition follows. The requesting attorney will take you through the direct examination. You will probably be asked how you came to examine this person, where the examination occurred, and about the dictated reports you generated after each visit, which will then be entered as evidence. Finally, you will be questioned directly about specific points of history, physical examination, test results, and your conclusions. Do not hesitate to refer to the file; no one expects you to remember details of an old case. If you have prepared properly, your responses will be prompt and confident. Your testimony should elaborate on yet be consistent with your notes. Inconsistencies will be apparent and will not reflect well on you.

If a topic falls out of your areas of expertise, simply say so.

A cardinal point during deposition is that the physician is not on anyone’s “side.” You are there to represent a professional opinion, not as an advocate for one attorney or the other, the employer, or an insurance company.

As mentioned earlier, all courtroom protocol holds here as well. After one attorney poses a question, the other might object and then give reasons for the objection. Once the objection has been noted, you will be asked to answer the question anyway. Whether the jury eventually hears the question and your response will depend on how the judge later rules on the contended points.

You might be asked to comment on a “hypothetical” situation, the details of which will closely resemble the case at hand. Objections are universally raised to hypotheticals, citing lack of foundation, and that the proposed scenario will not stand up at time of trial. As always, wait until you are instructed to answer, and then do so.

When the first attorney has finished, the opposing party starts cross-examination. Only points brought up during direct examination are meant to be open for cross-examination. This alternating direct, cross, redirect, recross questioning can theoretically continue indefinitely, although in reality there are rarely more than two or three such cycles, of decreasing length. A common tactic here is for the cross-examiner to ask you questions that are said to require only yes or no responses.
This is usually an attempt to keep particulars of the case that are not felt to be favorable off the record. As an expert, you are there to provide opinions, not simple yes or no answers. Feel free to respond as fully as you feel is needed, even if the questioning lawyer objects.

It is natural to feel defensive at times during a deposition, usually during cross-examination. You might be asked why certain aspects of physical examination or diagnostic testing were not performed. Keep a cool head during the entire episode. These are tried-and-true tactics to put you off guard and perhaps to lead you away from an opinion that one lawyer would rather not hear. You should make every attempt not to take the proceedings personally. The attorneys will sometimes attack each other like worst enemies and then, after the deposition, plan where to have lunch. Only occasionally is true animosity present. Your goal should be to accurately represent your opinions and leave legal maneuvers to the attorneys.

**EXPERT WITNESS REVIEWS**

You might be asked to review a file to provide your expert opinion about a particular issue. These often relate to possible overuse of physical medicine services (e.g., a practitioner provides hot packs and ultrasound three times weekly for several months to a patient reportedly injured in a car accident). If you agree that passive modalities are not indicated, your report will be used to retrospectively deny payment for those services. Given that the service provider could learn your identity, you must have the confidence and fortitude to perform this type of review.

Another form of chart review occurs with malpractice cases, which arise when a physician has been said to have breached, or failed to fulfill, a duty to the patient, and harm to the patient is the proximate result of that breach. In these cases, expert witnesses provide testimony as to the duty to the patient and its breach—the essential question is whether the standard of care owed by a physician to a patient has been met. Here, it is imperative that the salient issues fall within your areas of expertise. Medical practitioners usually review files on the side of the defendant, with the goal of showing there was no malpractice committed. As with IMEs, though, some witnesses are well known for providing an opinion that passive modalities are not indicated, usually with approval of the department chair.

**ADDITIONAL CONSIDERATIONS**

**Malpractice and Related Complaints**

Most physicians will be served with malpractice complaints during their careers. As noted earlier, these cases focus on whether the physicians have met the duty to their patients and whether any breach is the proximate cause of harm to the patient. Such complaints, even for cases that are quickly dismissed, make it sound as if your entire professional existence is an abomination and the facility where you work would be better off razed. It is never more important to remember not to take this personally. The attorney who authored the offensive passages is acting in the best interest of the plaintiff within the framework of our present system. Never respond to the complaint yourself; never contact the patient or attorney directly, even if you thought there was a terrific doctor-patient relationship before the day’s mail arrived. Apparently, the feeling was not mutual. You will be represented by a defense attorney arranged by your practice, your hospital, or your professional liability insurance company. It is important to contact your insurer and any risk manager responsible for your facility or practice as soon as you become aware of a potential malpractice claim, so that the defense, and possible settlement, can be handled.

It is critical to know ahead of time if your legal defense firm generally believes strongly in proceeding with litigation rather than agreeing to negotiated settlements. Plaintiff attorneys typically do not realize reimbursement unless there is a payoff; that is, their getting paid is contingent on the outcome of the case. The plaintiff attorney’s pocketbook is best rewarded by spending as little time as possible on a case with the largest payment possible. This often means avoiding a trial. Defense attorneys are paid hourly fees, regardless of outcome. It may also be tempting for the defense to settle a
case rather than to litigate it (for the insurance company they represent, some payout now vs. the chance of a much larger one later). It is important to understand, though, that all such settlements appear as negative outcomes for the defendant physician, even if facts point to no wrongdoing. Various federal and state laws and rules require public notification of settlements and adverse decisions when malpractice has been alleged (e.g., the National Practitioner Data Bank records adverse judgments and settlement amounts and is routinely reviewed when application for employment or privileges is made by physicians).

Physicians may also be held liable when they employ or supervise others, such as therapists and nurses. When a physician takes on duties to the patient that do not involve treatment, the other assumed duties may result in claims and liability. Often there will be allegations against several professionals and their facility, after which the defendants may attempt to shift responsibility from one to another.

In managed-care situations, where payers are involved in establishing treatment plans and allocating the “quantity of care,” the liability issues become more complicated, as the quantity of care may be directly related to the quality of care, or at least have an impact on it. These cases often involve jurisdictional and procedural issues, in addition to the substantive questions, and are complex and time-consuming to resolve.

To minimize claims of malpractice and related matters, it is critical for physicians to

- Understand the duties they undertake to patients (such as providing care, supervising others, hiring others who are qualified to provide care in their stead)
- Understand the constraints imposed on them by managed care and payers in general
- Document all health care decisions and actions in the appropriate location (patient chart, incident report, etc.)
- Limit liability by contract when possible (e.g., clarifying in a managed-care contract that the physician has no liability for administrative decisions relating to utilization)
- Insure against claims when possible, coordinating that insurance coverage with the insurance provided by an employer, a facility, or others
- Obtain protection from others, when possible (such as indemnification in a contract where other parties may be responsible for actions that result in claims against the physician)
- Obtain consents from patients when necessary or appropriate (such as informed consent when risky procedures are involved or more than one treatment may be available, and consents to disclose patient information as to health status or other matters)

Abandonment
Abandonment is another potential claim against physicians, in cases where a physician-patient relationship has commenced and the physician terminates it without giving the patient adequate opportunity to locate another provider of similar services. This is often a situation involving non-compliant patients, patients who are disruptive to others, or a patient for whom no source of payment for services is available.

Ascertaining problems before a physician-patient relationship begins is ideal but not always feasible. Establishing rules and an agreement with the patient at the outset of the relationship that allow for termination in certain events may be helpful. In all events, it is essential to consult regulations of the appropriate Board of Medical Examiners and rules of any facility or practice in which a physician sees the particular patient, before acting to terminate the physician-patient relationship. You may be required to make yourself available for a time, perhaps 1 month, after notifying the patient of your intent to end the treatment relationship, in order to handle urgent or emergent matters.

PRACTICE STRUCTURE ISSUES
Health care is a heavily regulated industry—various federal and state laws and regulations pertain to all facilities, personnel, and payment and insurance programs. The legal issues affecting physiatrists and others vary with the relevant practice structure. The following discussion provides a brief overview of certain issues commonly encountered.

Medical Staff/Hospital Privileges
When applying to join a medical staff, a physician should review the medical staff bylaws to become aware of how privileges are granted and how they may be revoked, suspended, or denied. The hospitals and other entities that grant and deal with privileges have the responsibility for ensuring that quality patient care is delivered through the medical staff and that responsibility gives the granting bodies great latitude in dealing with members of the medical staff on issues from quality of practice to behavioral issues (including impairment and substance abuse).

If economic issues are a component of obtaining or maintaining privileges, which may be the case when the medical staff privileges are an element of an employment relationship with a facility or hospital-based practice group, ascertain this at the outset of the relationship to determine whether you are willing and able to meet the economic aspects of the credentialing. If you do not believe that quality care can be delivered when you are required to see a high number of patients each day, for example, it is helpful to have that discussion with a prospective employer or facility administrator early on in the relationship, rather than after several unsatisfactory performance evaluations.

Group Practices and Employment
When physicians join a partnership or group, they often become employees of that practice. Issues that should be evaluated, negotiated, and documented in a written agreement
include the following: duties, compensation, the term of employment, how the employment relationship may be modified or terminated, insurance coverage and responsibility for defending the physician if malpractice complaints are made, ability to provide services outside the practice (e.g., medical services, teaching, research), compensation for those services, any restrictions on competitive activity during or after the employment relationship, and whether medical staff privileges terminate if the employment relationship ends.

In addition, many employed physicians seek assurances that they will be permitted to acquire an ownership interest in the practice, or at least a share of the practice’s profits, at some later time. The details of the negotiations vary greatly, depending on the particular practice group. For example, a small partnership taking on a new physiatrist early in her career may be unwilling to guarantee that she will become a partner, at least until the relationship has a chance to develop. On the other hand, a large and well-established practice group may have a well-structured approach for all new employed physicians, which establishes an equity (ownership) interest over time and which clearly delineates a schedule for how that interest is forfeited if employment ends. While employment agreements and practice documentation may look similar to many physicians who receive the offer to join a group, the rights of the parties as set forth in the final negotiated paperwork vary greatly, depending on many factors.

Payment and Compensation Issues
Although legal issues relating to physician compensation are complex and cannot be adequately covered in this chapter, some key concepts should be kept in mind:

1. **Referrals and compensation.** Compensation for referrals is prohibited under several federal laws, including the Medicare statute and regulations. For physiatrists and their facilities and practice groups, entities that receive patient referrals as a matter of course, this means that relationships with referral sources must be carefully analyzed to ensure that a legitimate basis exists for any benefits (not just monetary compensation) that flow to the referral source. Similarly, as physiatrists make referrals to therapists and others, the restrictions prohibiting compensation for the referrals apply to these situations. In cases where a legitimate basis for payments cannot be established, regulators may attempt to challenge the payments as impermissible compensation related to referrals.

2. **Billing issues.** Although it may be intuitive to say that bills should only be generated by physiatrists for services actually rendered, the complexity of medical billing (exacerbated in situations where more than one payer is involved) presents a constant challenge for clinicians. As payers seek to reduce expenditures for medical services and to eliminate fraud from the system, audits and challenges to payment increase in frequency. Physiatrists and those who prepare their billing must be knowledgeable as to what services may be billed, what restrictions are imposed by payers (such as prior authorizations and recertifications for prolonged courses of treatment), the processes for appealing when payment is denied, and the ability to recover from the patient, an employer, or others, if payment from one source is not forthcoming. In these situations, a clear understanding of the clinician’s right to payment is essential, not only for the clinician, but also for the patient or the employer seeking the medical services, who may be an effective advocate for payment.

Payer Contracts
Physicians often must be recognized by payers as “participating providers” before payment is made for clinical services. As payers seek to reduce expenditures for clinical activity, they have sought to keep the panel of participating providers small (making it easier to negotiate lower payment rates, justified by large volumes of patients seen by fewer clinicians). As health care costs continue to rise, there will be more pressure to reduce costs from payers, employers, patients, and others. Provider fees may decrease, capitated arrangements and other risk-shifting methods may change the types and quantities of services that are covered by health plans, and it is likely that receiving payment will become more difficult for physicians.

Physiatrists will need to

- Secure their participation in health plan provider panels
- Understand health plan limitations on coverage and payment
- Communicate those limitations to patients
- Limit their liability for claims related to administrative (nonclinical) decisions by health plan personnel
- Endure the administrative burden of dealing with payers, appealing payment denials, and so on
- Evaluate whether participation in various plans is economically justified

Compliance
It is now customary in the health care industry for providers to adopt compliance plans to ensure that their operations comply with legal and ethical requirements. Physiatrists should thoroughly understand their obligations under compliance plans that relate to their activities. For example, if a physiatrist suspects that improper payments are being made to a referral source, that physiatrist must understand whether he or she is obligated to report that suspicion, and how that report will be handled. Education efforts are generally in place in large health care facilities to inform employees, medical staff members, and others of their compliance obligations.

Antitrust
Although the topic of antitrust is beyond the scope of this chapter, medical professionals should have some awareness of its key concepts and prohibitions, especially as the health care
market is changing, aligning, and consolidating parties in ways that may violate antitrust laws.

Antitrust laws are intended to promote fair competition because consumers benefit from it. Antitrust laws prohibit monopolies, unfair competition, and unreasonable restraints of trade. In the current health care industry, practice and facility mergers that may adversely affect access to health care or the cost or price of health care are being challenged by antitrust regulators. When actual or potential competitors (such as two rehabilitation practice groups in a small geographic area) discuss fees charged for services, they may be open to allegations that they have “fixed” prices (tried to establish a floor price, which neither will undercut). This type of activity may result in severe consequences under the antitrust laws.

The changing nature of the health care industry does not lend itself to neat compliance with the antitrust laws. Although regulators are issuing current guidance to the industry, all providers of health care goods and services should exercise caution and obtain advice before engaging in discussions or transactions that may expose them to claims of antitrust violations.

SUGGESTED READINGS

Squillante NJ. Expanding the potential tort liability of physicians: a legal portrait of “nontraditional patients” and proposals for change. UCLA Law Rev. 1993;40:1617.
Not chaos-like, together crushed and bruised. But, as the world harmoniously confused:
Where order in variety we see, And where, though all things differ, all agree.
Alexander Pope (1688–1744), Windsor Forest, 13

Rehabilitation Medicine (RM) in its more modern modalities is nearly a century old. (Older modalities such as spa treatments have been practiced for at least 300 years in Europe and even longer by the Aztecs in Mexico.) The field deals with various pathological disorders that may result in debilitating chronic conditions in patients of both sexes throughout their life span. It utilizes different treatment modalities, classical and innovative, of both low and high technological content. It entails interacting not only with the patient but also with the patient’s family, work environment, and social milieu. As a result, rehabilitation professionals find themselves dealing with insurers as well as with legislative and political entities. RM is practiced in different parts of the world in different ways, depending upon the way both the individual practitioner and the entire rehabilitation system work.

In his talk at the first world congress of the society in 1952 in London (UK), Frank Krusen, the first president of the International Federation of Physical Medicine Society, cited a consensus of the participants by saying, “rehabilitation has become established as a new medical discipline which aims at restoration of the physically handicapped person...to normal life; that physicians should avoid an attitude of hopelessness or passive acceptance in the face of chronic illness or disability; that a dynamic approach to chronic illness frequently results in restoration of the chronically ill patient to a fair measure of self-sufficiency, self-respect and happiness; that physicians should be interested not only in adding years to life but also in adding life to years...the physician should always consider the psychological as well physical problems.”

This chapter describes the practice of RM, as well as its related programs and activities in different geographical regions in the world, with the aim of explaining how the profession is perceived and put into practice. It also discusses the attempt to bring globalization and standardization of the profession through coordinated efforts.

NORTH AMERICA: UNITED STATES AND CANADA

Globally, the field of physical medicine and rehabilitation (PMR) is committed to providing quality, compassionate, holistic, and team-oriented care to people with disabilities. Rehabilitation practitioners worldwide work to improve quality of life and enhance functional status through a combination of team work, dedication, and focused assessment of a patient’s functional strengths and deficits within a medical and psychosocial context. American physiatrists have a long history of involvement in international PMR activity. The process of rehabilitation has been compared to “performing surgery from the skin out” in that significant realignment, readjustment, and modification of a patient’s environment and circumstances must be performed in order to compensate for impairments and functional deficits.

Within the context of international aspects of RM, this section focuses exclusively on the North American experience. It provides a concise overview of salient topics including a brief history of PMR focused on the development and evolution of the specialty in the United States, an analysis of the major practice areas and specializations among US physiatrists, demographics, consideration of cooperative and harmonizing projects, review of US residency training and educational matters, discussion of problems and recent challenges, elaboration on international outreach efforts such as exchange programming, discussion of the Canadian system of rehabilitation, and a forecast for the future of rehabilitation in North America.

Due to the preeminent role that the United States and Canada have played as a beacon of education and enlightenment for international medical education, this chapter would not be complete without a description of the importance of international educational exchange programs in PMR. Under the auspices of the International Society of Physical Medicine and Rehabilitation (ISPMR), the Washington State Society of Physical Medicine and Rehabilitation (WSSPMR), the Canadian Physical Therapy Association (CPTA), and the American Physical Therapy Association (APTA), PMR programs have been developed.

The medical specialty of Physical Medicine and Rehabilitation (PMR) is referred to as such in the USA and Canada. Frequently, “PMR” and PRM are used interchangeably. In other parts of the world however (ie. Europe, Asia and many other international locations), the field is frequently referred to as Physical and Rehabilitation Medicine (PRM). For the purpose of uniformity and convention throughout this chapter, the specialty will be listed as PMR for when referring to the USA and Canada and PRM in other locations.
and Rehabilitation (ISPRM), the Faculty-Student Exchange Committee (www.isprm-edu.org) is composed of international members. It is presently directed by a United States–based elected physiatric leadership and has been responsible for the building of global educational and humanitarian bridges in physiatry. This topic is elaborated upon later in this chapter (page 542).

The field of PMR in North America traces its origins to the mid-20th century when a seismic shift in thinking among health care providers occurred. Due to the burgeoning numbers of wounded and injured soldiers emanating from the war, a new emphasis was placed on the value of comprehensive, team-oriented care for people with disabilities of all types. The critical importance of ministering to and caring for people with disabilities all across America soon came to be recognized as an essential societal and ethical obligation.

During the early days of PMR, the specialization was actually divided into two separate and distinct areas: physical medicine and rehabilitation (6). Frank Krusen (1898–1973), the author of the first rehabilitation textbook and the founder of the first residency training program in physical medicine (Mayo Clinic), is largely credited for his pioneering role in physical medicine which focused on the use of electricity, heat, light, mechanotherapy, exercise, and other modalities in the alleviation of disease and disability. To differentiate physical medicine physicians from physical therapists, Krusen originated the name “physiatrist”—an appellation that has stuck to this day.

Almost simultaneously, the legendary Dr. Howard Rusk (1901–1989), responding to the need for help with the medical and functional restoration of wounded, disabled soldiers returning from active duty, started the field of rehabilitation within the United States. Rusk's early pioneering efforts resulted in the establishment of one of the first comprehensive inpatient rehabilitation hospitals—the Rusk Institute of Rehabilitation Medicine at New York University.

Within the United States, physical medicine was first (primordially) recognized as a medical specialty in 1947, with the establishment of the American Board of Physical Medicine. In 1949, the word “rehabilitation” was added to the name. Over the years, PMR as practiced in the United States has evolved steadily. During the 1980s and 1990s, inpatient rehabilitation remained the mainstay of physiatric practice. More recently, there has been an increased shift in emphasis toward outpatient services.

PMR has prospered as a medical specialization, not only in mainland United States, but also in its territories. In Puerto Rico, for example, physiatry has grown by leaps and bounds.

Dr. Herman Flax, one of the early pioneers of PMR in Puerto Rico, chronicled the history and development of the field over a 25-year period (7).

Within the United States and Canada, PMR is a medical specialty approved by the Accreditation Council for Graduate Medical Education that emphasizes prevention, diagnosis, and treatment of persons with disabilities who experience restrictions in function resulting from disease, injury, or symptom exacerbation. Practitioners of PMR are known as “physiatrists” and utilize a holistic and team-oriented treatment approach which often combines medication, injections, exercise, physical modalities, and education customized to the patient’s unique requirements (6). PMR specialists offer care for persons with neuromuscular disorders who have acute and chronic disabilities. The overriding goal of PMR is to optimize patient function in all domains of life, including the medical, emotional, social, and vocational spheres.

With the “graying” and aging of the North American population, physiatry has grown significantly in its stature as a medical specialization and has been considered the “quality of life” medical specialization because of its unique focus on functional restoration and contribution to the conservative nonsurgical management of an aging population.

According to US Census 1990–2000 data, approximately 49 million noninstitutionalized Americans have a disability (one in five Americans). While there are many causes of disabilities, the most salient ones are those resulting in deficits of activities of daily living. Figure 22-1 below describes the most common conditions known to cause disabilities along with their respective prevalence.

Physiatrists in the United States practice within diverse and varied work settings. Although their work activities and specializations vary widely, there is still much consistency in work distribution. According to a 2005 study commissioned by the American Academy of Physical Medicine and Rehabilitation...
(AAPM&R), a majority of American physiatrists (according to time allocation) are involved in the general practice of PMR in the outpatient setting (50.1%), inpatient venue (23.1%), administration (9.8%), academic activities (4.4%), and research (2.8%). An AAPM&R comprehensive membership survey study conducted in 2002 revealed that for all respondents: 56.6% described themselves as “private practitioners” and 44.2% as “employed.” Approximately 0.8% of the above were both in private practice and employed.

Regarding physiatry employment settings for “employed physiatrists,” 47.2% worked in a hospital or rehabilitation institution, 41.8% in an academic setting, 9.4% as VA employees, and 6.1% as HMO employees. Many of the “employed physiatrists” worked in multiple settings. Regarding private practice physiatrists, 44.3% were in solo practice, 29.9% were in multiple specialty, 29.7% in single specialty, and 2.9% in other practice settings.

The pathway to becoming a board certified physiatrist in the United States begins in medical school. After completing 4 years of medical school, the medical school graduate enters an internship which is a 1-year sequence of coordinated rotations emphasizing basic medical skills within a transitional environment or an accredited training program in internal medicine, surgery, pediatrics, family medicine, obstetrics and gynecology, or some combination thereof. The 1-year internship period is followed by a 3-year residency training program in PMR.

At the conclusion of the 3-year residency, the qualifying graduate takes a board exam administered by the American Board of Physical Medicine and Rehabilitation (ABPM&R). Once this exam is successfully completed, the graduate takes another exam, the oral board, at the conclusion of the first year of practice. Recertification is required of each graduate every 10 years.

An increasing number of PMR residency graduates are choosing to pursue accredited fellowship training in PMR subspecialties by the ABPM&R including Spinal Cord Injury (SCI) (2002), Pediatrics (2002), Pain Medicine (2002), Neuromuscular Medicine (2006), Hospice and Palliative Care (2006), and Sports Medicine (2006). Other nonaccredited fellowships exist in many areas. The typical duration of these fellowships is between 1 and 3 years.

The ABPM&R has published residency training statistics that focuses on demographics, program size, and composition, as outlined in Table 22-1 (8).

Two key textbooks frequently utilized by physiatrists throughout North America (and the world, for that matter) are Physical Medicine and Rehabilitation: Principles and Practice by Joel DeLisa (9) and Physical Medicine and Rehabilitation Medicine by Randall Braddom (10). Other useful books include Essentials of Physical Medicine and Rehabilitation by Walter Frontera et al. (11). Useful handbooks for medical students and residents include PMR Secrets by Bryan O’Young et al. (12) and PMR Pocketpedia by Howard Choi et al. (13).


While the field as a whole continues to carry on its noble mission of providing quality, compassionate care for people with disabilities of all types, the specialization faces a major challenge from some insurers and third-party payers who have steadily decreased reimbursement. In addition, shrinkage of the allotted lengths of stay for a growing number of diagnostic categories has taken place. An illustrative example of this is that of an acute stroke patient who is allotted on average approximately 14 to 21 days of acute hospital rehabilitation prior to community discharge. Unlike many parts of the world, where such patients would be treated for 2 to 3 months, the US system of care offers much shorter lengths of stay.

Another major challenge has been the advent of prospective payment systems (14) for acute care hospitalization which has led to a shifting of patient care from the acute rehabilitation setting where rehabilitation is delivered in a focused concentrated fashion for approximately 3 hours a day to a subacute environment where the rehabilitation is less intense and often is of a “nursing home style” vintage. Throughout the United States, in order to promote vertical integration of health delivery systems, hospitals have transformed acute care beds into subacute care beds. In addition, many hospitals have purchased subacute care facilities including nursing homes and convalescence centers which serve a “surrogate rehabilitation” role.

Apart from the academic or hospital-based physiatric programs, an increasing number of private practice physiatrists (both solo practice and group practice) have now elected to focus a significant portion of their practice on pain management. This has occurred because of economic factors and arguably may pose a major threat to the roots and traditions of the field, since there may now be an evolving and diminishing number of available physiatrists to care for people with traditional disabilities such as stroke, SCI, and brain injury.

Within the educational and residency training realm, a growing and unprecedented number of graduating residents are pursuing fellowship training in PMR and a disproportionate number are choosing a pain management trajectory. Although some are restricting their practice to pain management, many still recognize the importance of remaining diversified and being prepared to care for persons with all types of disabilities. Since people with traditional disabilities (e.g., stroke, SCI, and Traumatic Brain Injury (TBI)) occupy an essential and ever growing portion of the health care bandwidth in the United States, rehabilitation “generalists” are very much in demand. A physiatrist’s higher calling of transcendent care and impassioned advocacy for persons with disabilities was recently noted in an issue of JAMA.

“As practitioners of the healing art, we are often swept away by the mundane minutiae of providing expert technical care to our patients. Often overlooked (however) is the human side of caring—that transcendent sense of seeing life through the eyes of our patients” (15).

The field of PMR continues to thrive and prosper within the United States because of the synergy that it has engendered.
### Table 22.1 ABPM&R Residency Training Statistics

<table>
<thead>
<tr>
<th>Year</th>
<th>Positions Offered</th>
<th>Positions Filled</th>
<th>First Year Int.</th>
<th>Second Year</th>
<th>Third Year</th>
<th>Fourth Year</th>
<th>Fifth Year</th>
<th>Sixth Year</th>
<th># in Combined Programs</th>
</tr>
</thead>
<tbody>
<tr>
<td>1996–1997</td>
<td>1368</td>
<td>1305 (95%)</td>
<td>96</td>
<td>409</td>
<td>402</td>
<td>379</td>
<td>18</td>
<td>1</td>
<td>88b</td>
</tr>
<tr>
<td>1997–1998</td>
<td>1302</td>
<td>1249 (96%)</td>
<td>74</td>
<td>373</td>
<td>386</td>
<td>392</td>
<td>25</td>
<td></td>
<td>86</td>
</tr>
<tr>
<td>1998–1999</td>
<td>1291</td>
<td>1239 (96%)</td>
<td>82</td>
<td>393</td>
<td>353</td>
<td>392</td>
<td>16</td>
<td></td>
<td>77</td>
</tr>
<tr>
<td>1999–2000</td>
<td>1257</td>
<td>1211 (97%)</td>
<td>86</td>
<td>390</td>
<td>366</td>
<td>351</td>
<td>18</td>
<td></td>
<td>48</td>
</tr>
<tr>
<td>2000–2001</td>
<td>1248</td>
<td>1217 (98%)</td>
<td>90</td>
<td>388</td>
<td>366</td>
<td>360</td>
<td>13</td>
<td></td>
<td>43</td>
</tr>
<tr>
<td>2001–2002</td>
<td>1233</td>
<td>1199 (98%)</td>
<td>81</td>
<td>383</td>
<td>373</td>
<td>350</td>
<td>12</td>
<td></td>
<td>37</td>
</tr>
<tr>
<td>2002–2003</td>
<td>1201</td>
<td>1185 (99%)</td>
<td>70</td>
<td>370</td>
<td>363</td>
<td>373</td>
<td>9</td>
<td></td>
<td>26c</td>
</tr>
<tr>
<td>2003–2004</td>
<td>1217</td>
<td>1194 (98%)</td>
<td>75</td>
<td>384</td>
<td>357</td>
<td>372</td>
<td>5</td>
<td>1</td>
<td>13b</td>
</tr>
<tr>
<td>2004–2005</td>
<td>1236</td>
<td>1211 (98%)</td>
<td>82</td>
<td>382</td>
<td>373</td>
<td>366</td>
<td>8</td>
<td></td>
<td>17</td>
</tr>
<tr>
<td>2005–2006</td>
<td>1227</td>
<td>1217 (99%)</td>
<td>75</td>
<td>388</td>
<td>380</td>
<td>371</td>
<td>3</td>
<td></td>
<td>10</td>
</tr>
<tr>
<td>2006–2007</td>
<td>1256</td>
<td>1244 (99%)</td>
<td>81</td>
<td>393</td>
<td>382</td>
<td>384</td>
<td>4</td>
<td></td>
<td>14</td>
</tr>
<tr>
<td>2007–2008</td>
<td>1257</td>
<td>1247 (99%)</td>
<td>86</td>
<td>386</td>
<td>391</td>
<td>379</td>
<td>5</td>
<td></td>
<td>19</td>
</tr>
</tbody>
</table>

### Demographics of PMR Residents

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>American Med Grads</td>
<td>1055 (85%)</td>
<td>1054 (85%)</td>
<td>996 (82%)</td>
<td>924 (76%)</td>
<td>821 (68%)</td>
<td>776 (65%)</td>
<td>753 (63%)</td>
<td>754 (62%)</td>
<td>774 (64%)</td>
<td>866 (70%)</td>
<td>925 (73%)</td>
<td>1010 (77%)</td>
<td></td>
</tr>
<tr>
<td>IMG—Non-US citizens</td>
<td>72 (6%)</td>
<td>72 (6%)</td>
<td>65 (5%)</td>
<td>103 (9%)</td>
<td>140 (12%)</td>
<td>238 (20%)</td>
<td>236 (20%)</td>
<td>258 (21%)</td>
<td>257 (21%)</td>
<td>240 (20%)</td>
<td>230 (18%)</td>
<td>211 (16%)</td>
<td></td>
</tr>
<tr>
<td>IMG—US citizens</td>
<td>120 (10%)</td>
<td>118 (9%)</td>
<td>156 (13%)</td>
<td>184 (15%)</td>
<td>233 (20%)</td>
<td>211 (15%)</td>
<td>210 (17%)</td>
<td>205 (17%)</td>
<td>180 (15%)</td>
<td>133 (10%)</td>
<td>104 (8%)</td>
<td>84 (7%)</td>
<td></td>
</tr>
<tr>
<td>Male residents</td>
<td>754 (60%)</td>
<td>759 (61%)</td>
<td>760 (62%)</td>
<td>753 (62%)</td>
<td>745 (62%)</td>
<td>730 (62%)</td>
<td>745 (62%)</td>
<td>772 (63%)</td>
<td>768 (63%)</td>
<td>796 (65%)</td>
<td>807 (65%)</td>
<td>861 (66%)</td>
<td></td>
</tr>
<tr>
<td>Female residents</td>
<td>485 (40%)</td>
<td>485 (39%)</td>
<td>457 (38%)</td>
<td>458 (38%)</td>
<td>449 (38%)</td>
<td>455 (38%)</td>
<td>454 (38%)</td>
<td>445 (37%)</td>
<td>443 (37%)</td>
<td>443 (35%)</td>
<td>442 (35%)</td>
<td>444 (34%)</td>
<td></td>
</tr>
</tbody>
</table>

### 79 Accredited Residency Training Programs Responded

- Programs with 0–9 residents: 11
- Programs with 10–19 residents: 51
- Programs with 20–29 residents: 13
- Programs with 30–39 residents: 2
- Program with 40–49 residents: 2

**Total Residents 1,247**

### Faculty

- Physicians at primary facilities: 728
- Physicians at affiliated facilities: 842
- Est. needed additional faculty (3 years): 167

### Diplomates

- New diplomates 2007: 349
- Total certificates issued (1997–2007): 8,818

Chapter 22 | International Aspects of the Practice of Rehabilitation Medicine

Most physiatry residents also choose to complete 6 months of neuromuscular disease/electromyography training during their residency. This allows them to sit for the Canadian Society of Clinical Neurophysiologists electromyography examination which is administered to both physiatry and neurology residents.

Canada, with a population of 33 million (19), currently has 373 physiatrists (20). As Royal College certified specialists, all physiatrists must participate in the RCPSC Maintenance of Certification program and must accumulate 400 hours of continuous medical education (CME) credits over a 5-year cycle.

The Canadian Association of Physical Medicine and Rehabilitation (CAPM&R, capmr.medical.org) is the national organization for physiatrists in Canada. Its main goals are to promote education, scientific collaboration, and practice development among physiatrists. A 2006 survey by the CAPM&R indicated that 53% of physiatrists are in private practice. Of those, one third do not carry inpatients. Seventy-eight percent of physiatrists have clinical or academic university appointments. Similar to the United States, there has been a shift in the practice patterns of physiatrists to an outpatient focus. The inpatient system is moving toward a model where hospitalists/family physicians are the attending physicians of record, and physiatrists act as the consultant.

In contrast to the United States, where health care is more managed care driven, Canada has a universal health care system. In this model, every citizen has equal access to health care and is fully insured by the government for all medically necessary care. This has both advantages as well as drawbacks with respect to the inpatient system. It ensures that anyone who benefits from rehabilitation services will be given the opportunity to participate in a program. Another advantage is that inpatient length of stay varies depending on the needs of the patient, and this is decided by the physiatrist and the rehabilitation team. For example, in the stroke population, the average length of stay in Canada is longer (40 days) (21). By contrast, a drawback of Canada’s social system is that wait times can be prolonged compared to the United States; the time from onset of stroke to inpatient rehabilitation admission is an average of 29 days (median: 15 days).

In conclusion, the North American contribution to international rehabilitation has been immense. As outlined earlier in this chapter, the humble origins of the field of PMR began in the United States with the work of dedicated pioneering physicians including Dr. Howard Rusk and Dr. Frank Krusen. Over time, the specialization has grown considerably.

Latin America

The origin of rehabilitation as a medical specialty in the developed countries and at the global level was linked to historical facts that promoted a great number of disabilities, such as the two World Wars in the first half of the last century. However, within Latin America the main impetus for the development of RM was the outburst of epidemics of transmissible illnesses...
such as poliomyelitis, which is now eradicated in many of the region’s countries. This brought about an improvement in sanitary conditions and a rapid development of rehabilitation as a medical specialty. These phenomena emerged simultaneously in many Latin American countries.

European countries such as England, France, and Germany, as well as the United States, served as models for many of the Latin American rehabilitation systems. Initially, the area of expertise was predominantly focused on aspects of Physical Medicine, and later on, RM started to develop as well. A continuous demand for the improvement in sanitary conditions and the possibilities for the development of the rehabilitation field at the national level hastened the development of these therapeutic modalities in this part of the world.

Scientific Societies of Physical Medicine and Rehabilitation were being established in the individual countries of Latin America and the Caribbean in the 40s and 50s, and consequently, the need of a Latin American Association that would bring these societies under one umbrella organization was recognized.

In 1961, during the first Mexican Congress of Rehabilitation, the distinguished Mexican Doctor, Alfonso Tohen Zamudio, brought together representatives from all Latin American countries. At that congress, they laid the basis of the Latin American Medical Association of Rehabilitation (AMALAR) and decided to hold the first AMLAR Congress in Mexico 2 years later in 1963, the second in 1967 in Peru, and the third in 1969 in Uruguay. Ever since that time, 23 congresses of the AMLAR took place. The AMLAR’s presidency is in Uruguay, since the XXIII Congress, held in the city of Punta del Este (Uruguay), in October of 2008.

Rehabilitation in Uruguay has its origins in the 1940s. Prof. Alvaro Ferrari Forcade saw that there was a need for a functional measurement tool to assess the patient’s status. He developed a scale that registers the patient’s functional clinical state in three basic categories: somatic, psychological, and social. The scale assigns a score to each of the categories, in such a way that the results depict the patient’s “profile of disability.” Moreover, it allows for predicting the patient’s functional outcomes (22,23).

In this way, the first functional measure of disability in Latin America was developed. Due to the fact that the implementation of this scale was done early in the development of RM in the region, its importance was not fully acknowledged at the time. This can explain the reason why this scale did not catch on while other such scales did, shortly afterward, in other parts of the world.

Latin America spans a large area with respect to political, geographic, demographic, and epidemiologic aspects. This is the reason for the varied degrees in the implementation of the RM and access to technology in the Latin American health systems. The demographic projections for Latin America and the Caribbean show that there will be progressive and dynamic changes during the course of next 40 years. These trends show a shift in the demographic profile, with significant aging of the population, due to a decrease in the birth rate and an increase in life expectancy for those over 60 years of age.

Even though in all the countries there is a marked increase in the population 60 years of age and above, the demographic change is not homogeneous and while some countries are still in an initial stage of increase, others are in a more advanced phase (24). Cuba in the Caribbean and the countries of the south cone (Argentina, Chile, and Uruguay) in Latin America, have a more aged population (Fig. 22-2), showing a more advanced transition.

This demographic change has a consequence from the epidemiological point of view. An increase in the elderly population will increase the prevalence of chronic nontransmissible illnesses related with the elderly. There is also an increase in the occurrence of nontransmissible illnesses linked to trauma and habits of life (accidents, social violence, accidents at work, and addiction) that coexist in the region with the transmissible illnesses. The increased number of high-risk newborns, along with malnutrition and the societal exclusion of indigenous ethnic groups (25), all contribute to the increase in the demand for qualified RM practitioners. The growing demand for rehabilitation professionals, and that these professionals be trained to the highest standards and continue their training after certification with a CME program, are the main focuses of the Scientific Societies of Rehabilitation and Physical Medicine of each country and AMLAR (26).

It is estimated that the population with disabilities in Latin America is close to 85 million people. Out of those, approximately 3 millions are in Central America (25). Although we know about the social and economic impact of disabilities in Latin America, it becomes more difficult to plan accurate strategies due to the lack of trustworthy epidemiological data. The reported data vary and it is not certain that they describe the true epidemiological situation in many of the countries, and neither are they statistically comparable. This is a consequence of using different disability concepts and different methodologies for gathering the information. Some countries have included questions regarding disability in the population censuses. Others have carried out specific surveys for disability considering the articles corresponding to the ICF (International Classification of Functioning, Disability, and Health). Generally speaking, the censuses have shown a lower prevalence of disabilities than the surveys. The discrepancy in the results is also due to what is considered a disability. Is disability based on the impairment a mainly medical concept? Or is it based on the functioning, integration, and the individual’s participation with consideration to the patient’s interaction with the environment? (This definition is based on the ICF paper “Measuring Disability Prevalence”. Social Protection Discussion Paper, No. 0706. The World Bank, March 2007). This heterogeneity in concepts and methods makes the comparison of the results difficult when referring to the prevalence of disabilities in Latin America (Table 22-2).

RM has developed in Latin American countries but not uniformly. The availability of rehabilitation facilities is found in 78% of the countries. There is specific legislation regarding rehabilitative medicine in 62% of the countries and implementation of specific programs in 51% of the countries. The field
is further hindered by very limited documentation of relevant data and limited research work. The practitioners are either specialists in rehabilitation or technicians consisting mainly of physical therapists, in addition to speech-pathologists, psychologists, nurses, social workers, and occupational therapists. The medical schools and colleges offer little training and teaching in the field of rehabilitation. In addition to the factors mentioned above, there are additional roadblocks to

### TABLE 22.2 Prevalence of Disability in Countries of Latin America, According to Statistical Method and Disability Concept

<table>
<thead>
<tr>
<th>Country</th>
<th>Method</th>
<th>Year</th>
<th>Disability Concept</th>
<th>Prevalence (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Argentina</td>
<td>Survey</td>
<td>2003</td>
<td>Functioning</td>
<td>7.1</td>
</tr>
<tr>
<td>Brasil</td>
<td>Census</td>
<td>2000</td>
<td>Functioning</td>
<td>14.5</td>
</tr>
<tr>
<td>Colombia</td>
<td>Census</td>
<td>2005</td>
<td>Functioning</td>
<td>2.0</td>
</tr>
<tr>
<td>Costa Rica</td>
<td>Survey</td>
<td>2000</td>
<td>Impairment</td>
<td>5.4</td>
</tr>
<tr>
<td>Chile</td>
<td>Survey</td>
<td>2004</td>
<td>Functioning</td>
<td>13</td>
</tr>
<tr>
<td>El Salvador</td>
<td>Survey</td>
<td>2003</td>
<td>Functioning</td>
<td>1.5</td>
</tr>
<tr>
<td>México</td>
<td>Census</td>
<td>2000</td>
<td>Impairment</td>
<td>2.4</td>
</tr>
<tr>
<td>Nicaragua</td>
<td>Survey</td>
<td>2003</td>
<td>Impairment</td>
<td>10.3</td>
</tr>
<tr>
<td>Panamá</td>
<td>Census</td>
<td>2000</td>
<td>Impairment</td>
<td>1.8</td>
</tr>
<tr>
<td>Paraguay</td>
<td>Census</td>
<td>2002</td>
<td>Impairment</td>
<td>3.4</td>
</tr>
<tr>
<td>Uruguay</td>
<td>Survey</td>
<td>2003</td>
<td>Functioning</td>
<td>7.6</td>
</tr>
</tbody>
</table>

Source: National Statistical Institutes.
the progress of RM in Latin America. Some of those problems are there are too few practitioners of RM in most countries; people with disabilities have a hard time integrating themselves into society; the private sector plays a preeminent role in the health care systems and is reluctant to pay for rehabilitative medicine (25); large sections of the population do not have accessibility to medical aid, let alone to rehabilitation facilities. There is a distinct association among disability, poverty, and social exclusion.

On the whole, the number of specialists in RM is in fact greater than the one registered by each national scientific society. Table 22-3 refers to the approximate numbers of physicians in some countries of the region, which have exceptionally low rates of specialists per inhabitants.

RM as a medical specialty was strengthened by unifying approaches to rehabilitative medicine which took place at the study group on The Training of Specialists, summoned by the PAHO/WHO in Santiago, Chile, in October 1969 (23).

There are three basic models of rehabilitative care in Latin America: (a) hospitals of acute care, (b) rehabilitation centers, or (c) community-based rehabilitation (CBR). Most countries have an emphasis on one of these models and subsequently affect the specialist’s profile in each country. Uruguay is an example of a country that uses the rehabilitative model based on hospitals of acute care while promoting the interdisciplinary approach of the rehabilitation. They have recently started to work on rehabilitation centers related to pediatric care. Argentina and Chile have developed the models based on hospitals of acute care and rehabilitation centers. Nicaragua, El Salvador, Colombia, Argentina, and Guyana have developed the model based on CBR. Together with the help of non-governmental organizations, the respective governments are developing this model in Mexico, Ecuador, Peru, and Bolivia. Brazil is a remarkable example where all three models are used. The multidisciplinary diagnostic and therapeutic approach of disability treatment in Brazil has been an important factor in the development of rehabilitation in this country. The Rehabilitation Institute of the Hospital das Clínicas of São Paulo, the ABBR in Rio de Janeiro, and the Rehabilitation Institute of San Salvador were pivotal academic training sites in the specialty of rehabilitative medicine for Brazil and also for the region. In the early 70s, Brazil was one of the first countries that developed a residency program in rehabilitation training (27). This residency program was used as a model for other such programs in others countries of the region.

The technology used in rehabilitation has improved from the 90s in several countries such as Argentina, Chile, Brazil, Colombia, and Mexico, especially in the field of orthopedic and neurological rehabilitation. Improvements in these countries had a positive effect throughout the region, in the practice of the specialty, in the academic training programs, and in the makeup of the specialist’s profile.

Recently, AMLAR has begun to work toward unifying the specialist doctor’s profile in the different countries of Latin America. They are trying to achieve this through the Societies of Rehabilitation and Physical Medicine of the individual countries, as well as in the specialist doctor’s CME, the professional certification, and certificate reexamination. In the XXI Congress of the AMLAR (Venezuela, 2004), it was agreed that specialists in rehabilitation should be qualified to perform the following:

1. To identify and evaluate, in an integrative way, disabling conditions and processes
2. To devise a therapeutic program with the goals of preventing the subsequent deterioration in the somatic, psychic, and social realms of the patient
3. To implement and coordinate actions that will compensate for lost function by means of reeducation or compensation by functions that the patient still has, having as an objective to return the individual to the environment, with the highest possible level of independence
4. To acquire the skill and ability of prescribing, managing, and supervising the different procedures of the physical medicine and the therapeutic reeducation, making use of medical resources, technicians, family, and community
5. To work with an interdisciplinary team of experts in many diverse fields: nursing, physical therapy, occupational therapy, speech-pathology, neuro and clinical psychology, social workers, and others
6. To undergo special training in the use of exercise and physical modalities, orthopedic prescription, orthosis and technical helps, in the appropriate use of technology in rehabilitation. Training will be given in the use and application of instruments of functional measure and quality of life. In addition, special training in electrodiagnosis will be given
7. To conduct research in the areas of basic science, outcome measures and applied technology
8. To correctly handle the basic concepts of management and administration of rehabilitation services in all the levels of practice
9. To develop CME programs and to make sure that the practitioners will successfully complete specialty recertification, according to national and international regulations and standards
10. To practice in an ethically appropriate manner

### TABLE 22.3 Rate of Physiatrists Per 100,000 Inhabitants in Several Countries of Latin America

<table>
<thead>
<tr>
<th>Country</th>
<th>Physiatrists</th>
<th>Population</th>
<th>Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Argentina</td>
<td>240</td>
<td>3,260,130</td>
<td>0.6/100,000</td>
</tr>
<tr>
<td>Brazil</td>
<td>820</td>
<td>169,799,170</td>
<td>0.5/100,000</td>
</tr>
<tr>
<td>Chile</td>
<td>96</td>
<td>15,116,435</td>
<td>0.6/100,000</td>
</tr>
<tr>
<td>Colombia</td>
<td>200</td>
<td>42,888,592</td>
<td>0.5/100,000</td>
</tr>
<tr>
<td>Ecuador</td>
<td>120</td>
<td>18,755,650</td>
<td>0.6/100,000</td>
</tr>
<tr>
<td>Uruguay</td>
<td>68</td>
<td>3,314,466</td>
<td>2.0/100,000</td>
</tr>
<tr>
<td>Venezuela</td>
<td>400</td>
<td>27,483,200</td>
<td>1.5/100,000</td>
</tr>
</tbody>
</table>
In conclusion, the absence of trustworthy epidemiological data related to disability prevalence has hindered the development of specific programs of rehabilitation in most Latin American countries and it has impeded the access to such services. RM has developed differently in the various Latin American countries in accordance with various policies, legislative aspects, and technological developments. In the last decade, Latin America has improved its epidemiological data acquisition to make sure that the data is both more detailed and more uniformed for comparison purposes. A better understanding of the concept of disability in all of its aspects starting from the problems in functioning, of integration and participation in society in general, and by using better epidemiological sources seems to be the way toward standardizing and harmonizing the approaches used. This in turn will hopefully lead to the development of better policies and programs of rehabilitation. This will, in turn, improve the integration of the medical rehabilitation specialty in each of the country's health care systems.

**EUROPE**

The first significant event in RM that took place in Europe was the International Physiotherapy Conference in Rome that took place from 13th to 16th October, 1907. After the conference, several text books on Physical Therapy were published in Italy and many other countries between 1925 and 1935. The development of therapy with physical methods based on medical-scientific origin is the continuation to the previous history of hydrotherapy and thermal treatments. These former treatments had already developed into a very important sector oriented toward the use of movement as an element of health and cure. These treatments were used by the Romans all the way through the medical modalities of Napoleon's army and the Habsburg Empire. In 1931, in Italy and in German speaking countries, authorities were appointed to deal with insurance for accidents in the workplace. This led to the development of functional recuperation and prosthesis. The terms used were physiotherapy and later physical medicine. In the 1950s, the terms recovery and rehabilitation were also used. The areas of clinical application have since become more and more numerous, but in particular the activity is aimed at the rehabilitation of disabled children, rehabilitation following traumatic lesions and injuries, and rehabilitation following peripheral and central neurological traumas and degenerative diseases.

Schools were opened for doctors and other rehabilitation-related professionals such as massotherapists and physiotherapists. The 1st Conference of Physical Medicine was held in Marseille in 1957, then in Milan in 1958. In 1962, the European Union of Medical Specialists (UEMS) recognized physical medicine as an autonomous discipline under the name of physical and rehabilitation medicine or PRM. A characterizing element of that period, which was very rich in clinical experience and innovative intuition, was the development of many distinct methodologies such as Brunnstrom, Vojta, Bobath, Kabat, Salvini-Perfetti, Maigne, and others.

In these years, the first national societies of physical medicine in Italy, France, Portugal, Belgium, and Holland were formed. In addition, a European Academy was formed on the 25th April 1963 in Brussels, and with an officializing act by the Royal Belgian Bulletin, the European Federation of Physical Medicine and Rehabilitation was formed, which initially grouped together the nine national societies which were already present on the continent. Large areas of research developed. One of the first areas of interest involved the treatment of scoliosis, the care and treatment of subjects with war amputations and work injuries, and the thousands of cases of poliomyelitis that were prevalent in the 1950s.

After that, the field developed rapidly and spread over all of Western Europe. The development of rehabilitative medicine was helped by important governmental legislation as well as by the Council of Europe and then by the European Union (28). The European Federation, with the UEMS and the Academy, published the first White Book in 1989 (29). A new edition came out in 2006 (30,31).

Some of the reasons for the great strides in both the scientific and operational advances in RM have been:

- The contribution and integration of ‘adjunct’ health workers, in particular, nurses, orthopedic technicians, physiotherapists, speech therapists, and occupational therapists to RM.
- The thrust and richness of the commitment and the proposals by disabled persons associations and their families who have been able to transform pain, damage, and possible desperation into a clear request for services and treatment. Thus, requiring the national and European community to carry out their cultural as well as economic duties.
- The contribution of many institutes, religious and private foundations who have dedicated themselves to this sector, must not be forgotten.

On a global level, the WHO based in Geneva and other organizations are pushing to gain attention for the problems of people with disabilities and for the documentation of rehabilitation resources in countries world.

The European Parliament has taken a positive interest in policies that safeguard the rights of disabled people. The “Recommendation 1185 (1992)—Rehabilitation Policy for the Disabled” (28) was decisive. The European Union decided upon the value, content, and quality that must be guaranteed to all patients at each medical examination.

The European Definition of a Medical Act is “The medical act encompasses all the professional action, e.g., scientific teaching training and educational, clinical and medico-technical steps to promote health and functioning, prevent diseases, provide diagnostic, therapeutic and rehabilitative care to patients, individual groups or communities and is the responsibility of and must always be performed by a registered medical doctor/physician or under his or her direct supervision and/or prescription.” (3 November 2006—UEMS Meeting in Budapest).
There is wide consensus on the importance of rehabilitative medicine on the ethical, cultural, and political level on the continent. For this reason, there is a progressive improvement at the operational level in both research and treatment and services offered in the 27 countries of the European Community. There are, however, some differences between countries due in particular to different economical conditions. As a result, the PMR and all of its professionals have the responsibility to communicate and understand all pertinent aspects in the Community (32), in addition to their daily clinical tasks of assisting individuals and solving their problems.

In fact, all over Europe great changes have taken place in the perception of health. Disorders, symptoms, and phenomena that perhaps in the past were supported and considered inevitable are now no longer accepted and demands are rightly made for any treatment that can eliminate or alleviate such problems. RM is, more than any other medical specialty, in the center of this transformation. Scientific breakthroughs can potentially modify disabilities that previously were not successfully treated and that have spurred a growth in the demand for rehabilitative medicine services (33).

The rights of the individual to a treatment for any problem related to his or her overall biopsychosocial situation, limitation, or disability that alters even only transiently his or her autonomy, self-sufficiency, and self-determination is the starting point of all rehabilitation activities and associated professional and organizational responsibilities (34). Social Participation is a term that aptly represents the individual’s fulfillment of these rights (35). The individual’s rights are inextricably bound to the duty of society to guarantee every person all possible treatments for maintaining, for as long as possible and at the highest level possible, personal autonomy. It is also society’s duty to optimize and at the same time verify the appropriate use of the many available rehabilitative instruments with respect to parameters of efficacy, efficiency, and sustainability.

Physiatry in Europe is not a very old discipline, and precisely for this reason, it has vigorous roots. Its taproot, which as always in medicine is part science, part research, experience in the field, and critical reflection, highlights the duality of the physical and the rehabilitative components of physiatry. The physical component deals with the whole physical environmental context which interacts positively or negatively with the person and his or her problems, while the rehabilitative component aims to restore balance to the physical components by means of intervention. These two components act synergistically with respect to research and in clinical practice. In the current transformation of health care demands, rehabilitative medicine is showing its maximum potential and enormous intrinsic value.

The traditional medical intervention approach focuses on a disease process. By contrast, physiatry in both practice and research never separates disease processes from the individual as a whole and his or her active and passive relations with his or her surrounding context. Indeed, the physiatric approach is constantly based on the awareness that the determining factor is precisely this relation and not the isolated biological conditions.

Diagnosis, evaluation, and prognosis in traditional medicine are based on analyzing separate symptoms and pathological factors and integrating them all into an overall picture. By contrast, a primarily rehabilitative view considers the whole person in all aspects, functional, emotional, motivational, and behavioral as the primum movens using these parameters for determining the modality, limits, and aims of the care of the patient. Since a patient’s health status is the result of a large and complex number of different factors, the therapeutic pathway to reach the best possible level of recovery and maintenance must be equally complex, synergic, and multifactorial and that is the goal of RM.

Since the 1970s, several forces converged which have stimulated a renewed focus for the work of rehabilitation professionals. These include several European initiatives. Among them are (a) the disability classification systems of the WHO, including the International Classification of impairment, Disability and handicap and the ICF (36); (b) a recent international document on CBR and (c) the Council of Europe’s Directives (37).

The ICF has formalized for all medical practitioners a language and a way of thinking that were always ours in rehabilitation. The research and the applications that are currently coming to fruition provide new elements for growth in all fields, as is happening with the Disability and Rehabilitation Action Plan WHO 2005/2010 “A world for all” (38). PRM claim to be qualified for the overall study of the function and health of a person (and of a community). They are also qualified to do research into those elements that characterize and modify the person and his or her relational context, and for managing the multidisciplinary and multiprofessional collaboration necessary both in order to carry out this research and for performing the therapeutic interventions (39).

The task of PRM is to help the functionality of a patient reach an optimal level, compatible with the potential and aspirations of the patient, using some specific therapeutic strategies:

- treating impaired body structures and functions,
- overcoming obstacles arising from changes in body functions, limitations of activities, or restriction to participation,
- preventing the onset of further disability and limitations in functioning.

The central point for PRM is the Individual’s Rehabilitation Plan which collects and distills all the previously mentioned concepts and concretely expresses the true epicentric status that the individual must hold in all stages and in all interventions. The Individual’s Rehabilitation Plan represents the possible pathway for reaching an optimal state of well-being, participation, and health. It combines the skills and duties of the various members of the Team (which is the indispensable unit for rehabilitation care) and the different treatment settings and structures for the different periods of this process.
The Individual Rehabilitation Plan defines, recapitulates, and verifies the entire content of this tangible pathway toward recovery of as much autonomy as possible; it must, above all, be experienced and agreed by the individual, but it must also be based on precise and validated scientific elements.

Rehabilitation, both on an individual and on a population level, must represent the meeting point between scientific evidence and epidemiological evidence. Using indicators of health in different countries (e.g., the European community as a whole), we can derive at a plan that could list the priorities of possible interventions in relation to their efficacy. This is undoubtedly a fundamental aim for us all (40).

From the point of view of health care, there is a need for the diagnostic, prognostic, and therapeutic evaluation to have a unifying capacity, collecting together all the different elements in order to best understand the way in which the different pathologies, biological, psychic, emotional, relational, and affective problems, can interfere with the person’s functions, potential, and desire to regain activity and health.

In brief, European Health System must move from emphasizing the epidemiology of the disease to stressing the epidemiology of limitations in participation, which is much more significant in terms of real impact.

The WHO’s ICF, is nowadays the best instrument, on both the clinical and community levels, for integrating individual and population data. It offers a uniform methodology of interpretation and representation for all the determinants of the health status of a person and of a population. The European Union provides the means for representing and collating all the data on the health of its member countries, in scientific, administrative, social policy, cultural, educational, and occupational terms (35).

### Education and Training

PRM is an independent medical specialty in all European countries, except Denmark and Malta, but its name and focus varies somewhat according to different national traditions and laws. Training usually lasts for between 4 and 6 years depending on the country (UEMS Charter on Training, EC Directive 93/16/EEC, 5 April 1993) (Table 22-4). The number of specialist per inhabitants is not uniform throughout Europe. The numbers range from a low of less than 0.5 per 100,000 inhabitants in the United Kingdom and Ireland to more than 4.5 per 100,000 inhabitants in Portugal and in the Czech Republic (Table 22-5). Specialists in PRM have freedom of mobility across UEMS member states, but require certification from the national training authorities. Those with national certification are eligible to be certified by the European Board of PRM. The European Board of PRM has developed a comprehensive system of postgraduate education for PRM specialists that consists of:

- A curriculum for postgraduate education containing basic knowledge and the application of PRM in specific health conditions

### Table 22.4: Education on PRM in European Countries 2005 ESPRM

<table>
<thead>
<tr>
<th>Country</th>
<th>Number of PRM Specialists</th>
<th>Total Number of Trainees</th>
<th>Number of Specialists in my Country is</th>
<th>Number of Specialists in my Country is</th>
</tr>
</thead>
<tbody>
<tr>
<td>Austria</td>
<td>202</td>
<td>41</td>
<td>TOO LOW</td>
<td>INCREASING</td>
</tr>
<tr>
<td>Belgium</td>
<td>440</td>
<td>72</td>
<td>ENOUGH</td>
<td>DECREASING</td>
</tr>
<tr>
<td>Czech Rep</td>
<td>483</td>
<td>130</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>Denmark</td>
<td>120</td>
<td>0</td>
<td>ENOUGH</td>
<td>DECREASING</td>
</tr>
<tr>
<td>Finland</td>
<td>152</td>
<td>20</td>
<td>ENOUGH</td>
<td>INCREASING</td>
</tr>
<tr>
<td>France</td>
<td>1,880</td>
<td>60</td>
<td>TOO LOW</td>
<td>INCREASING</td>
</tr>
<tr>
<td>Germany</td>
<td></td>
<td></td>
<td>TOO LOW</td>
<td>INCREASING</td>
</tr>
<tr>
<td>Greece</td>
<td>120</td>
<td>30</td>
<td>TOO LOW</td>
<td>INCREASING</td>
</tr>
<tr>
<td>Hungary</td>
<td>230</td>
<td>36</td>
<td>TOO LOW</td>
<td>INCREASING</td>
</tr>
<tr>
<td>Ireland</td>
<td>13</td>
<td>2</td>
<td>TOO LOW</td>
<td>INCREASING</td>
</tr>
<tr>
<td>Italy</td>
<td>2,200</td>
<td>640</td>
<td>TOO LOW</td>
<td>INCREASING</td>
</tr>
<tr>
<td>Norway</td>
<td>127</td>
<td>37</td>
<td>TOO LOW</td>
<td>INCREASING</td>
</tr>
<tr>
<td>The Netherlands</td>
<td>287</td>
<td>97</td>
<td>TOO LOW</td>
<td>INCREASING</td>
</tr>
<tr>
<td>Poland</td>
<td>1,000</td>
<td>115</td>
<td>TOO LOW</td>
<td>STEADY</td>
</tr>
<tr>
<td>Portugal</td>
<td>480</td>
<td>37</td>
<td>ENOUGH</td>
<td>STEADY</td>
</tr>
<tr>
<td>Slovenia</td>
<td>81</td>
<td>17</td>
<td>ENOUGH</td>
<td>INCREASING</td>
</tr>
<tr>
<td>Spain</td>
<td>1,500</td>
<td>240</td>
<td>ENOUGH</td>
<td>INCREASING</td>
</tr>
<tr>
<td>Sweden</td>
<td>200</td>
<td>55</td>
<td>TOO LOW</td>
<td>STEADY</td>
</tr>
<tr>
<td>Switzerland</td>
<td>220</td>
<td>50</td>
<td>ENOUGH</td>
<td>INCREASING</td>
</tr>
<tr>
<td>Turkey</td>
<td>962</td>
<td>346</td>
<td>ENOUGH</td>
<td>INCREASING</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>129</td>
<td>57</td>
<td>TOO LOW</td>
<td>INCREASING</td>
</tr>
<tr>
<td>TOTAL</td>
<td>10,826</td>
<td>2,082</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
A standardized training course of at least 4 years in a PRM department and detailed documentation in a uniform official logbook

A single written annual examination throughout Europe

A system of national managers for training and accreditation to foster good contacts with trainees in their country

Standard rules for the accreditation of trainers and a standard process of certification

Quality control of training sites performed by site visits of accredited specialists

Continuing professional development (CPD) and continuing medical education (CME) within the UEMS. Recertification occurs every 10 years.

**Interventions in PRM as mainly Accepted in European Countries**

A PRM specialist can either perform or prescribe the given intervention. Interventions include:

- **Medical interventions**
- Medications aimed at restoring or improving body structures and/or function, for example, pain therapy, inflammation therapy, regulation of muscle tone, improvement of cognition, improvement of physical performance, or treatment of depression
- Practical procedures—including injections and other techniques of drug administration

- **Assessment and review of interventions**
- **Prognostication**
- **Physical treatments**
- Manual therapy techniques for reversible stiff joints and related soft tissue dysfunctions
- Kinesiotherapy and exercise therapy
- Electrotherapy
- Other types of physical treatments including ultrasound, heat and cold applications, phototherapy (e.g., Laser therapy), hydrotherapy and balneotherapy, diathermy, massage therapy, and lymph therapy (manual lymphatic drainage)
- **Occupational therapy** to:
- analyze activities, such as those of daily living and occupation, support impaired body structures (e.g., splints)
- teach the patient skills to overcome barriers to activity of daily living (e.g., adjusting the private facilities)
- train in the presence of impaired function and cognition
- enhance motivation
- Speech and language therapy within the framework of complex specialized rehabilitation programs
- **Dysphagia management**
- **Neuropsychological interventions**
- **Psychological assessment and interventions, including counseling**
- **Nutritional therapy**

---

### Table 22.5 Number of PRM Doctors in Europe 2005 ESPRM

<table>
<thead>
<tr>
<th>Country</th>
<th>Inhabitants</th>
<th>Physicians</th>
<th>Number of PRM</th>
<th>Final Year in Education of PRM</th>
<th>Total</th>
<th>Number of PMR Specialists per 100,000 Inhabitants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Austria</td>
<td>8,110,244</td>
<td>36,264</td>
<td>202</td>
<td>202</td>
<td>202</td>
<td>2.49</td>
</tr>
<tr>
<td>Belgium</td>
<td>10,239,085</td>
<td>42,036</td>
<td>440</td>
<td>456</td>
<td>456</td>
<td>4.30</td>
</tr>
<tr>
<td>Cyprus</td>
<td>700,000</td>
<td>1,800</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>1.29</td>
</tr>
<tr>
<td>Czech Republic</td>
<td>10,300,000</td>
<td>35,000</td>
<td>483</td>
<td>518</td>
<td>518</td>
<td>4.69</td>
</tr>
<tr>
<td>Denmark</td>
<td>5,500,000</td>
<td>22,000</td>
<td>120</td>
<td>120</td>
<td>120</td>
<td>2.18</td>
</tr>
<tr>
<td>Finland</td>
<td>5,200,000</td>
<td>19,336</td>
<td>152</td>
<td>152</td>
<td>152</td>
<td>2.82</td>
</tr>
<tr>
<td>France</td>
<td>50,000,000</td>
<td>196,000</td>
<td>1,880</td>
<td>1,880</td>
<td>1,880</td>
<td>3.13</td>
</tr>
<tr>
<td>Germany</td>
<td>80,000,000</td>
<td>36,264</td>
<td>202</td>
<td>202</td>
<td>202</td>
<td>2.50</td>
</tr>
<tr>
<td>Greece</td>
<td>11,000,000</td>
<td>50,000</td>
<td>120</td>
<td>120</td>
<td>120</td>
<td>1.09</td>
</tr>
<tr>
<td>Hungary</td>
<td>10,000,000</td>
<td>35,000</td>
<td>230</td>
<td>230</td>
<td>230</td>
<td>2.30</td>
</tr>
<tr>
<td>Ireland</td>
<td>3,500,000</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>0.09</td>
</tr>
<tr>
<td>Italy</td>
<td>56,000,000</td>
<td>200,000</td>
<td>2,200</td>
<td>2,200</td>
<td>2,200</td>
<td>3.93</td>
</tr>
<tr>
<td>Norway</td>
<td>4,300,000</td>
<td>16,915</td>
<td>127</td>
<td>127</td>
<td>127</td>
<td>2.95</td>
</tr>
<tr>
<td>Netherlands</td>
<td>15,500,000</td>
<td>287</td>
<td>24</td>
<td>24</td>
<td>24</td>
<td>1.85</td>
</tr>
<tr>
<td>Poland</td>
<td>39,000,000</td>
<td>115,000</td>
<td>1,000</td>
<td>1,000</td>
<td>1,000</td>
<td>2.56</td>
</tr>
<tr>
<td>Portugal</td>
<td>10,000,000</td>
<td>25,000</td>
<td>480</td>
<td>480</td>
<td>480</td>
<td>4.80</td>
</tr>
<tr>
<td>Slovenia</td>
<td>2,000,000</td>
<td>4,980</td>
<td>81</td>
<td>81</td>
<td>81</td>
<td>4.05</td>
</tr>
<tr>
<td>Spain</td>
<td>39,000,000</td>
<td>1,500</td>
<td>60</td>
<td>60</td>
<td>60</td>
<td>3.85</td>
</tr>
<tr>
<td>Sweden</td>
<td>8,900,000</td>
<td>25,000</td>
<td>200</td>
<td>200</td>
<td>200</td>
<td>2.25</td>
</tr>
<tr>
<td>Switzerland</td>
<td>7,000,000</td>
<td>14,178</td>
<td>220</td>
<td>220</td>
<td>220</td>
<td>3.14</td>
</tr>
<tr>
<td>Turkey</td>
<td>65,000,000</td>
<td>85,000</td>
<td>962</td>
<td>962</td>
<td>962</td>
<td>1.48</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>57,000,000</td>
<td>188,000</td>
<td>129</td>
<td>129</td>
<td>129</td>
<td>0.23</td>
</tr>
</tbody>
</table>
Activities of PRM Doctors in Europe 2005 ESPRM

<table>
<thead>
<tr>
<th>Private Practice &lt; 9%</th>
<th>Private Practice between 10%–39%</th>
<th>Private Practice &gt; 39%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Finland</td>
<td>Belgium</td>
<td>Austria</td>
</tr>
<tr>
<td>Germany</td>
<td>Denmark</td>
<td>Czech</td>
</tr>
<tr>
<td>Hungary</td>
<td>France</td>
<td>Greece</td>
</tr>
<tr>
<td>Ireland</td>
<td>Italy</td>
<td>Switzerland</td>
</tr>
<tr>
<td>The Netherlands</td>
<td>Norway</td>
<td></td>
</tr>
<tr>
<td>Poland</td>
<td>Portugal</td>
<td></td>
</tr>
<tr>
<td>Slovenia</td>
<td>Spain</td>
<td></td>
</tr>
<tr>
<td>Sweden</td>
<td>Turkey</td>
<td></td>
</tr>
<tr>
<td>United Kingdom</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Disability equipment, assistive technology, prosthetics, orthotics, technical supports and aides
- Patient education
- Rehabilitation nursing

PRM specialists work in both public and private health services to varying degrees. This varies between different European countries (Table 22-6).

The European Bodies in PRM

UEMS PRM Section

The purpose of the UEMS PRM section is to promote the specialty in a professional capacity and to harmonize the specialty at the European level through specialist training and continual professional development through revalidation. They work to develop clinical standards in practice and to help facilitate the required research to develop it further. They are accountable (as with other specialties) to the UEMS and have started to work closely with the European Commission and the Council of Europe. They have three main committees under an Executive Committee.

- Training and Education Committee (a statutory committee—the European Board of PRM)
- Clinical Affairs Committee
- Professional Practice Committee

The Académie Européenne de Médecine de Réadaptation—European Academy

This body of up to 50 senior doctors in the specialty across Europe was created in 1969. Academics are invited on the basis of their distinguished contribution to the specialty, particularly to its humanitarian aspects. The aim of the Académie is to improve all areas of rehabilitation for the benefit of those who need it.

The European Society of Physical and Rehabilitation Medicine

The European Society of Physical and Rehabilitation Medicine (ESPRM) was founded in 2003 and is concerned with research and teaching of PRM in Europe. It succeeded the European Federation of Physical Medicine and Rehabilitation (established in 1963) and aims to coordinate European activities and be a vehicle for scientific exchange. The society offers individual membership to all eligible PRM specialists and federated membership of the national PRM societies in Europe. Individual membership is free of charge. The ESPRM is establishing an interactive electronic platform (www.esprm.org), where information can be found on research projects, grants, and funding. The Web site offers updated information about courses, congresses, exchange funding, etc.

The ESPRM organizes biennial scientific congresses. The main topics of the previous congresses were clinical standards, measurement of outcomes and effective interventions in neurological rehabilitation, musculoskeletal rehabilitation, and amputee rehabilitation, advances in PRM—traditional and modern concepts, evidence-based rehabilitation, PRM in lung transplant and in diabetes mellitus, and the scientific basis of PRM research.

MEDITERRANEAN, MIDDLE EAST, AND AFRICA

The creation of the “Mediterranean Forum of Physical and Rehabilitation Medicine” (MFPRM) and its role in the development of Rehabilitation Services in the Mediterranean Countries: Over 20 years ago, Rehabilitation experts from around the Mediterranean Basin, expressed their wish to do something to promote PRM and rehabilitation services in the Mediterranean countries. They wanted something to unite them as a family, a mechanism for mutual help and for strengthening their relationships. The first step was taken in Israel when Prof. Haim Ring proposed to the Israeli Association of PRM to organize the first Mediterranean Congress in Herzliya, Israel. This First Mediterranean Congress of PRM took place in May 1996. The symbol chosen was a map of the Mediterranean Basin with a dove of peace flying above and the motto of “Rehabilitation without frontiers” underneath. About 500 people from 42 countries from all continents participated. The Mediterranean Basin was especially well represented with participants from 11 countries. During that congress, a special meeting took place under the title: “A Mediterranean PMR Society: Is it viable?” It was decided at that meeting to continue all agreed upon activities and to hold meetings every 2 years. Subsequent meetings were held in Valencia, Spain; Athens, Greece; Syracuse, Sicily; Antalya, Turkey; and Vilamoura, Portugal. The number of participants grew from one conference to another. More importantly the number of representative Mediterranean countries grew steadily. During the 3rd Mediterranean Congress of PRM, the MFPRM was created. At the 4th Mediterranean Congress, the MFPRM statute was accepted. It contained the following provisions:

A. Mission and Goals:

1. MFPRM will be the scientific Mediterranean body for physicians who work in the fields of PRM in the countries surrounding the Mediterranean Sea as well as adjacent countries.
TABLE 22.7 PRM Manpower in Mediterranean Countries in 2006

<table>
<thead>
<tr>
<th>PRM in Mediterranean Countries—October 2006</th>
<th>In Training</th>
<th>In Practice</th>
<th>Per 100,000 h</th>
<th>Inhabitants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Croatia</td>
<td>41</td>
<td>288</td>
<td>7.5</td>
<td>4,381,000</td>
</tr>
<tr>
<td>Greece</td>
<td>32</td>
<td>174</td>
<td>1.87</td>
<td>11,000,000</td>
</tr>
<tr>
<td>Spain</td>
<td>280</td>
<td>1,600</td>
<td>4.18</td>
<td>45,000,000</td>
</tr>
<tr>
<td>Italy</td>
<td>500</td>
<td>5,000</td>
<td>9.8</td>
<td>56,000,000</td>
</tr>
<tr>
<td>Slovenia</td>
<td>20</td>
<td>70</td>
<td>4.49</td>
<td>2,003,000</td>
</tr>
<tr>
<td>Portugal</td>
<td>48</td>
<td>489</td>
<td>5.18</td>
<td>10,356,000</td>
</tr>
<tr>
<td>France</td>
<td>120</td>
<td>1,816</td>
<td>3.23</td>
<td>60,000,000</td>
</tr>
<tr>
<td>Turkey</td>
<td>300</td>
<td>1,500</td>
<td>1.18</td>
<td>72,000,000</td>
</tr>
<tr>
<td>Cyprus</td>
<td>5+4</td>
<td>0.9</td>
<td></td>
<td>1,000,000</td>
</tr>
<tr>
<td>Malta</td>
<td>1</td>
<td>0.25</td>
<td></td>
<td>400,000</td>
</tr>
<tr>
<td>Servia</td>
<td>120</td>
<td>600</td>
<td>7.59</td>
<td>10,000,000</td>
</tr>
<tr>
<td>Israel</td>
<td>30</td>
<td>150</td>
<td>1.2</td>
<td>7,000,000</td>
</tr>
<tr>
<td>Egypt</td>
<td>50</td>
<td>600</td>
<td>2</td>
<td>75,000,000</td>
</tr>
<tr>
<td>Jordan</td>
<td>14</td>
<td>65</td>
<td>1.44</td>
<td>5,500,000</td>
</tr>
<tr>
<td>Romania</td>
<td>70</td>
<td>700</td>
<td>3.5</td>
<td>22,000,000</td>
</tr>
<tr>
<td>Algeria</td>
<td>90</td>
<td>285</td>
<td>1.14</td>
<td>32,790,000</td>
</tr>
<tr>
<td>Albania</td>
<td>20</td>
<td>0.65</td>
<td></td>
<td>3,100,000</td>
</tr>
<tr>
<td>Totals</td>
<td>1,739</td>
<td>13,412</td>
<td>3.54 mean</td>
<td>427,530,000</td>
</tr>
</tbody>
</table>

2. MFPRM will facilitate exchange of information regarding different aspects of rehabilitation research, multicenter trials, national and regional projects, and meetings and congresses throughout the Mediterranean Basin.

3. MFPRM will organize a biennial Mediterranean Congress of PRM in one of the countries that has active members in the Forum.

4. On a national level, MFPRM, in close cooperation with the national society of PRM, will influence governments to support initiatives and collaborations in the field of teaching, development of services, and research in PRM. It will provide governments with information about the contents and evidence-based efficacy of PRM.

B. Membership in MFPRM:

MFPRM is open to all qualified physicians who specialize in PRM and work in one of the countries bordering the Mediterranean or in one of the adjacent countries. Membership is free to all.

C. The society is conducted by a board and a general assembly is organized every 2 years during the Mediterranean Congress.

In November 2005, the *Euro-Mediterranean Rehabilitation PRM School* opened its gates, under the auspices and involvement of the MFPRM, and was successful in part due to the active and successful efforts of Franco Cirillo and Alessandro Giustini. The school is run by a consortium consisting of the Italian PRM Society “SIMFER” and the Sicilian regional group members of MFPRM. In addition, the school is supported by a number of Sicilian Universities and the Syracuse Regional province authorities. The Scientific Committee is made up of members from the Scientific Associations. The school is financially supported by the Syracuse Town Council, the regional province, and the Tourist Board of Syracuse. Thanks to these contributions, the students need only to pay their airfare. Tuition, food, and lodging are free. Every year, the curriculum focuses on a specific topic and is presented by specialists in that area. The courses are a week long and English is the official language of the courses. The first year’s sessions were on education and on research in rehabilitation. In November 2006, the second session took place. The main topic of the course was “Pain and Mobility.” Thirty-six young professionals, members of the Forum, from Italy, Greece, Egypt, Portugal, Serbia, Israel, Slovenia, Turkey, and Spain came and received both theoretical and practical training. The third session was held in November 2007 and hosted about 40 students. The topic was “Emergency Medicine & Rehabilitation.” At the 6th Mediterranean Congress of PRM which took place in Vilamoura, Portugal, in October 2006, all participants from developing countries were waived registration fees and participated for free. The 7th Mediterranean Congress of PRM was held in Portorose, Slovenia, in September 2008. The number of members of the MFPRM, under the presidency of Prof. Nicolas Christodoulou, residing president since 2002, has been continuously increasing. The members hail from the following countries: Portugal, Spain, France, Italy, Malta, Slovenia, Croatia, Serbia, Bosnia, Montenegro, Albania, Greece, Bulgaria, Romania, Moldova, Turkey, Cyprus, Syria, Lebanon, Israel, Jordan, Egypt, Tunisia, and Algeria. In the course of the six Mediterranean congresses it became possible through interactive sessions to collect information about the PRM manpower in the Mediterranean countries and to design the desired model of rehabilitation services that would be presented to the governments in the area. Table 22.7 illustrates the manpower situation in the Mediterranean countries which responded to the survey.

The median number of professionals in practice and in training in 2006 was 3.54 per 100,000 inhabitants, but the mean desired number according to the needs, as expressed by the participants in the survey, is 4.29 per 100,000 inhabitants. This means that in the Mediterranean Area at least 3,200 additional specialists in PRM are needed. The conclusions reached from the survey resulted in a recommendation for a minimal network of medical rehabilitation services to be developed in each Mediterranean country. The model of such a network is as follows:

1. **Phases of rehabilitation intervention:**
   a. Intensive rehabilitation
   b. Completion of the recovery process and of the rehabilitation project
   c. Maintenance and prevention of further disability
2. Organization levels of Rehabilitation Services in three phases:
   a. hospitalization
   b. residential
   c. outpatient
3. Network of rehabilitative services delivered in
   a. Long-term hospital settings
   b. Outpatient recovery and functional reeducation clinics and hospitals
   c. Extra-hospital rehabilitation settings
   d. Outpatient rehabilitation centers
4. Intensive rehabilitation interventions delivery:
   a. Multiple specialty and/or single specialty hospital facilities
   b. Extra-hospital rehabilitation settings
5. Rehabilitation Centers focused on special PRM branches:
   a. Spinal cord injuries
   b. Pediatrics
   c. Neuropsychological problems
   d. Traumatic brain injuries (including strokes)

It was mutually agreed by MFPRM members to try to implement this simple basic model of a rehabilitation services network in each country. It was also agreed upon to further the development of rehabilitation services, until each Mediterranean country implements comprehensive rehabilitation services. MFPRM has gained the support of world and international bodies. Since the first Mediterranean PRM Congress in Herzliya in 1996, all of the “in-between congresses meetings” of the International Rehabilitation Medicine Association (IRMA), the International Federation of Physical Medicine and Rehabilitation (IFPMR) and their merged body: International Society of Physical and Rehabilitation Medicine (ISPRM), as well as those of the European Federation of Research in Rehabilitation (EFRR), take place during the Mediterranean PRM Congresses (41–43).

Middle East—Kuwait
There is a very active PMR Society which has organized several scientific events under the presidency of Dr. Abdulla Eyadeh. The main PRM premises are in the PMR Hospital in Kuwait, which is an 81-bed facility with Dr. Abdulla Eyadeh as the head of the institution. It caters to persons with various physical disabilities from Kuwait as well as from other Middle Eastern countries. It has several specialized services and is also a center for various academic activities. The PMR Hospital was granted accreditation in 2004 after having fulfilled the guidelines set by the Kuwaiti Ministry of Health. The First Conference of the PMR Society of Kuwait and the First Regional Meeting of ISPRM was conducted in April 2000. The PMR Hospital in Kuwait has, to date, organized eight workshops and conferences since 2001. There is an expansion of the diagnostic facilities in the hospital which already has a fully equipped EMG and Urodynamics clinic. A comprehensive cardiological rehabilitation program has been started at PMR Hospital in Kuwait as well. This program aims to improve the cardiac health in those with high risk factors and rehabilitates cardiac and pulmonary patients (44).

Saudi Arabia
The existing rehabilitation services in the country organized the 4th National Rehabilitation Symposium and the Regional Meeting of the ISPRM at the Armed Forces Hospital in Riyadh, Saudi Arabia, in February 2001. There were 25 speakers from Saudi Arabia as well as lecturers from other Arab and non-Arab countries. Numerous relevant topics were presented reflecting the rehabilitation activities within the country. These included recent developments in stroke, traumatic brain injury, and SCI rehabilitation. A workshop on Training in Rehabilitation Medicine was held, which provided opportunities for the participants to obtain information and discuss rehabilitation training options available in Saudi Arabia and other countries (45).

Iraq
The National Spinal Cord Injuries Centre (formerly IBN Al-Kuff Military SCI Hospital) is the only rehabilitation center in the country. It is located in Baghdad and has a 130-bed capacity. It was established and opened in October 1982. Its Director is Dr. Kaydar M. Al-Chalabi, a member of the ISPRM. Until March 2003, about 3,200 SCI patients underwent a comprehensive rehabilitation program and were followed up regularly. Initially, a Danish medical rehabilitation team worked together with the Iraqi team for 4 years and established the working system. When they left at the end of 1986, the Iraqi rehabilitation team successfully took over the operation. The success rate of this center was similar to other spinal centers. During the last 13 years (1991 to 2003) a lot of difficulties were encountered, including a shortage of medicines, assistive devices, and supplies. However, the center continued operating and a lot of scientific papers from that center were presented in local Arab and international conferences. The hospital academically is linked to the College of Medicine of Baghdad University. Because of the war, the center had stopped operating for a while due to severe damages to the buildings. Rebuilding of the center was completed in 2004 and resumed operations. Despite the difficulties the personnel faced during the imposed sanctions, during and immediately after the last war and the current unstable security status of Iraq, they continue working under the leadership of Dr. Kaydar Al-Chalabi, keeping the PRM specialty in Iraq alive (46).

Africa (Sub-Saharan)
In this region the rehabilitation services do not yet follow the rehabilitation models of the rest of the world. Although in some countries, such as in South Africa, Zimbabwe, and very recently Ethiopia, the rehabilitation services are better organized and will be mentioned separately. Centers for protection and rehabilitation of torture victims, which often extend their rehabilitation services to nontortured individuals, can be found in the following countries: Democratic Republic of the Congo (Centre Psycho Médical pour la Réhabilitation des Victimes de la Torture—SOPROP, OASIS Centre for Treatment and Rehabilitation of Victims of Torture and Trauma, SAVE
Congo DR, Cameroon (Center for Rehabilitation and Abolition of Torture, Trauma Centre Cameroon), Ethiopia (Rehabilitation Center for Victims of Torture in Ethiopia), Kenya (Independent Medico Legal Unit), Nigeria (Consulting Centre for Constitutional Rights and Justice), Rwanda (Forum des Activistes Contre la Torture), Senegal (Victimes de Violences Rehabilitées, le Centre de Soins du CAPREC), Republic of South Africa (The Trauma Centre for Survivors of Violence and Torture), Sudan (Amel Center for Treatment and Rehabilitation of Victims of Torture), Uganda (African Centre for Treatment and Rehabilitation of Torture Victims), and Zimbabwe (AMANI Trust, Mashonaland Programme) (47).

South Africa
There are several sites that offer medical rehabilitation services. In Johannesburg, within the Barney Morwitz Hospital there is a 21-bed ward for general rehabilitation. The Johannesburg Cardiac Rehabilitation Center and the Hillbrow Acute Hospital Pain Clinic are set up in an interdisciplinary manner utilizing all aspects of rehabilitation; The Acorndock/Tintswald Community Hospital offers a CBR program model called “WITS/Tintswald rehab worker training”. In Cape Town, there is the Groote Shuur Hospital with a wing devoted to stroke rehabilitation, the Conracid Hospital has an SCI unit with 169 beds; In Port Elisabeth Aurora Hospital, a registered private hospital is dedicated to PMR and subacute care is a collaboration between the QuadPara Association of South Africa and the private sector. Throughout South Africa there are several private RM clinics dealing with all kinds of rehabilitation problems (48,49).

Zimbabwe
Public and privately owned rehabilitation hospitals, schools, vocational training centers, and associations for and of people with disabilities comprise key rehabilitation service providers. The Zimbabwean Ministry of Labour and Social Welfare (MLSW) provides grants for the vocational training of people with disabilities who are registered with training centers in the country. The Ministry of Health and Child Welfare (MHCW) has rehabilitation units at each of the 15 referral hospitals in the country. It also shares responsibility with the MLSW for rehabilitation services provided at the nine national rehabilitation centers. The MLSW has responsibility over the administration and management of the centers as well as the vocational skills programs, whereas the MHCW runs the hospital within the rehabilitation centers. A majority of the large corporations in industry and commerce have rehabilitation departments for their workers. Organizations for and of people with disabilities are more involved with vocational and psychosocial rehabilitation than with occupational and physiotherapy (50).

Ethiopia
In collaboration with “Light for the World” (Austria), a CBR program started. The program runs now in six Kebeles (the smallest administrative unit in Ethiopia) in North Gondar: two in the city of Gondar and four in the rural villages of North Ethiopia. It is focused on giving proper care and follow-up to children who have clubfeet or hydrocephalus, to women who suffer from fistulas after obstructed labor, children who become disabled due to obstructed labor or unsafe and unsuccessful abortions, and to others. The priorities of this program are to educate the patients, the medical providers and convince the Ethiopian community that just as important as saving lives is rehabilitating and adding quality to life. The CBR program is a service program but they also train physiotherapy students. The University aims to produce graduates who are equally willing to work in urban and rural areas. In 2006, the first 80 physiotherapists graduated. This started a new and exciting period in Ethiopia, as the new graduates face a challenge together with other health professionals in developing rehabilitation services in Ethiopia (51).

ASIA AND OCEANIA
About 60% of the world’s population, of which 5% to 10% are disabled, reside in Asia and Oceania (52). The wide variation in culture and economic development within the region has a significant impact upon the breadth and scope of rehabilitation services. In the developed countries of the region, the increasing demand for rehabilitation care is related to the aging population, particularly musculoskeletal disorders and stroke rehabilitation. Nevertheless, most Asian countries are developing countries and have limited health care and welfare service resources. Malnutrition, infectious disease, traffic and work accidents, natural disasters, and even wars continue to play an important causative role for disasters in many Asian countries (53). As a result of limited resources and limited expertise in rehabilitation, a CBR approach has been advocated and practiced in some developing countries in Asia (54–59).

Given the diversity in the practice of medical rehabilitation among the Asian and Oceanic countries, a formal collaboration has not been developed until recently. The ASEAN (West Asian Region) Rehabilitation Medicine Association Congress was established in 1998 and held its first meeting the same year in Thailand. A subsequent meeting was held in 2002 in the Philippines. Between these two meetings, a Millennium Asian Symposium, sponsored by the Japanese Association of Rehabilitation (60), was held in Japan in 2001. At this meeting, representatives from various Asian countries discussed the establishment of an Asian Society of Rehabilitation Medicine. Thereafter, the Korea-Japan Joint Conference on Rehabilitation Medicine was held in Korea in 2002 and Japan in 2004. These joint conferences stimulated the establishment of the Asian Society of Physical and Rehabilitation Medicine. These meetings, coupled with a rapid economic development in some Asian countries, prompted a growing interest to establish a regional society to facilitate the development and promotion of the PRM Specialty. As a result, a regional society, the Asian Oceania Society of Physical and Rehabilitation Medicine (AOSPRM), was officially established in Seoul during the 4th World Congress of ISPRM 2007 after initial task force
meetings that were held in Sao Paulo, Brazil, and Zhuhai, China, in 2005 and 2006, respectively. Hereafter, AOSPRM plans to hold biannual congresses, with the first held in 2008 in Nanjing, China. This first regional conference of the Asian and Oceanic region attracted more than 700 participants from its 19 member countries and included attendees from Europe and North and South America. During the conference, further closer collaboration among the countries in the region in terms of training, research, and exchange was put forward. A decision on a minor amendment of the official name of AOSPRM as “Asia-Oceanian Society of Physical and Rehabilitation Medicine” was also made during the Nanjing conference. The Asia-Oceanian conference is biannual from 2008 onward to promote academic exchange within the region and collaboration with other regions around the world.

As shown in Table 22-8, 19 countries or administrative regions are currently members of this newly established regional society and more countries are expected to join in the near future. Different countries use different designations for PMR. The three most commonly used designations are RM, PMR, and PRM, while RM is the more commonly used term. Japan and Philippines established academic societies in the 1960s while two thirds of the other countries established their academic societies in 1980 or later. The specialty of PMR is relatively new in this region. Physiatrists number 1% or less of the total number of practicing doctors in most countries (61,62). In some countries the specialty was only established in the last few years, and some are still lacking sufficient physiatrists for the specialty and academic societies to develop.

The degree of maturity of training programs in the specialty of PRM also varies in an analogous fashion to the specialty itself. Within the countries or administrative regions that have well-established training programs, there are also variations in entry into and the years of required PMR training. In some countries, 2 to 3 years of general training in medicine along with passing an entry examination is required before commencing formal PMR training, after which a board certification examination is required for accreditation as a physiatrist. In general, the years spent in core training for PMR varies from 3 to 5 years in most countries that offer training programs (61,62). Features of convergence and divergence of training programs between and within countries occur. For example,

### Table 22.8: International PMR: Comparative Data & Characteristics

<table>
<thead>
<tr>
<th>Country or Administrative Region</th>
<th>Population (Millions)</th>
<th>PMR Association</th>
<th>Year of Establishment</th>
<th>Number of Physiatrists</th>
<th>PMR Training (Years)</th>
<th>Board Certification Examination</th>
<th>CME/CPD Activities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia</td>
<td>21.2</td>
<td>Australasia Faculty of RM&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1993</td>
<td>300</td>
<td>2 + 4</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>New Zealand</td>
<td>4.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bangladesh</td>
<td>158.7</td>
<td>Bangladesh Association of PRM</td>
<td>1995</td>
<td>25</td>
<td>3</td>
<td>Yes</td>
<td>n/a</td>
</tr>
<tr>
<td>Brunei</td>
<td>0.39</td>
<td>n/a</td>
<td></td>
<td>1</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>China</td>
<td>1,322.5</td>
<td>Chinese Society of PMR&lt;sup&gt;b&lt;/sup&gt;</td>
<td>1985</td>
<td>10,000</td>
<td>5</td>
<td>In preparation</td>
<td>Yes</td>
</tr>
<tr>
<td>Chinese Taipei</td>
<td>23</td>
<td>Chinese Association of RM</td>
<td>1983</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hong Kong SAR</td>
<td>7.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>India</td>
<td>1,129</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indonesia</td>
<td>231.6</td>
<td>Taiwanese Academy of PMR</td>
<td>1971</td>
<td>640</td>
<td>4</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Iran</td>
<td>71.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Japan</td>
<td>127.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Korea</td>
<td>48.2</td>
<td>Korean Academy of RM</td>
<td>1972</td>
<td>947</td>
<td>4</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Laos</td>
<td>5.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Malaysia</td>
<td>27.5</td>
<td>Malaysian Association of RM</td>
<td>2004</td>
<td>15</td>
<td>2 + 4</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Mongolia</td>
<td>2.7</td>
<td>Mongolian Society of PRM</td>
<td>2005</td>
<td>130</td>
<td>1</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>Philippines</td>
<td>88.7</td>
<td>Philippine Academy of RM&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1974</td>
<td>300</td>
<td>3</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Singapore</td>
<td>4.7</td>
<td>Socioity of RM Singapore</td>
<td>2005</td>
<td>16</td>
<td>3 + 3</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Thailand</td>
<td>62.9</td>
<td>Royal College of Physiatrists of Thailand</td>
<td>1998</td>
<td>338</td>
<td>3</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Vietnam</td>
<td>87.4</td>
<td>Thai RM Association</td>
<td>1988</td>
<td>1400</td>
<td>3</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Vietnam Rehabilitation Association</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup> Formerly Australia College of Rehabilitation Medicine 1980–1992.

n/a, not established yet; PRM, physical and rehabilitation medicine; PMR, physical medicine and rehabilitation; RM, rehabilitation medicine; CME/CPD, Continuous Medical Education/Continuous Professional Development.
Australia and New Zealand share the same academic society and training system while China has a different training program in the mainland from her special administrative region, Hong Kong. Activities for CME or CPD have been established in most countries that have the board certification examination.

In some Asian countries, the practice of PMR has been to some extent influenced by Eastern Medicine. Given the long history of Traditional Chinese Medicine (TCM) or Oriental Medicine in some countries such as China, some components of these systems have been incorporated into the daily practice of PMR. An international survey on stroke rehabilitation showed that in some of the Asian countries, Oriental Medicine techniques such as acupuncture and massage were more commonly used for stroke patients in comparison with countries in other parts of the world (63). This has triggered interest in the “East-meets-West” program whereby medical expertise is exchanged. The other important aspect of providing rehabilitation services to the disabled in some Asian countries has occurred through CBR. However, there have not been any formal training components in terms of CBR within the PMR training programs in any of these countries. This will be one of the important areas of PMR training to develop in the region.

With the advocacy of evidence-based medicine and practice, TCM and Oriental Medicine has opened up areas for research, which will likely lead to evidence for the role of TCM and Oriental Medicine in PMR. Instead of separate practices of Western and Oriental Medicine, a combined approach for a “Universal Medicine” will be an area of interest to develop among the physiatrists in the region. This foreseeable development will likely be brought about by the strengthening of collaboration within the AOSPRM to facilitate PMR specialty training and research in the region. Interregional collaboration such as between AOSPRM and European Society of PMR is also developing. Even though the PMR specialty is relatively immature in the region, it will have a foreseeable bright future through a joint effort within the region and collaboration with other regions.

INTERNATIONAL EXCHANGE PROGRAMS

International PMR exchange programs exert a powerful force on world health both in practical terms and with respect to humanistic values (64). Pragmatically speaking, global exchange programs can significantly enhance care for international citizens, by altruistically promoting the sharing of information, knowledge, and technical expertise (“teach others and learn from others”) among disparate nations across the globe. Additionally, international PMR exchange programs may broaden research horizons and collaborative vistas, potentially leading to new scientific discoveries. Humanistic projects and global exchange endeavors can unite nations (3).

A series of biennial international interactive workshops: “Building Global Bridges: International PMR Educational Exchange” were conducted in the Czech Republic (2003), Brazil (2005), and Korea (2007) and included representatives from many of the member nations of the international rehabilitation community. These workshops served as a candid forum for discussion of views and insights about visiting opportunities, educational programs, exchange possibilities, humanitarian initiatives, local rehabilitation resources, and residency curriculum highlights. The historical event included panelists from sample countries, representing the five continents of Africa, America, Australia, Asia, and Europe, who were committed to participating in global physiatric didactic exchange programs. To further the goal of international cooperation and education “across the oceans”, representatives of “developing” and “developed” countries were active participants (65,66).

Promotion of international goodwill through education, faculty-student exchange, and humanitarianism is a noble physiatric goal and is a major objective of the ISPRM Faculty-Student Exchange Committee.

The committee’s activity began in earnest in 2002 when then President John Melvin (United States) and President-elect Haim Ring (Israel) recognized the vital importance of creating an ISPRM outreach arm dedicated to promoting international educational exchange. That year, the Exchange Committee was founded. With the ongoing encouragement and mentorship of the ISPRM Board and leadership including Professor Linamara Batistella and Professor Chang-il Park, the committee has continued to promote international exchange activity among physiatrists internationally. Through its various international humanitarian projects, it has touched the lives of many people with disabilities on a global level.

The ISPRM Faculty-Student Exchange Committee serves as an information resource, international liaison, and active placement service for educational, advocacy, and humanitarian opportunities in PRM, across all continents. The committee facilitates placement of physiatrists in rotations and observerships as well as in rehabilitation-related humanitarian projects based on geography and need.

The committee maintains a Web site www.isprm-edu.org which provides an activity overview as well as a photo album highlighting outcomes and successes. The exchange committee serves a diverse group of rehabilitation professionals and facilitates international placement in educational, research, and humanitarian rotations. Participants in the program represent virtually all stages of education and levels of academic achievement. The core groups served include senior medical students, PMR residents, and attending physiatrists in private practice, PMR academic faculty, and PMR researchers. The committee also supports other nonphysician rehabilitation professionals in obtaining rotational opportunities (Fig. 22-3).

The ISPRM Exchange Committee serves several fundamental roles including facilitating education and research opportunities, coordinating disability advocacy and humanitarian projects as well as other functions shown in Figure 22-4.

Specific objectives include advancing knowledge of RM and techniques, improving quality of care for developed and developing nations, arranging lectures, workshops and seminars, and providing humanitarian assistance during and after times of crisis, for example, tsunamis, hurricanes (67), and cyclones.
The committee performs a variety of tasks consistent with its mission and vision. Some of these activities are:

1. Handling daily e-mail and postal mail correspondence from rehabilitation physicians interested in additional information about global exchange opportunities
2. Gathering background information about prospective students and visitors (CVs, resumes, letters of reference) and determining areas of interest to ensure an appropriate "match" and a meaningful educational experience
3. Building networks of "exchange partners" and securing participating academic faculty and institutions by sending invitational e-mail queries
4. Monitoring "day-to-day" progress and logistics of actively rotating students, residents, and faculty, especially when those doing rotations spend time in more than one observership during a particular "placement". In this capacity, the chair of the committee can be viewed as a cross between a "camp counselor" and a residency program director
5. Developing preplacement administrative infrastructure documents for the support of placements, e.g., waiver forms, liability disclaimer documents, and academic institution sign-offs
6. Interfacing with institutional physicians, administrators, as well as governmental officials on matters pertaining to educational and humanitarian exchange opportunities. These activities are conducted in concert with input and guidance from the ISPRM executive leadership
7. Maintaining and updating an outcome focused web-based photo journal and textual narrative which reviews particular placements that the committee has made in addition to tracking global exchange activities
8. Submission of articles and updates pertaining to committee activity to "News & Views" and other academic publications
9. The activities of the committee are conducted in a team fashion with individual members of the committee providing input, direction, and collaborative assistance
10. Interfacing with other ISPRM members (including past and present leadership) to apprise them of committee activity and to seek their input
11. Submission of an annual written report to the ISPRM Board as well as meeting regularly with the plenary ISPRM Board to provide a committee report

The process by which physiatrists are placed in global opportunities is represented in Figure 22-5.

The committee remains committed to promoting international placements for traveling physiatrists and welcomes the active help and input of potential host institutions abroad.

The exchange committee proposes the establishment of an ISPRM recognized "global faculty," comprised of physiatrists who have hosted visiting exchange students and physicians. International institutions and physicians interested in participating may contact the committee, the Chair of the Committee, Mark Young: markyoung123@gmail.com or through the Web site: www.isprm-edu.org.
THE INTERNATIONAL REHABILITATION ORGANIZATIONS

As time has elapsed since the beginning of the modern era, colleagues from different parts of the world have looked to come closer and organize themselves into international organizations, meetings, courses, and the like. The International Federation of Physical Medicine was created in 1950 as a federation of existing national societies subsequently changing its name to the IFPMR. The first international meeting took place in 1952 in London (United Kingdom). Later congresses took place in 1956 (Copenhagen), 1960 (Washington), 1964 (Paris), 1968 (Montreal), 1972 (Barcelona), 1976 (Rio de Janeiro), 1980 (Stockholm), 1984 (Jerusalem), 1988 (Toronto), 1992 (Dresden), and 1995 (Sydney).

In 1968, the IRMA was created in order to provide a professional and social framework on an individual basis to those rehabilitation professionals who were not part of a national society. The first congress took place in 1970 in Milan and Turin (Italy). The second IRMA congress took place in 1974 in Mexico City (Mexico), then in 1978 in Basel (Switzerland), in 1982 in San Juan (Puerto Rico), in 1986 in Manila (Philippines), in 1990 in Madrid (Spain), in 1994 in Washington, DC (the United States), and in 1997 in Kyoto (Japan). The last IRMA's congress took place in 1999 in Washington, DC (the United States). Since both organizations shared many common officials on their respective boards, they merged to become the ISPRM—www.isprm.org, currently based in Assenede (Belgium).

The first ISPRM meeting took place in 2001 in Amsterdam (Holland) with 1,200 participants, the second in 2003 in Prague (Czech Republic) instead of the original designated place Jerusalem (Israel) due to local conditions at the time with 1,400 participants (95% of active participants), the third one in 2005 in Sao Paolo (Brazil) with 1,800 participants, and the fourth one in 2007 in Seoul (South Korea) with 2,400 participants from 75 countries. The number of participants, as well scientific presentations, steadily increased from one congress to the next. This gave the opportunity to colleagues in different parts of the world to participate in this unique event. The scope of the presentations also changed over time (68) with the incorporation of advanced technologies in rehabilitation (robotics, virtual reality, functional electrical stimulation) as well the new WHO-ICF, along with the more classical topics.

The last ISPRM congress was held in 2009 in Istanbul (Turkey). Future congresses will take place in 2011 in San Juan (Puerto Rico) and in 2013 in Beijing (China). Besides the pure international organization congresses, an increasing number of regional or continental rehabilitation organizations are conducting meetings, with a broad professional scope, and most of them are described in the respective subchapters: AMLAR in Latin America, the ESPRM and the EFRR in Europe, the MFPRM in the Mediterranean region, and the most recent being the AOSPRM composed of 17 national organizations that held its first meeting in Nanjing (China) in May 2008. Additional regional meetings are the so-called “Middle East Chapter” for Arab countries colleagues, the Trans-Pyrenean Congress (NE Spain and SW France) and the Baltic Rehabilitation congress for countries around the Baltic Sea, having its last meeting in Riga (Estonia) in 2008. Other societies deal with more specific aspects (movement analysis, etc.).

It can be safely said that these regional congresses serve as a means of integration at the regional level (“regionalism”) and the basis for the strengthening of the world ones.
THE INTERNATIONAL REHABILITATION JOURNALS—PUBLISHING IN PRM JOURNALS

For many years, rehabilitation professionals have strived to express their views, work, and scientific production in scientific PRM journals. Some of them have done and still do so in national and regional journals in their own languages. It is difficult to accurately determine the number of these journals on a world level but there are surely several dozens. Over 33 indexed journals appeared doing a computer search for the term rehabilitation journals in English. About two third of these journals are published in the United States and the rest were mainly in the United Kingdom. In addition, there are PRM journals published in English by publishing houses based in Holland, Sweden, and Italy. The declared main topic is either wide (general rehabilitation), or involve specific topics such as assistive technology, research in rehabilitation, bioengineering, nursing, psychological/psychiatric, occupational therapy, rehab management and counseling, etc. Other journals focus on a main or defined pathology: stroke, brain injury/head trauma, genera; neurological, spinal cord, prosthetic and orthotics, cardiopulmonary, back and musculoskeletal, burn care and rehabilitation, or cancer rehabilitation. Some are devoted to specific patient ages with their related pathologies such as pediatric or geriatric. The highest Impact Factor (determined by the number of citations and the general number of publications over a given period of time—usually 1 year) is held by the following general international rehabilitation journals in English: Neurorehabilitation and Neural Repair, Journal of Rehabilitation Medicine, Archives of Physical Medicine and Rehabilitation, Clinical Rehabilitation, American Journal of Physical Medicine and Rehabilitation, Brain injury, Disability and Rehabilitation, International Journal of Rehabilitation Research, among others.

In 2009, the AAPM&R initiated a new publication: The Journal of Injury, Function and Rehabilitation with a clinical emphasis. The issue of publishing in PRM journals, with the problems inherent to the submission, review (peer-review), and acceptance/refusal decision processes, were first discussed in a special session titled “Meet the Editor,” in the 2nd ISPRM World Congress in Prague (transferred from Jerusalem). This resulted in a joint paper by participants and chair, published in the congress proceedings (69) and at the 4th World Congress in Seoul in 2007. With a record attendance for that congress, participants generated a peer-reviewed article published simultaneously in three different first-line journals (70): The American Journal of Physical Medicine and Rehabilitation (USA), The Journal of Rehabilitation Medicine (Sweden), and the Disability and Rehabilitation Journal (UK). Topics discussed in the session included the reasons for publishing in peer-review journals, important considerations in submitting a manuscript, the peer-review process, the effect of electronic publishing—which leads to shorter publication times, the opportunity to preview papers, and more open access to journals. The discussions concluded that the field of PRM will continue to expand, with an audience having a broader range of scientific and clinical interests. Another important conclusion was that the ICF may be increasingly used as a framework in reporting. Intermediate meetings of this kind were also held in the 6th and 7th Mediterranean PRM Congresses held in Vilamoura (Portugal) in 2006 and Portorose (Slovenia) in 2008 respectively, bringing the discussion close to potential writers and reviewers in the region.

CONCLUSION

There are some words whose meanings are so universally accepted that definition is unwarranted and might be considered pedantic by some. Physical medicine is such a term to some but not to others, for most specialists in physical medicine have their own idea of its meaning. But there are as many definitions as there are countries, and within some countries there are many definitions, for this is a specialty that stems from many older specialties, each of which has its own national and international history.

Sidney Licht, 1960

From the historical perspective (71,72), there is no doubt that RM has come a long way since its inception. Although differences still persist among regions and systems, there is no doubt that significant and rapid changes occur in parallel across all the regions in the world, developed and developing countries alike. Regional and world congresses, the international PRM journals, and the ISPRM International Exchange Program have made great contributions to an international harmonization in residency and teaching programs, the understanding and building of common concepts and terms, the establishment of new rehabilitation systems, and the development of strategic plans to strengthen the stability of those systems in jeopardy.

Research in rehabilitation is thriving and the quality of the research programs as well their scope is dramatically improving, including animal experimental studies (73). One of the most salient features of the last years is the impressive recruitment, in number and mainly in quality of a cadre of active members that, whether they fill some formal position or not, are increasingly devoting themselves to the regional and international organizations. This has brought about a dramatic upgrade of standards of work and activities. By now a kind of “international carrier” has been created in the rehabilitation realm. Above all, if to judge by the different regional reports, there is a bright future for RM. The profession is closing the gaps among regions and is actively becoming global (2). It is the dawn of a new day.

ACKNOWLEDGMENT

Special thanks to Jeff Shames (Israel), for his help in all the stages of chapter preparation. The authors are grateful to the following people for their invaluable contribution in
the preparation of the chapter: Graciela Borelli (Uruguay), Tai Ryou Han (Korea), Jianan Li (China), Juan Lacuague (Uruguay), Rodney Li Pi Shan (Canada), Bryan O’Young (the United States), Mauro Zampolini (Italy).

Memorial:
The coauthors of this chapter wish to acknowledge posthumously the significant contributions of the late Professor Haim Ring in the initial iteration of this chapter. Professor Ring’s leadership skills and academic contributions to International Physiatry will always be remembered and treasured.

REFERENCES

39. UEMS-PRM-Section. UEMS-PRM-Section. 2005 [cited; Available from: www.euro-prm.org].
49. Gambell R. www.aurorahospital.co.za, rob@aurorahospital.co.za.
PART III

Major Conditions
INTRODUCTION
Stroke has been recognized since antiquity, and it remains a major cause of disability today. In this chapter, we review the causes, mechanisms, and symptoms of stroke, including a discussion of common stroke syndromes. We then address the evaluation of stroke cause and strategies to prevent recurrent stroke, an area that requires full integration into the rehabilitative phase of care. An overview of acute stroke management is provided, which segues into the various stages of rehabilitation after a stroke. The expanding emphasis on facilitating neurological recovery after stroke is highlighted as it has become a major focus of rehabilitation research since the prior edition of this book was published, and this shift is expected to continue. Lastly, we discuss some of the long-term sequelae of stroke and their management.

DEFINITION
A stroke is the sudden occurrence of permanent damage to an area of the brain caused by a blocked blood vessel or bleeding within the brain. Other causes of focal brain damage, such as traumatic injury to the brain, demyelinating lesions, brain tumors, brain abscesses, and others, can produce stroke-like symptoms but are not included in this definition. Strokes can be divided into two major categories: ischemic, generally caused by a vascular occlusion, and hemorrhagic, caused by bleeding within the parenchyma of the brain. Some classify nonparenchymal hemorrhage, such as subarachnoid hemorrhage (SAH) due to a ruptured intracranial aneurysm, as a form of hemorrhagic stroke as well.

A significant number of people who sustain a stroke do not reach medical attention due to a lack of symptoms or the failure to recognize the symptoms as requiring medical attention. (It is estimated that 28% of individuals aged 70 to 74 have had a silent cerebral infarction) (1). The most common symptom of stroke is focal weakness, though stroke can produce a wide range of symptoms such as sensory loss, speech and language disturbance, visual loss, etc. The resultant neurological deficits are generally referred to as impairments, which may or may not result in functional limitations often characterized as disability.

The objectives of stroke rehabilitation are to achieve a maximum level of functional independence; facilitate neurological recovery, minimize disability; successfully reintegrate back into home, family, and community; and reestablish a meaningful and gratifying life. Education of the stroke survivor and his or her family regarding secondary stroke prevention, including risk factor modification, and compliance with medical therapy for stroke prevention are responsibilities shared by the rehabilitation team. These goals are accomplished through exercise and other treatments to facilitate recovery and reduce impairments; functional training to compensate for residual impairments; and use of assistive devices, such as braces or wheelchair, to substitute for lost function. Successful rehabilitation also requires management of the many psychosocial issues that surround integration of the patient back to home and community.

EPIDEMIOLOGY
Each year approximately 780,000 Americans suffer a stroke, of which 600,000 are an initial event and the remainder are recurrent strokes (1). While the number of strokes in the United States continues to rise as a result of the aging of the population, their incidence has been stable in recent years at about 158 per 100,000 population. This contrasts with the death rate from stroke, which has fallen 24% from 1994 to 2004 (with a reduction in the total number of stroke deaths of 6.8% due to population growth and aging of the population). Stroke remains the third leading cause of death in the United States, after heart disease and cancer. Incidence is age related, being uncommon before age 50 but doubling each decade after age 55. Stroke is more common among men than among women in younger cohorts, but more common among women for individuals over age 85 (1).

RISK FACTORS AND PREVENTION
Prevention remains of critical importance in reducing the burden of morbidity and mortality from stroke. Risk factors for ischemic stroke have been identified and can be separated into two categories: those that are potentially modifiable and those that are not modifiable. The latter includes age, race/ethnicity, sex, and family history, whereas the potentially modifiable risk factors include hypertension, smoking, atrial fibrillation, diabetes, diet, obesity, sedentary lifestyle, and hyperlipidemia.

Risk factors for hemorrhagic stroke have also been identified, and while primary stroke prevention efforts generally focus...
on the risk factors for ischemic stroke, secondary prevention efforts for individuals who have already sustained a hemorrhagic stroke should include education with a particular focus on treating hypertension, refraining from excessive alcohol consumption, and avoiding the use of anticoagulant medications.

The following discussion focuses on potentially modifiable risk factors for ischemic stroke, listed in Table 23-1.

### Modifiable Risk Factors in Asymptomatic Individuals

**Hypertension** is the most important risk factor. The degree of risk increases with elevation and becomes particularly strong with levels higher than 160/95 mm Hg. Systolic hypertension and high mean arterial pressure represent parallel risks. In the Framingham study, a sevenfold increased risk of cerebral infarction was observed in patients who were hypertensive (2). Hypertension increases the risk of thrombotic, lacunar, and hemorrhagic stroke and increases the likelihood of SAH. Successful long-term treatment of hypertension greatly reduces the risk of these events, even after a prior stroke (3). The hypothesis that angiotensin receptor blockers might provide specific stroke prevention benefits beyond their blood pressure lowering effects has not been borne out (4).

Because of concerns regarding maintaining flow distal to critical cerebrovascular stenoses, aggressive management of hypertension immediately after stroke should be avoided. Studies have shown that abrupt reductions in blood pressure within the first 24 hours after stroke are associated with poorer outcomes (5). Clinicians typically aim to gradually bring blood pressure into the normotensive range over a period of weeks after a stroke, though little data exist to guide this practice.

**Heart disease** is an important risk factor for stroke. To some degree, this reflects the shared risk factors and pathophysiology of stroke and heart disease: hypertension and atherosclerosis. The risk of stroke is doubled in individuals who have coronary artery disease (6), and coronary artery disease accounts for the majority of subsequent deaths among stroke survivors. Atrial fibrillation and valvular heart disease increase the risk of cerebral infarction because both may cause cerebral emboli. In patients with chronic, stable atrial fibrillation, the risk of stroke is increased fivefold (7). When atrial fibrillation is a manifestation of rheumatic heart disease, the risk of embolic stroke is increased up to 17 times normal (7). Prevention of embolic stroke in these patients is best achieved by long-term anticoagulation with warfarin. Treatment carries the danger of intracranial hemorrhage, especially in elderly individuals and in those with impaired balance and who are likely to fall. When the risk of hemorrhage appears to be high, aspirin in a dose of 325 mg daily may be used as an alternative to warfarin in patients with nonvalvular atrial fibrillation, although aspirin is much less effective than warfarin in preventing embolism. Younger patients (below age 75) without valvular heart disease and who lack other risk factors (e.g., hypertension, diabetes, left ventricular hypertrophy, history of TIA or stroke) are at low risk for stroke and may be treated with aspirin rather than with warfarin (8).

**Diabetes, as an independent risk factor, doubles the risk of stroke.** While achieving good glycemic control is advised for all diabetic patients, studies have not shown a clear relationship between glycemic control and stroke risk (9,10).

**Smoking** has been shown to increase the risk of both ischemic stroke and SAH, with greater amounts of smoking carrying greater risk. Individuals who discontinue smoking have substantial reductions in stroke risk, approximating that of nonsmokers after several years (11). Counseling regarding smoking cessation is an essential component of programs to reduce the risk of recurrent stroke.

**Hyperlipidemia** increases the risk for stroke to a small extent. Treatment with cholesterol-lowering medications other than “statins” has not been found beneficial in reducing the risk of stroke. Conversely, treatment with statin medications has been found to reduce the risk of stroke, in individuals with or without hyperlipidemia (12), raising the possibility that the benefits of statins in stroke prevention may not be related solely to their lipid-lowering effects. Thus, treatment with statins is indicated for primary stroke prevention in patients with coronary heart disease or symptomatic atherosclerotic disease with low-density lipoprotein (LDL) cholesterol levels of 100 mg/dL or greater. In individuals at very high risk for stroke with multiple risk factors, the threshold for treatment is an LDL of 70 mg/dL (13).

**Homocysteine.** Elevated homocysteine have been found to be associated with a higher risk of ischemic stroke. While homocysteine levels can be lowered with supplementary vitamins (folate, vitamins B6 and B12), treatment has not been found to reduce stroke risk (14). Given the low cost and minimal risk of vitamin treatment to reduce homocysteine levels, the current recommendation is that such treatment is “reasonable” despite the lack of evidence of efficacy (15).

**C-reactive protein:** C-reactive protein (CRP) levels have been shown to correlate with risk of cardiovascular disease, including stroke, as have elevated fibrinogen levels. A recent study found that treating individuals with elevated CRP with statins reduced major cardiovascular events, despite normal lipid levels prior to treatment (16). While the role of testing CRP levels to guide stroke prevention remains a subject of debate, it may prove a useful adjunct to other tools for assessing risk.

---

**TABLE 23.1 Modifiable Risk Factors for Stroke**

<table>
<thead>
<tr>
<th>Risk Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hypertension</td>
</tr>
<tr>
<td>Heart disease</td>
</tr>
<tr>
<td>Ischemic/hypertensive</td>
</tr>
<tr>
<td>Valvular</td>
</tr>
<tr>
<td>Arrhythmias</td>
</tr>
<tr>
<td>Smoking</td>
</tr>
<tr>
<td>Diabetes mellitus</td>
</tr>
<tr>
<td>Elevated fibrinogen</td>
</tr>
<tr>
<td>Erythrocytosis</td>
</tr>
<tr>
<td>Hyperlipidemia</td>
</tr>
<tr>
<td>Erythrocytosis</td>
</tr>
<tr>
<td>Elevated fibrinogen</td>
</tr>
<tr>
<td>Hyperlipidemia</td>
</tr>
<tr>
<td>Smokers</td>
</tr>
<tr>
<td>Non-smokers</td>
</tr>
<tr>
<td>Smokers</td>
</tr>
<tr>
<td>Non-smokers</td>
</tr>
<tr>
<td>Younger patients (below age 75)</td>
</tr>
<tr>
<td>Without valvular heart disease</td>
</tr>
<tr>
<td>Left ventricular hypertrophy</td>
</tr>
<tr>
<td>History of TIA or stroke</td>
</tr>
<tr>
<td>Low risk for stroke</td>
</tr>
<tr>
<td>May be treated with aspirin</td>
</tr>
<tr>
<td>Rather than warfarin</td>
</tr>
<tr>
<td>Treatment carries the danger of</td>
</tr>
<tr>
<td>Intracranial hemorrhage</td>
</tr>
<tr>
<td>Especially in elderly individuals and those with impaired balance and who are likely to fall.</td>
</tr>
<tr>
<td>When the risk of hemorrhage appears to be high, aspirin in a dose of 325 mg daily may be used as an alternative to warfarin in patients with nonvalvular atrial fibrillation, although aspirin is much less effective than warfarin in preventing embolism. Younger patients (below age 75) without valvular heart disease and who lack other risk factors (e.g., hypertension, diabetes, left ventricular hypertrophy, history of TIA or stroke) are at low risk for stroke and may be treated with aspirin rather than with warfarin (8).</td>
</tr>
<tr>
<td><strong>Diabetes, as an independent risk factor, doubles the risk of stroke.</strong> While achieving good glycemic control is advised for all diabetic patients, studies have not shown a clear relationship between glycemic control and stroke risk (9,10).</td>
</tr>
<tr>
<td><strong>Smoking</strong> has been shown to increase the risk of both ischemic stroke and SAH, with greater amounts of smoking carrying greater risk. Individuals who discontinue smoking have substantial reductions in stroke risk, approximating that of nonsmokers after several years (11). Counseling regarding smoking cessation is an essential component of programs to reduce the risk of recurrent stroke.</td>
</tr>
<tr>
<td><strong>Hyperlipidemia</strong> increases the risk for stroke to a small extent. Treatment with cholesterol-lowering medications other than “statins” has not been found beneficial in reducing the risk of stroke. Conversely, treatment with statin medications has been found to reduce the risk of stroke, in individuals with or without hyperlipidemia (12), raising the possibility that the benefits of statins in stroke prevention may not be related solely to their lipid-lowering effects. Thus, treatment with statins is indicated for primary stroke prevention in patients with coronary heart disease or symptomatic atherosclerotic disease with low-density lipoprotein (LDL) cholesterol levels of 100 mg/dL or greater. In individuals at very high risk for stroke with multiple risk factors, the threshold for treatment is an LDL of 70 mg/dL (13).</td>
</tr>
<tr>
<td><strong>Homocysteine.</strong> Elevated homocysteine have been found to be associated with a higher risk of ischemic stroke. While homocysteine levels can be lowered with supplementary vitamins (folate, vitamins B6 and B12), treatment has not been found to reduce stroke risk (14). Given the low cost and minimal risk of vitamin treatment to reduce homocysteine levels, the current recommendation is that such treatment is “reasonable” despite the lack of evidence of efficacy (15).</td>
</tr>
<tr>
<td><strong>C-reactive protein:</strong> C-reactive protein (CRP) levels have been shown to correlate with risk of cardiovascular disease, including stroke, as have elevated fibrinogen levels. A recent study found that treating individuals with elevated CRP with statins reduced major cardiovascular events, despite normal lipid levels prior to treatment (16). While the role of testing CRP levels to guide stroke prevention remains a subject of debate, it may prove a useful adjunct to other tools for assessing risk.</td>
</tr>
</tbody>
</table>
Risk Factors for Recurrent Stroke

The probability of stroke recurrence is highest early after a stroke. For survivors of an initial stroke, the annual risk of a second stroke is approximately 5%, with a 5-year cumulative risk of recurrence around 25% (17,18), although it may be as high as 42% (17). Risk factors for initial stroke also increase the risk of recurrence, especially hypertension, heart disease, and diabetes mellitus (19,20). Heavy alcohol consumption is also a risk factor for recurrent stroke.

Mortality after stroke is high, with between 32% and 58% of initial survivors dead within 5 years after a first stroke, with survival varying based on age, sex, and race (1). Leonberg and Elliott (19) were able to achieve 16% reduction in stroke recurrence rate by an energetic and sustained program of control of multiple risk factors. Therefore, efforts should be made to reduce risk of recurrent stroke and mortality by controlling risk factors.

Types of Strokes

Transient Ischemic Attacks

Transient ischemic attacks (TIAs) have historically been defined as a strokelike event that completely resolves within 24 hours. The symptoms of a TIA begin abruptly and may persist for only a few seconds or minutes, followed by apparent full resolution. Despite the fact that TIAs are (by definition) associated with complete resolution of symptoms, a substantial percentage of diffusion-weighted MRIs in patients with TIA will show small acute infarcts corresponding to the individual’s symptoms (21). TIAs are not trivial events and require a thorough investigation. Those TIAs associated with MRI evidence of tissue damage carry a particularly high risk of further vascular events (22). In the first month after onset of TIAs, 4% to 8% of patients will develop a completed stroke, and the risk is 30% in the next 5 years (23,24). Generally, stroke prevention in patients with TIAs includes the use of an antiplatelet drug, such as aspirin (24).

Cerebral Thrombosis

Thrombosis of the large extracranial and intracranial vessels as the result of atherosclerotic cerebrovascular disease accounts for approximately 30% of all cases of stroke (Table 23-2) (20). Atherosclerotic plaques are particularly prominent in the large vessels of the neck and at the base of the brain. Sudden occlusion of one of these large vessels, in the absence of good collateral circulation, usually results in a large brain infarction. Risk factors for atherosclerosis include hypertension, smoking, diabetes, and hyperlipidemia. In some cases, the gradual progressive stenosis of a major blood vessel provides sufficient time for collateral circulation to develop, and complete occlusion may be asymptomatic or produce fewer effects than might otherwise be expected.

Unlike embolic strokes, which tend to have an abrupt onset, atherothrombotic strokes often begin more subtly, and affected individuals may only become aware that they have weakness or other impairment when they attempt to walk or get out of bed. The extent of the clinical deficit usually worsens over some hours or sometimes days, followed by stabilization and then gradual improvement.

Cerebral Embolism

Embolism is responsible for about 30% of all cases of stroke. Emboli may arise from thrombi within the heart or on the heart valves, from paradoxical embolism, or from an ulcerated atherosclerotic plaque within an extracranial artery (Table 23-3). Cardiogenic embolism may occur from thrombus formation at the site of a recent myocardial infarction, an area of myocardial hypokinesis, within the left atrium in patients with atrial fibrillation, or on diseased or prosthetic valves. Paradoxical embolism results from a deep venous thrombosis in the pelvis or leg that embolizes into the right side of heart, then passes through a patent foramen ovale into the left atrium, and ultimately to the cerebral circulation.

Cerebral embolism presents with an abrupt onset that is due to sudden loss of arterial perfusion to a focal area of the brain. The blood flow and anatomy of the cerebral circulation favor passage of an embolism into the middle cerebral artery territory, though any vascular territory may be affected. Many emboli are friable and may break into smaller pieces as they travel through the cerebral circulation, resulting in multiple smaller infarcts affecting several distal branches of the main vessel.

The initial clinical deficit may change rapidly. If the embolus undergoes lysis and fragmentation, the neurologic

<table>
<thead>
<tr>
<th>TABLE 23.2 Causes of Stroke</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cause</td>
</tr>
<tr>
<td>Large vessel occlusion/infarction</td>
</tr>
<tr>
<td>Embolism</td>
</tr>
<tr>
<td>Small vessel occlusion, lacunar</td>
</tr>
<tr>
<td>Intracerebral hemorrhage</td>
</tr>
<tr>
<td>Subarachnoid hemorrhage</td>
</tr>
</tbody>
</table>


<table>
<thead>
<tr>
<th>TABLE 23.3 Sources of Cerebral Embolism</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cardiac</td>
</tr>
<tr>
<td>• Atrial fibrillation, other arrhythmias</td>
</tr>
<tr>
<td>• Mural thrombus—recent MI, hypokinesis, cardiomyopathy</td>
</tr>
<tr>
<td>• Bacterial endocarditis</td>
</tr>
<tr>
<td>• Valve prosthesis</td>
</tr>
<tr>
<td>• Nonbacterial valve vegetations</td>
</tr>
<tr>
<td>• Atrial myxoma</td>
</tr>
<tr>
<td>Large artery</td>
</tr>
<tr>
<td>• Atherosclerosis of aorta and carotid arteries</td>
</tr>
<tr>
<td>Paradoxical</td>
</tr>
<tr>
<td>• Peripheral venous embolism with R-to-L cardiac shunt</td>
</tr>
</tbody>
</table>
signs may fade rapidly. While treatment with anticoagulants is appropriate for long-term secondary prevention of recurrent stroke, immediate anticoagulation with heparin has been found to increase the risk of symptomatic hemorrhagic conversion (25). For this reason, anticoagulation therapy is frequently delayed from the acute onset for up to 2 weeks in individuals with large cardioembolic infarctions.

**Lacunar Stroke**

Lacunar lesions constitute approximately 20% of all strokes and are small, circumscribed lesions, at most 1.5 cm in diameter resulting from occlusions in the deep penetrating branches of the large vessels that perfuse the subcortical structures, including internal capsule, basal ganglia, thalamus, and brainstem. Small lacunar infarcts may produce major neurologic deficits if they occur in key regions but generally cause more minor symptoms than large vessel infarcts and may in fact be asymptomatic. Lacunar strokes are highly associated with hypertension and may result from either microatheroma or lipohyalinosis.

**Cerebral Hemorrhage**

Intracerebral hemorrhage accounts for approximately 11% of all cases of stroke. The most common cause of cerebral hemorrhage is hypertension, and intracerebral hypertensive hemorrhage most commonly occurs at the site of small, deep, penetrating arteries. It is thought that hemorrhage occurs through rupture of microaneurysms (Charcot-Bouchard aneurysms) that develop in these vessels in hypertensive patients. The majority of lesions occur in the putamen or thalamus, and in about 10% of patients, the spontaneous hemorrhage occurs in the cerebellum.

Cerebral amyloid angiopathy is another important cause of cerebral hemorrhage, representing 5% to 20% of cases. It is most common after age 65, although may occasionally affect individuals as early as age 45. Hemorrhages tend to be lobar rather than deep, and there is usually evidence of hemosiderin deposition on MRI imaging indicative of prior microhemorrhages at the time of initial clinical presentation. No specific treatments are yet available, aside from avoidance of medications that may predispose to bleeding and appropriate control of hypertension if present.

The clinical onset of the hemorrhage is often dramatic, with severe headache and rapidly progressive neurologic deficits. In patients with larger hemorrhages, consciousness becomes progressively impaired leading to coma. Brain displacement from the hematoma and cerebral edema may give rise to transtentorial herniation and death, generally within the first few days poststroke. While cerebral hemorrhage is associated with a higher mortality rate than infarction, there is some evidence that the neurological deficit from a hemorrhage may recover to a greater degree than a comparable initial deficit from an infarction (26).

Intracerebral hemorrhages have been demonstrated to continue to expand after initial presentation in a substantial portion of patients. A recent trial of recombinant activated factor VII reduced the growth of the hematomas but failed to show any improvement in functional status or reduction in mortality (27).

Intracerebral hemorrhage is a well-recognized complication of anticoagulant therapy and may occur spontaneously or after minor trauma. For patients on warfarin, the risk is related to the degree of elevation of the INR. In one study, the adjusted odds ratio for sustaining an intracranial hemorrhage was 4.6 for INRs in the range of 3.5 to 3.9 (compared with individuals with INRs between 2.0 and 3.0) and increased to 8.8 for INRs above 4.0 (28).

Other causes of intracerebral hemorrhage include trauma, vasculitis, and bleeding into a tumor. Patients with a bleeding diathesis—for example, thrombocytopenia or coagulation disorders—may develop an intracranial hemorrhage.

Patients with acute cerebellar hemorrhages typically develop a sudden headache and inability to stand, along with nausea, vomiting, and vertigo. With large posterior fossa lesions, the hematoma and edema may occlude the fourth ventricle, causing acute hydrocephalus. Urgent decompression with evacuation of the hematoma can be lifesaving. Patients who survive surgical evacuation of a cerebellar hemorrhage, or who have a less severe lesion, usually make a good functional recovery, as do patients with cerebellar infarcts (29).

**Subarachnoid Hemorrhage**

In about 7% of all stroke patients, the lesion is an SAH, usually resulting from rupture of an arterial aneurysm at the base of the brain with bleeding into the subarachnoid space. Aneurysms develop from small defects in the wall of the arteries and slowly increase in size. The risk of rupture rises as the size of the aneurysm increases, and for this reason, intervention is generally advised for asymptomatic aneurysms greater than 10 mm in diameter (30). Major rupture of an aneurysm may be preceded by headache from a small bleed or by localized cranial nerve lesions caused by direct pressure by the expanding aneurysm. When rupture occurs, the clinical onset is usually dramatically abrupt. There is severe headache followed by vomiting and signs of meningeal irritation. Focal signs are usually not observed initially but may develop as a result of associated intracerebral bleeding or cerebral infarction occurring as a complication of arterial vasospasm. Coma frequently occurs, and as many as one third of patients may die acutely. Rebleeding is very common, and therefore, early surgical/invasive radiological intervention has become routine, with the objective of obliterating the aneurysm to prevent recurrent hemorrhage. Blood in the subarachnoid space may cause arterial vasospasm, leading to localized areas of cerebral infarction with associated focal neurologic deficits. Nimodipine is routinely administered after SAH to reduce the likelihood and/or severity of vasospasm. Hydrocephalus may develop immediately after SAH due to obstruction of the ventricular system from intraventricular hemorrhage, or as a later complication several weeks after the acute event as a result of arachnoiditis from blood in the CSF.
Obliteration of aneurysms may be performed surgically, by clipping the neck of the aneurysm or by the use of detachable coils placed through an angiographic approach to thrombose the aneurysm.

SAH may also result from bleeding from an arteriovenous malformation (AVM), which is a tangle of dilated vessels found on the surface of the brain or within the brain parenchyma. These lesions are congenital abnormalities and tend to bleed in childhood or young adulthood. In about half the cases, the hemorrhage is the first clinical indication of the lesion. In approximately one third of patients, the AVM presents as a seizure disorder or with chronic headaches. In most patients, the lesions eventually bleed. Most patients survive a single hemorrhagic event. The rate of rebleeding is about 6% in the first year and 2% to 3% per year thereafter. Treatment options include surgical excision of the AVM, proton beam therapy, or neurovascular embolization through obliteration.

**Stroke Syndromes**
The relatively predictable anatomy of the brain’s vascular supply, localization of particular functions to certain areas of the brain, and the predilection of stroke for certain vascular territories result in a number of commonly occurring ischemic stroke syndromes. These stroke syndromes can be recognized when they occur and assist in localization of the stroke lesion as well as in predicting functional outcome. While a myriad of stroke syndromes have been described, we discuss only a few of the most common ones in this chapter.

**Internal Carotid Artery Syndrome**
The clinical consequences of complete occlusion of an internal carotid artery vary from no observable deficit to catastrophic. In cases where there is good collateral circulation, no neurological consequences may occur with carotid occlusion. By contrast, massive cerebral infarction in the distribution of the anterior and middle cerebral arteries may present with rapid obtundation, with dense contralateral motor and sensory deficits. In some cases (particularly in younger people), severe cerebral edema may lead to transtentorial herniation and death. In such cases, decompression of the swollen brain with craniectomy may be lifesaving.

Less extensive infarctions result in partial or total lesions in the distribution of the middle cerebral artery. The anterior cerebral circulation may be preserved through flow from the opposite side via the anterior communicating artery.

**Middle Cerebral Artery Syndromes**
The internal carotid artery divides into the middle and anterior cerebral arteries. The middle cerebral artery supplies the lateral aspect of the frontal, parietal, and temporal lobes and the underlying corona radiata, extending as deep as the putamen and the posterior limb of the internal capsule. As the main stem of the middle cerebral artery passes out through the Sylvian fissure, it gives rise to a series of small branches called lenticulostriate arteries, which penetrate deeply into the subcortical portion of the brain and perfuse the basal ganglia and internal capsule. At the lateral surface of the hemisphere, the middle cerebral artery divides into upper and lower divisions, which perfuse the lateral surface of the hemisphere.

When the middle cerebral artery is occluded at its origin, a large cerebral infarction develops involving all the structures mentioned above. Because of the cerebral edema that usually accompanies such a large lesion with brain displacement, the patient frequently shows depressed consciousness, with head and eyes deviated to the side of the lesion, with contralateral hemiplegia, decreased sensation, and homonymous hemianopia. If the dominant hemisphere is involved, aphasia is usually present, which may be severe if the entire territory of the MCA is infarcted. As the patient’s mental status improves, other features become evident, namely dysphagia, contralateral hemianopia, and, in patients with nondominant hemisphere lesions, perceptual deficits and neglect. Patients who survive the acute lesion regain control of head and eye movements, and normal level of consciousness is restored. However, severe deficits involving motor, visuospatial, and language function usually persist.

Occlusion of the middle cerebral artery branches, except for the lenticulostrate, is almost always embolic in origin, and the associated infarctions are smaller and more peripherally located than those seen after occlusion of the MCA trunk. The superior division of the middle cerebral artery supplies the Rolandoic and pre-Rolandic areas, and an infarction in this territory will result in a dense sensory-motor deficit on the contralateral face, arm, and leg, with less involvement of the leg. As recovery occurs, the patient is usually able to walk with a spastic, hemiparetic gait. Little recovery occurs in motor function of the arm. If the left hemisphere is involved, there is usually severe aphasia initially with eventual improvement in comprehension, although an expressive aphasia is likely to persist. Small focal infarctions from occlusions of branches of the superior division will produce more limited deficits such as pure motor weakness of the contralateral arm and face, apraxia, or expressive aphasia.

The inferior division of the middle cerebral artery supplies the parietal and temporal lobes, and lesions on the left side result in severe involvement of language comprehension. The optic radiation is usually involved, resulting in partial or complete contralateral homonymous hemianopia. Lesions affecting the right hemisphere often result in neglect of the left side of the body. Initially, the patient may completely ignore the affected side and even assert that his left upper extremity belongs to someone else. Such severe neglect often gradually improves but may be followed by a variety of persisting impairments such as deficits in attention, constructional apraxia, dressing apraxia, perceptual deficits, and aprosodia.

Several characteristic and rather common syndromes have been described when lacunar strokes occur in the distribution of the lenticulostriate branches of the middle cerebral artery. Among the most common is a lesion in the internal capsule causing a pure motor hemiplegia. An anterior lesion in the internal capsule may cause dysarthria with hand clumsiness, and a lesion of the thalamus or adjacent internal capsule causes
a contralateral sensory loss with or without weakness. The neurologic deficits in these lesions often show early and progressive recovery with good ultimate outcome.

**Anterior Cerebral Artery Syndromes**

Branches of the anterior cerebral arteries supply the median and paramedian regions of the frontal cortex and the strip of the lateral surface of the hemisphere along its superior border. There are deep penetrating branches that supply the head of the caudate nucleus and the anterior limb of the internal capsule. Occlusions of the anterior cerebral artery are not common, but when they occur, there is contralateral hemiparesis with relative sparing of the hand and face and greater weakness of the leg. There is associated sensory loss of the leg and foot. Lesions affecting the left side may produce a transcortical motor aphasia characterized by diminution of spontaneous speech but preserved ability to repeat words. A grasp reflex is often present along with a sucking reflex and paratonic rigidity (also known as gegenhalten—muscle hypertonia presenting as an involuntary variable resistance during passive movement of a limb). Urinary incontinence is common. Large lesions of the frontal cortex often produce behavioral changes, such as lack of spontaneity, distractibility, and tendency to perseverate. Affected individuals may have diminished reasoning ability. Bilateral anterior cerebral artery infarctions may cause severe abulia (lack of initiation).

**Vertebrobasilar Syndromes**

The two vertebral arteries join at the junction of the medulla and pons to form the basilar artery. Together, the vertebral and basilar arteries supply the brainstem by paramedian and short circumferential branches and supply the cerebellum by long circumferential branches. The basilar artery terminates by bifurcating at the upper midbrain level to form the two posterior cerebral arteries. The posterior communicating arteries connect the middle to the posterior cerebral arteries, completing the circle of Willis.

Some general clinical features of lesions in the vertebrobasilar system should be noted. In contrast to lesions in the hemispheres, which are unilateral, lesions involving the pons and medulla often cross the midline and produce bilateral symptoms. When motor impairments are present, they are often bilateral, with asymmetric corticospinal signs, and they are frequently accompanied by cerebellar signs. Cranial nerve lesions are very frequent and occur ipsilateral to the main lesion. There may be dissociated sensory loss (involvement of the spinothalamic pathway with preservation of the dorsal column pathway or vice versa), dysarthria, dysphagia, disequilibrium and vertigo, and Horner’s syndrome. Of particular note is absence of cortical deficits, such as aphasia and cognitive impairments. Visual field loss and visuospatial deficits may occur if the posterior cerebral artery is involved, but not with brainstem lesions. Identification of a specific cranial nerve deficit allows precise anatomic localization of the lesion.

Lacunar infarcts are common in the vertebrobasilar distribution, arising from occlusion of small penetrating branches of the basilar artery or posterior cerebral artery. In contrast to cerebral lacunes, most brainstem lacunes are symptomatic. There is a variety of characteristic brainstem syndromes associated with lesions at various levels in the brainstem. Pontine lacunar infarcts frequently result in a pure motor hemiparesis. The reader is referred to neurologic texts for a comprehensive discussion of these lesions. Several brainstem syndromes are relatively common among patients referred for rehabilitation, and these are described in some detail.

The **lateral medullary syndrome** (Wallenberg’s syndrome) is produced by an infarction in the lateral wedge of the medulla. It may occur as an occlusion of the vertebral artery or the posterior inferior cerebellar artery. The clinical features of this syndrome, along with the corresponding anatomic structures involved, are impairment of contralateral pain and temperature (spinothalamic tract); ipsilateral Horner’s syndrome consisting of miosis, ptosis, and decreased facial sweating (ipsilateral sympathetic tract); dysphagia, dysarthria, and dysphonia (ipsilateral paralysis of the palate and vocal cords); nystagmus, vertigo, nausea, and vomiting (vestibular nucleus); ipsilateral limb ataxia (spinocerebellar fibers); and ipsilateral impaired sensation of the face (sensory nucleus of the fifth nerve). Patients with this syndrome are frequently quite disabled initially because of vertigo, disequilibrium, and ataxia, but they often make a good functional recovery.

Occlusion of the basilar artery may result in severe deficits with complete motor and sensory loss and cranial nerve signs from which patients do not recover. Patients are often comatose. **Locked-in syndrome** is an uncommon but devastating stroke syndrome involving the brainstem. The infarction in such cases affects the upper ventral pons, involving the bilateral corticospinal and corticobulbar pathways but sparing the reticular activating system and ascending sensory pathways. Patients have normal sensation and can see and hear but are unable to move or speak. Blinking and upward gaze are preserved, which provides a very limited but usable means for communication. The patient is alert and fully oriented. Some patients do not survive, and those who do are severely disabled and dependent. Slow improvement and partial recovery may occur in this group of patients, justifying appropriate levels of rehabilitation intervention.

Focal infarctions may occur in the midbrain and affect the descending corticospinal pathway, sometimes also involving the third cranial nerve nucleus (Weber’s syndrome), resulting in ipsilateral third nerve palsies and paralysis of the contralateral arm and leg.

Eye movement abnormalities are seen in a variety of brainstem stroke syndromes due to the location of the nuclei for these cranial nerves in the midbrain (third and fourth cranial nerves) and pons (sixth nerve) and their interconnections.

The posterior cerebral artery perfuses the thalamus through perforating arteries, as well as the temporal and occipital lobes with their subcortical structures, including the optic radiation. An occipital lobe infarction will cause a partial or complete contralateral hemianopsia, and when these visual deficits involve the dominant hemisphere, there may be associated...
difficulty in reading or in naming objects. When the thalamus is involved, there is contralateral hemisensory loss. A lesion involving the thalamus may cause a syndrome characterized by contralateral hemianesthesia and central pain, although only about 25% of cases of central pain in stroke are caused by lesions of the thalamus. Other lesion sites reported to be associated with central pain are the brainstem and parietal lobe projections from the thalamus. In the thalamic syndrome, patients report unremitting, unpleasant, burning pain affecting the opposite side of the body. Examination of the patient reveals contralateral impairment of all sensory modalities, often with dysesthesia. There may be involvement of adjacent structures, such as the internal capsule (hemiparesis, ataxia) or basal ganglia (choreoathetosis).

**Evaluation of the Stroke Etiology**

For many patients, a stroke represents only one facet of systemic atherosclerotic vascular disease. Such patients will often have multiple risk factors for an event, such as increased age, hypertension, diabetes, and smoking. Stroke, however, can occur in the absence of common risk factors or atherosclerosis.

The investigation undertaken to identify the cause of a stroke depends on the patient's age and presence or absence of risk factors. For example, a 79-year-old man with hypertension and a long history of smoking who sustains a lacunar infarct will not likely require assessment for a hypercoagulable state, whereas this may be an important component of the investigation of the cause of stroke in a 32-year-old woman without identifiable risk factors.

A basic evaluation of stroke cause includes a thorough physical and neurological examination, cerebral imaging (CT or preferably MRI), an electrocardiogram, noninvasive carotid studies, and an echocardiogram. In cases where there is concern regarding a possible cardiogenic embolism, a 24-hour Holter monitor is appropriate. Transesophageal echocardiography provides superior imaging of left atrium, mitral valve, and the aortic arch. In cases where the possibility of a paradoxical embolism has been raised, the echocardiogram should include a “bubble” study to evaluate for right to left shunting, and studies for deep venous thrombosis are frequently performed. CT or MR angiography may be indicated when concerns regarding large vessel occlusion, stenosis, or dissection are present. Occasionally, conventional angiography may be indicated, although this is increasingly reserved for use during intravascular interventions.

In younger patients, evaluation frequently includes tests for a hypercoagulable state and screening for vasculitis or rheumatological disorder (e.g., lupus).

**STROKE IN CHILDREN AND YOUNG ADULTS**

Stroke can occur at any age and is an important cause of disability in children and young adults. In a significant number of cases, 40% to 50%, no obvious risk factors—such as sources of cardiogenic emboli and atherosclerosis—are found. These patients should be thoroughly investigated for primary etiology of the stroke. A list of possible diagnoses is given in Table 23-4. Coagulation disorders may be inherited or acquired, but they may account for up to 20% of cases with thrombotic infarction in young adults. Deficiencies of antithrombin III, protein C, and protein S are among the most important coagulopathies, as each of these substances is part of the naturally occurring anticoagulant system. Each of these coagulopathies requires long-term treatment with warfarin.

Carotid dissection may occur with minimal or no antecedent trauma and result in stroke in the MCA territory. Vertebrobasilar artery dissection may occur after high-velocity chiropractic manipulation of the cervical spine. A variety of vasculitides may occur, some of which are part of multisystem autoimmune disease such as systemic lupus erythematosus. Another uncommon cause of stroke early in life is the inherited disorder homocystinuria, which predisposes individuals to early atherosclerosis. Lastly, stroke has been reported as an occasional complication occurring during pregnancy or in the postpartum period.

**Acute Stroke Management**

The goals of acute stroke management are (a) to limit or reverse neurologic damage through thrombolysis or neuroprotection and (b) to monitor and prevent secondary stroke complications such as elevated intracranial pressure.

Intravenous thrombolysis with recombinant tissue plasminogen activator has been known to be effective when administered to appropriate individuals within 3 hours of symptom onset since the mid-1990s (31). While a recent study found

---

**TABLE 23.4 Causes of Stroke in Children and Young Adults**

<table>
<thead>
<tr>
<th>Cause</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cerebral embolism</td>
</tr>
<tr>
<td>Trauma to extracranial arteries</td>
</tr>
<tr>
<td>Thromboembolic occlusion</td>
</tr>
<tr>
<td>Dissection</td>
</tr>
<tr>
<td>Subarachnoid hemorrhage</td>
</tr>
<tr>
<td>Aneurysm</td>
</tr>
<tr>
<td>Arteriovenous malformation</td>
</tr>
<tr>
<td>Sickle cell anemia</td>
</tr>
<tr>
<td>Vasculopathy</td>
</tr>
<tr>
<td>Moyamoya disease</td>
</tr>
<tr>
<td>Systemic lupus erythematosus</td>
</tr>
<tr>
<td>Drug induced</td>
</tr>
<tr>
<td>Vasculitis</td>
</tr>
<tr>
<td>Coagulopathy</td>
</tr>
<tr>
<td>Deficiency of antithrombin III</td>
</tr>
<tr>
<td>Deficiency of protein C</td>
</tr>
<tr>
<td>Deficiency of protein S</td>
</tr>
<tr>
<td>Homocystinuria</td>
</tr>
<tr>
<td>Oral contraceptives</td>
</tr>
<tr>
<td>Postpartum</td>
</tr>
<tr>
<td>Drug induced</td>
</tr>
</tbody>
</table>
thrombolysis useful and acceptably safe as late as 4.5 hours poststroke in a selected population, it is clear that this intervention is most effective when given as early as possible (32). This has prompted major centers to develop stroke teams that can respond rapidly within 24 hours a day with immediate clinical evaluation, urgent CT scan, and infusion as soon as possible within the therapeutic window. The ability of MRI to identify areas of mismatch between perfusion and diffusion has been studied as a means of identifying candidates for intravenous or intra-arterial thrombolysis outside of the standard time windows (33). Ultrasound has been used as a therapeutic tool in studies to enhance the thrombolytic effects of infused medications, although it is not yet been incorporated into routine clinical practice (34).

Thrombolytic agents are being administered intra-arterially in clinical trials and may prove an alternative to intravenous thrombolysis in patients who are not candidates for the latter procedure or as a salvage procedure in patients whose affected artery fails to recanalize with intravenous thrombolytic therapy (35). Clot removal devices, such as the Merci device, have been studied but are not yet in routine clinical use (36). A major challenge in the delivery of all of these therapies remains the delay in obtaining medical care among stroke victims (often due to a lack of recognition of the significance of their symptoms or by the severity of the deficits themselves), inability to identifying the time of onset, and the difficulty in providing these complex therapies in the required, time-dependent manner.

The development of an effective neuroprotective agent remains one of the major goals in acute stroke care but has thus far not been successful. Control of blood pressure, fever, and hyperglycemia have, however, been shown to improve outcome in acute stroke. Other approaches, such as brain cooling, remain under investigation.

Excitotoxicity is believed to play a role in the death of ischemic neurons in the penumbra around an infarction. Dying cells release excitatory amino acids, particularly glutamate, which activate cell membrane channels, allowing toxic levels of calcium to accumulate inside cells that are injured but not yet dead. The elevated intracellular calcium initiates an array of neurochemical changes within the neurons that generate free radicals and result in cell death. Clinical trials of glutamate receptor antagonists and free radical scavengers for acute stroke have been disappointing in humans, despite evidence of efficacy in animal models.

**Secondary Stroke Prevention**

Secondary prevention involves a multipronged effort at risk factor reduction, which may involve behavioral change, such as smoking cessation, aerobic exercise, and dietary modifications in addition to optimizing treatment of associated medical risk factors, such as hypertension and diabetes.

The third aspect of secondary prevention is the use of specific medications for stroke prevention.

Antiplatelet medications are appropriate for the majority of patients for secondary prevention of ischemic stroke. Aspirin in doses of 50 to 325 mg provides a reduction in stroke of approximately 25% (13). Gastrointestinal toxicity (bleeding, dyspepsia) is the most common side effect of aspirin, followed by allergies.

Clopidogrel is another antiplatelet agent with a different mechanism of action from aspirin. In clinical trials, its efficacy in preventing stroke is comparable to aspirin, though it is more costly. It is generally well tolerated, though its use has been associated with rare cases of thrombotic thrombocytopenic purpura. Combining clopidogrel and aspirin is generally not recommended, as the combination appears to raise the risk of hemorrhagic complications without further reduction in stroke risk (37). Ticlopidine is a related drug, with similar efficacy. It has largely fallen out of use due to concerns regarding neutropenia (24).

Another agent, dipyridamole, has some efficacy when taken alone but is generally prescribed as part of a fixed dose combination with aspirin (Aggrenox). In a large European trial, there was a 37% reduction in risk of stroke in patients prescribed both aspirin and dipyridamole (38), which compared favorably with treatment with aspirin alone. A recent study comparing dipyridamole with aspirin versus clopidogrel failed to show any difference in efficacy between these two therapies (39), and uncertainty regarding the ideal selection of antiplatelet agent persists (40).

Stroke can occur despite the use of antiplatelet agents. There is no consensus regarding the management of individuals who sustain a stroke despite preventive antiplatelet therapy, although substitution of combined aspirin plus dipyridamole therapy for aspirin monotherapy or substitution of clopidogrel for aspirin monotherapy is commonly instituted.

Warfarin use for stroke prevention is generally restricted to patients with atrial fibrillation or other known cardiac or other embolic source. For most indications, a target INR of 2 to 3 is used, although a higher range is needed for patients with certain types of mechanical heart valves.

Carotid endarterectomy reduces the risk of stroke in those patients with single or multiple TIAs and with 70% or greater stenosis of the ipsilateral internal carotid artery (41). Patients with stenosis of 50% to 70% may be considered for surgery if symptomatic, that is, if they are having TIAs or a stroke ipsilateral to the carotid lesion. The evidence for carotid endarterectomy is less compelling for patients with asymptomatic carotid stenosis, with a recommendation that it be considered for asymptomatic patients with stenoses of 60% to 99% (42).

Carotid stenting has been studied as an alternative to carotid endarterectomy, particularly in patients who may be at poor surgical risks. At present, there is no consensus regarding the relative efficacy, although some data suggest that endarterectomy provides better outcomes overall (43).

**Rehabilitation during the Acute Phase**

The care of stroke survivors is organized in a variety of different systems around the world. In many European countries, stroke units provide a combination of acute stroke management and subsequent intensive rehabilitation in a single unit.
Studies of stroke units have consistently found improved outcomes when compared with care on general medical units (44). Interestingly, much of the benefits in mortality appear to relate to prevention and/or earlier recognition of medical complications of stroke and earlier mobilization (45).

In the United States, acute stroke care is often transitioned rapidly to rehabilitative care, often within a matter of days. Despite this, it is important that rehabilitation not be considered a separate phase of care, that only begins after acute medical intervention. Rather, it is an integral part of medical management and continues longitudinally through acute care, postacute care, and community reintegration. Although diagnosis and medical treatment are the principal focus of early treatment, rehabilitation measures should be offered concurrently. Many of these can be considered preventive in nature. For example, patients who are hemiplegic, lethargic, and incontinent are at high risk for developing pressure ulcers. Deliberate strategies should be followed to prevent skin breakdowns, including protection of skin from excessive moisture, the use of heel-protecting splints, maintenance of proper position with frequent turning, and daily inspection and routine skin cleansing (46).

Many patients with acute stroke have dysphagia and are at risk for aspiration and pneumonia. In the able-bodied, aspiration usually results in vigorous coughing, but as many as 40% of patients with acute stroke experience silent aspiration. Protection against aspiration (and resulting pneumonia) includes avoiding oral feeding in patients who are not alert. Even in alert patients, the ability to swallow should be assessed carefully before oral intake of fluids or food is begun. This is done with a bedside screening assessment that can be efficiently completed by a physician or nursing staff and generally includes taking a small drink of water and observing for coughing or change in vocal quality (47). If any doubt exists about aspiration, a swallowing videofluoroscopy examination or flexible endoscopic evaluation of swallowing (FEES) is performed. During the acute phase, nasogastric tube feeding or gastrostomy tube placement may prove necessary. Patients who are lying flat in bed are at significant risk for regurgitation and aspiration, and the head of the bed should be kept elevated (48).

Impairment of bladder control is frequent following a stroke, which may initially cause a hypotonic bladder with overflow incontinence. If an indwelling catheter is used, it should be removed as soon as possible, with careful monitoring to insure that appropriate voiding resumes. For the occasional patient with persistent urinary retention after stroke, regular intermittent catheterization is preferable to an indwelling catheter (49).

Patients with hemiplegia are at high risk for development of contractures due to immobility. Spasticity, if present at this early stage, may contribute to the development of contractures through sustained posturing of the limbs. The harmful effects of immobility can be ameliorated by regular passive stretching and moving the joints through a full range of motion, preferably at least twice daily.

The risk of deep venous thrombosis is high, especially in patients with hemiplegia. Every patient should, therefore, have some form of deep vein thrombosis (DVT) prophylaxis, either subcutaneous heparin or external pneumatic compression boots or both.

Early mobilization is beneficial by reducing the risks of DVT, deconditioning, gastroesophageal regurgitation and aspiration pneumonia, contracture formation, skin breakdown, and orthostatic intolerance. Positive psychological benefits are also likely. Mobilization involves a set of physical activities that may be started passively but that quickly progress to active participation by the patient. Specific tasks include turning from side to side in bed and changing position, sitting up in bed, transferring to a wheelchair, standing, and walking. Mobilization also includes self-care activities such as self-feeding, grooming, and dressing. The timing and progression in these activities depend on the patient’s condition. These activities should begin as soon as possible (generally within 24 to 48 hours of admission (49)) unless the stroke survivor is unresponsive or medically/neurologically unstable.

### Evaluation for Rehabilitation Program

Evaluation of longer-term rehabilitation needs should occur within the first few days after stroke. Many stroke survivors will benefit from admission to an acute rehabilitation hospital or unit (also known as an Inpatient Rehabilitation Facility), and criteria for admission into such a program are listed in Table 23-5. Some individuals may be more appropriate for a subacute rehabilitation program (based in a Skilled Nursing Facility), which provides a less intense rehabilitation program with a lesser degree of medical supervision over a longer period of time. These programs are most appropriate for individuals who are unlikely to return home due to premorbid conditions (such as dementia), who are too frail to undergo an intensive rehabilitation program, or whose neurological impairments are so profound as to prevent participation in a hospital-level program. Some individuals complete a period of intensive “acute” rehabilitation, followed by “subacute” rehabilitation, before returning home. Stroke survivors with isolated disabilities such as a partial aphasia, visual loss, or monoparesis may more appropriately receive rehabilitation on an outpatient basis or through a homecare agency.

<table>
<thead>
<tr>
<th>TABLE 23.5</th>
<th>Criteria for Admission to a Comprehensive Rehabilitation Program</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stable neurologic status</td>
<td></td>
</tr>
<tr>
<td>Significant persisting neurologic deficit</td>
<td></td>
</tr>
<tr>
<td>Identified disability affecting at least two of the following: mobility, self-care activities, communication, bowel or bladder control, or swallowing</td>
<td></td>
</tr>
<tr>
<td>Sufficient cognitive function to learn</td>
<td></td>
</tr>
<tr>
<td>Sufficient communicative ability to engage with the therapists</td>
<td></td>
</tr>
<tr>
<td>Physical ability to tolerate the active program (at least 3 h/d)</td>
<td></td>
</tr>
<tr>
<td>Achievable therapeutic goals</td>
<td></td>
</tr>
</tbody>
</table>
The prognosis for return of useful hand function is poor when predictors of eventual motor recovery in the arm do not result in a substantial level of function. Movements (as are often recovered in the proximal portion of the upper limb) do not result in a substantial level of hand function for functional tasks, and gross movements. By contrast, the upper limb relies on the exquisite fine motor control of the hand for functional tasks, and gross movements (as are often recovered in the proximal portion of the upper limb) do not result in a substantial level of function.

The severity of arm weakness at onset and the timing of the return of movement in the hand are both important predictors of eventual motor recovery in the arm. If motion in one segment is not followed within a week by the appearance of motion in a second segment, the prognosis for recovery of full motion is poor. If no initial motion is noticed during the first 3 weeks, or if motion in one segment is not followed within a week by the appearance of motion in a second segment, the prognosis for recovery of full motion is poor.

The patterns of motor recovery have been described in detail by the authors of the Copenhagen Stroke Study. Summary data are shown in Figure 23–1. Ninety-five percent of stroke survivors reached their best neurologic level within 11 weeks of onset. Individuals with milder stroke recovered more quickly, and those with severe strokes reached their best neurologic level in 15 weeks on average. The course of motor recovery reaches a plateau after an early phase of progressive improvement, and only minor additional measurable improvement occurs after 6 months postonset. However, in some stroke survivors who have significant partial return of voluntary movement, recovery may continue over a longer period of time. However, recovery varies from person to person, and it is important that these data are not misinterpreted as dictating a fixed course. The observation that stroke survivors with apparently chronic stable motor deficits can experience improved motor function as the result of an intensive exercise program even years after a stroke also creates a conundrum for clinicians seeking to determine a reasonable therapeutic end point for motor rehabilitation efforts. Aphasia occurs in about one third of patients with acute stroke, although a substantial portion of affected individuals}

![FIGURE 23-1. The time course of recovery in stroke survivors shown as the cumulated rate of patients having reached their best neurologic outcome. The course of recovery is given for patients whose initial stroke severity was mild, moderate, severe, very severe, or mild. All patients are represented by m.](image)
will improve. By 6 months or more after stroke, only 12% to 18% of stroke survivors have identifiable aphasia (58,59). The time course of recovery of language function is slower and more prolonged than motor abilities, with improvements continuing for more than a year in some cases (59). The course of recovery varies depending on the nature of the language impairments. Patients with Broca’s (nonfluent) aphasia with large hemisphere lesions tend to have little recovery, whereas patients with smaller lesions (e.g., confined to the posterior frontal lobe) often show evolution into a milder form of aphasia with anomia and word-finding difficulty. Individuals with global aphasia (affecting both comprehension and verbal expression) tend to progress slowly, with comprehension often improving more than expressive ability. Communicative ability of stroke survivors who initially have global aphasia improves over a longer period of time, up to 1 year or more poststroke. Patients with global aphasia associated with large lesions may show only minor recovery, but those with smaller lesions may improve substantially. Language recovery in Wernicke’s (fluent) aphasia is variable.

About 20% of patients have a visual field defect. In general, the degree of visual improvement following stroke is modest, and if the field defect persists beyond a few weeks, late recovery is less likely.

Mechanisms of Neurologic Recovery

Neurological improvement begins immediately after the stroke and is believed to result from a number of mechanisms. Improvements seen in the first days to weeks after stroke appear to include recovery of function in portions of the ischemic penumbra and resolution of edema and associated mass effect. Overlapping with this early resolution and continuing for at least several months is the process of cerebral plasticity and functional reorganization of the cerebral cortex. Both animal and human studies with fMRI and other techniques have demonstrated alterations in cortical maps associated with this plasticity. Similarities exist between the processes of motor learning and motor recovery, and some overlap between these processes is likely. Some or all of the improvements in motor function observed with intensive training in stroke survivors with apparently stable deficits may be a result of motor learning, rather than “motor recovery” per se (60).

Plasticity is strongly influenced by use, as shown in both animal and human studies (61). Research is ongoing to determine the optimal training program to achieve maximal motor recovery. (See Chapter 81 for more details about neuroplasticity.)

Specific Stroke Impairments and Their Rehabilitation

The initial examination of a patient with an acute stroke includes a thorough neurologic examination. The neurologic findings are used by the rehabilitation team for prognostication, development of the rehabilitation plan, and selection of the appropriate setting for rehabilitation. Reassessment of the patient during rehabilitation provides a means of monitoring progress and subsequently evaluating outcome.

Cognition and Communication

The impact of stroke on cognition and communication depends on several factors, including the size and location of the lesion and the presence or absence of premorbid factors such as prior stroke(s), dementia, or other neurologic conditions. Particularly among older stroke survivors, there is a high prevalence of premorbid cognitive decline, often undiagnosed prior to the stroke.

Delirium is also common during the early phases of acute hospitalization and may result from infection, sleep disruption, fluid or electrolyte disturbances, or medication side effects. Elderly patients are at increased risk of delirium, although stroke patients can be affected at any age. This delirium is generally reversible and should be recognized so that the underlying cause can be treated, so that efforts to prognosticate are not based on faulty impressions of the magnitude of the deficits directly related to the stroke.

Stroke can also affect a person’s ability to maintain an alert and attentive state. Large hemispheric strokes may cause drowsiness, as can some lesions affecting the brainstem. Stimulants are frequently employed as treatment (such as dextroamphetamine or methylphenidate), as well as modafinil, often with partial benefit. Most patients with reduced arousal after a stroke tend to improve substantially over a period of days to weeks.

A bedside mental status assessment is an essential part of the assessment of every stroke patient. The Mini-Mental State Examination (MMSE) developed by Folstein et al. (62) is a useful bedside tool that screens a variety of mental demands quickly and gives a well-validated measure of overall mental function. It requires less than 10 minutes to administer. The NIH Stroke Scale (NIHSS) is another widely used measure that includes some elements of cognitive and language abilities. Neither the MMSE nor the NIHSS constitutes a full cognitive and linguistic assessment, and clinicians should tailor their examination to further explore and delineate the impairments identified on a basic assessment. Formal neuropsychological testing may be useful to develop a comprehensive assessment of cognitive and linguistic functions but is usually best deferred until stroke survivors have completed a portion of their rehabilitation and are approaching a stable plateau in ability.

Perceptual impairments, in which the primary sensory system is intact but the processing of the sensory information is impaired, are common among stroke survivors, particularly (but not exclusively) those with nondominant hemisphere lesions. Left hemispatial neglect may be accompanied by a right gaze preference in severe cases, but more subtle impairments may require formal testing to identify. Commonly used bedside tests for visual neglect include letter cancellation tasks or line bisection.

Hemispatial neglect may affect some forms of stimuli more than others, such as visual, tactile, or auditory. Therapies commonly include the use of visual and/or verbal cuing, including having the patient learn to cue himself or herself verbally to look to the left side. Fresnel prisms applied to eyeglasses are
sometimes used to shift images in the affected hemispherical field toward the center of the retina and hence into the field of view, although benefits are unclear, patient acceptance is limited, and use is not widespread (63).

Strokes affecting the right parietal lobe also commonly cause deficits in constructional ability, and patients may be unable to copy or draw simple figures (or omit elements on the left side). Other bedside tests of perceptual function are failure to recognize palm writing (graphesthesias), inability to identify objects in the hand (astereognosis), and extinction of simultaneous bilateral stimulation.

Anosognosia, a lack of ability to recognize the deficits resulting from the stroke, is common among individuals with right parietal lesions and may include a lack of awareness of hemiplegia, a lack of insight regarding their need for continued hospitalization, and, in severe cases, an inability to recognize their hemiplegic limb(s) as their own.

The term apraxia describes the inability of an individual to execute an intended movement when motor and sensory functions are apparently preserved. Apraxia is a disorder of motor planning, rather than of coordination (such as ataxia) or of motor strength and control (such as hemiparesis). Ideomotor apraxia may be detected when a person is unable to carry out a task on command such as “comb your hair” or “wave good-bye,” even though there is no paralysis. This type of apraxia is most common in individuals with dominant hemisphere strokes. Patients with lesions of the nondominant parietal lobe may have apraxia of dressing.

A variety of behavioral changes can occur after stroke and reflect the anatomy of the brain damage. Individuals with frontal lobe damage may display reduced attention and abulia. In more severe frontal lesions (generally bilateral), severe abulia may result in akinetic mutism. Apathy is common, manifesting as a lack of concern regarding one’s personal condition and reduced insight and concern regarding the impact on friends or family.

Deficits in attention and executive function are very common following a stroke and can occur in association with frontal, parietal, or temporal lesions. Stroke survivors may show reduced ability to maintain attention to a task, inability to switch from one task to the next, or inability to filter out unimportant stimuli in the environment with resulting distractibility. Deficits in attention are commonly misinterpreted as memory impairments, as individuals who are not attending to stimuli and information provided will be unable to recall it at a later time.

Dementia, particularly in its milder forms, may have been present before a stroke and is often unrecognized and undiagnosed by clinicians during the early recovery period. Stroke itself may cause multi-infarct dementia, and patients with multiple lesions, especially when bilateral, are more likely to show features of dementia. Numerous studies have shown that in patients with a history of stroke, those with more extensive white matter changes are at higher risk for dementia.

Poststroke depression is very common and is discussed in greater detail below. Emotional lability occurs in as many as 20% of patients poststroke and is more common in patients with right hemisphere lesions. Individuals with emotional lability often are unable to control or suppress their emotional response to common environmental stimuli and may cry or laugh very easily even though the stroke survivor recognizes that this response is inappropriate. Emotional lability (sometimes termed “emotional incontinence”) is often mistaken for depression. Education of the stroke survivor and his or her family is particularly important, and many individuals find they can tolerate these symptoms reasonably well as long as they understand their cause. Emotional lability tends to improve with time and may respond partially to treatment with tricyclic or selective serotonin reuptake inhibitors (SSRI) antidepressants (64–66).

Communication Disorders
Communication is a complex function involving reception, central processing, and expression of information (see Chapter 15). Communication occurs through the use of language and consists of a system of symbols that are combined to convey ideas, that is, letters, words, or gestures. Impairment of language is called aphasia, and its presence reflects an abnormality in the dominant hemisphere. Speech, on the other hand, is a term that refers to the motor mechanism involved in the production of spoken words, namely breathing, phonation, and articulation. Dysphonia and dysarthria are disorders of speech. The inflection and intonation of speech (prosody) are also important in communicating and may be impaired after right hemisphere stroke (aprosodia).

There are a number of aphasia classifications. In the simplest classification, aphasia is divided into two main categories: motor aphasia (sometimes called nonfluent, expressive, or anterior aphasia), characterized by nonfluent speech, and sensory aphasia (sometimes called fluent, receptive, posterior, or Wernicke’s aphasia), characterized by fluent speech (see Table 23–6) (67). Anomic aphasia is a milder form of aphasia, where difficulty recalling the correct words or names is most prominent. Simple bedside tests can allow the clinician to categorize the communication disorder. Questions typically addressed during the bedside assessment of aphasia are provided in Table 23–7. Comprehensive evaluations of patients can be performed using formal aphasia tests. Table 23–8 summarizes the more commonly used formal aphasia test instruments.

Language therapy is based on the detailed evaluation of the patient’s cognitive and linguistic capabilities and deficits. In the early stages of rehabilitation, it is important for the therapist to help the patient establish a reliable means for basic yes/no communication. The therapist then progresses to specific techniques based on the patient’s deficits. Specific techniques have been described for improving comprehension, word or phoneme retrieval, and gestures to supplement verbal communication. Recovery of functional communication may have a
more protracted course than motor recovery, and meaningful improvements may be seen in some stroke survivors 6 and 12 months after stroke and beyond. A meta-analysis found speech therapy effective for aphasia (68), although a Cochrane review (69) has found insufficient evidence to draw this conclusion. More recently, a dose-response effect has been identified, with more intensive therapy correlated with greater improvement (70). Newer methods of therapy, including Constraint-Induced Aphasia Therapy (involving intensive practice of speech tasks with which the stroke survivor has difficulty) are being explored (71).

**Visual Deficits**

Visual field deficits are common among patients with strokes in the middle cerebral artery or posterior cerebral artery distributions. Bedside testing of visual fields, including examination for extinction with simultaneous stimulation, is an important element of the clinical evaluation of stroke patients undergoing rehabilitation. More formal visual field testing may prove useful in selected patients but is seldom necessary. Patients with isolated visual field deficits, but without concomitant hemispatial neglect, generally learn to compensate effectively. A number of therapist-directed and computer-based training (e.g., Vision Restoration Therapy, Novavision, Inc.) approaches have been developed to facilitate improving the size of the functioning visual field, although their overall benefits remain debated.

Strokes in the brainstem can cause extraocular palsies and associated diplopia. Patching is often useful to prevent diplopia; patches may be alternated between eyes when appropriate to allow practice of extraocular movements and facilitate recovery of the paretic muscles.

**Dysphagia**

Dysphagia may affect 40% of patients with a unilateral hemispheric stroke (72) and is an important risk factor for pneumonitis and aspiration pneumonia. It generally has a favorable prognosis (48,73) but may be more severe and persistent in patients with brainstem lesions or with bilateral hemispheric strokes (74).

Dysphagia may affect either or both the oral preparatory pharyngeal phases of swallowing. Bedside evaluation of a patient includes observation of the function of the lips, tongue, cheeks, and jaw and elevation of the larynx during swallowing, as well as vocal quality. Having the patient swallow a small amount of water (30–90 mL) at the bedside provides considerable information about oral control, the timeliness of swallowing, presence or absence of a wet vocal quality, and the presence or absence of coughing. Patients who are drowsy should be considered unsafe for oral intake by virtue of their altered level of consciousness, and further assessment deferred until they are more awake.

If an initial screening examination suggests dysphagia, a more extensive bedside evaluation by a speech and language pathologist (or, in some facilities, an occupational therapist) is generally advisable. Some patients aspirate “silently” (without coughing), and the overall results of the bedside assessment should guide decisions regarding more extensive assessment (75). A videofluoroscopic swallowing study (VFSS) (also called a modified barium swallow) provides detailed information about the swallowing mechanism and allows for

---

**TABLE 23.6 Classification of Aphasia**

<table>
<thead>
<tr>
<th>Classification</th>
<th>Fluency</th>
<th>Comprehension</th>
<th>Repetition</th>
<th>Naming</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global</td>
<td>Poor</td>
<td>Poor</td>
<td>Poor</td>
<td>Poor</td>
</tr>
<tr>
<td>Broca</td>
<td>Poor</td>
<td>Good</td>
<td>Variable</td>
<td>Poor</td>
</tr>
<tr>
<td>Isolation</td>
<td>Poor</td>
<td>Poor</td>
<td>Good</td>
<td>Poor</td>
</tr>
<tr>
<td>Transcortical motor</td>
<td>Poor</td>
<td>Good</td>
<td>Good</td>
<td>Poor</td>
</tr>
<tr>
<td>Wernicke</td>
<td>Good</td>
<td>Poor</td>
<td>Good</td>
<td>Poor</td>
</tr>
<tr>
<td>Transcortical sensory</td>
<td>Good</td>
<td>Poor</td>
<td>Poor</td>
<td>Poor</td>
</tr>
<tr>
<td>Conduction</td>
<td>Good</td>
<td>Good</td>
<td>Poor</td>
<td>Poor</td>
</tr>
<tr>
<td>Anomic</td>
<td>Good</td>
<td>Good</td>
<td>Good</td>
<td>Poor</td>
</tr>
</tbody>
</table>

---


**TABLE 23.7 Bedside Aphasia Assessment**

<table>
<thead>
<tr>
<th>Questions</th>
<th>Clinical Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Does the patient understand?</td>
<td>Give verbal commands, ask the patient to point to objects.</td>
</tr>
<tr>
<td>Is the patient able to talk?</td>
<td>Ask the patient to name objects, describe them, and count.</td>
</tr>
<tr>
<td>Can the patient repeat?</td>
<td>Listen for spontaneous speech.</td>
</tr>
<tr>
<td>Can the patient read?</td>
<td>Ask the patient to repeat words.</td>
</tr>
<tr>
<td>Can the patient write?</td>
<td>Give commands in writing.</td>
</tr>
<tr>
<td></td>
<td>Ask the patient to copy or to write dictated words.</td>
</tr>
</tbody>
</table>
direct visualization of aspiration. Moreover, modified food consistencies, such as thickened liquids, and compensatory swallowing maneuvers, such as a chin tuck, may be tested during the course of the fluoroscopic swallowing study to determine if they result in safer swallowing.

FEES is another technique that allows more direct visualization of the swallowing process and is used in some settings as an alternative to VFSS, particularly when VFSS is unavailable (e.g., at a nursing home) orlogistically impractical (a patient who cannot be moved to the radiology suite for a VFSS for medical reasons). While it does not allow direct visualization of aspiration, FEES does provide considerable information that complements the bedside examination and avoids the radiation exposure present with VFSS.

Both compensatory and recovery strategies are useful for dysphagia. Modified food consistencies, such as pureed foods or thickened liquids, are widely used, as are physical maneuvers, such as a controlling food bolus size and pacing, or using a "chin tuck" or double swallow. It appears likely that practicing swallowing over time improves performance and may enhance the process of spontaneous recovery.

For patients who cannot swallow safely, a nasogastric, gastrostomy, or jejunostomy tube is indicated for enteral feedings. None of these tubes are entirely effective in preventing aspiration pneumonia, in part because oral secretions may be the source of aspiration pneumonia, rather than feedings. Nasogastric and gastrostomy tubes both permit aspiration through the mechanism of gastroesophageal reflux as well, especially when patients are placed flat in bed. The selection of a feeding tube depends on several factors, including patient preference, anticipated duration of the use of a feeding tube, and procedure-related risks. Nasogastric tubes often prove problematic for longer term feeding due to frequent dislodgement, tube blockages, and patient discomfort, and therefore, gastrostomy tubes are often favored if the duration of the feeding tube use is expected to be lengthy. While most individuals with hemispheric stroke recover good swallowing function, those with brainstem lesions or bilateral hemisphere lesions may progress more slowly, and some require indefinite use of a feeding tube. Some individuals recover the ability to ingest sufficient food before they are able to drink enough liquids to maintain adequate hydration—in these cases, a feeding tube may continue to be required for fluid supplementation.

A surprising number of patients admitted for stroke rehabilitation are malnourished—as high as 22% in one report (76). Elderly patients may have marginal nutritional status prior to their stroke, and if not closely monitored during postacute rehabilitation, their fluid and nutritional status may be further compromised because of dysphagia, reliance on others for oral or tube feedings, lack of interest in food, depression, and problems with communication. The risk of malnutrition and dehydration is very real in even the most alert appearing patient and intake of fluid, protein, total calories, and body weight should be monitored closely in all patients. Oral nutritional supplements may have to be prescribed, and if a patient continues to have inadequate intake, enteral tube feeding may be necessary.

### Motor Impairment

**Strength, power (the speed with which force can be generated), motor control and coordination, muscle tone, and balance may all be affected by stroke.** The most widely used clinical scale to assess strength is the Medical Research Council's (MRC) 6-point scale, 0 to 5, in which 0 represents complete paralysis, 3 is the ability to fully move the limb against gravity, and 5 indicates normal strength (77). This scale is useful in grading muscle strength in patients with lower motor neuron lesions or myopathies but is problematic for assessing stroke survivors. Stroke survivors will often be able to activate a muscle group to varying degrees depending on the position of the limb and coactivation of synergistic muscles. Thus, a stroke survivor may be able to generate considerable force when grasping in association with wrist and elbow flexion but have much less ability to selectively flex the fingers in isolation. Also, the failure of this scale to indicate the degree of motor control may lead to misinterpretation, as in the example of a patient who can generate considerable force, but who has limited motor control, and whose MRC score is recorded as normal or near normal. Despite these shortcomings, the MRC scale is widely used by clinicians, often supplemented by qualifying descriptions (e.g., “elbow flexion is 4/5 on MRC scale in a flexion
synergy pattern with moderately reduced motor control”"). Use of a dynamometer (78) to measure force generation requires more time and similarly fails to assess motor control and thus has a limited role in the measurement of paresis in stroke survivors.

Brunnstrom (79) adopted a different approach for assessment of motor function in hemiplegic patients in which movement patterns are evaluated and motor function is rated according to stages of motor recovery (Table 23-9) (67). While this rating can be performed very quickly, the scale defines recovery only in broad categories. Moreover, not all hemiparetic stroke survivors progress through these “stages” sequentially, with some survivors skipping stages during their recovery. For these reasons, this scale has not achieved widespread clinical or research usage.

Despite the limitations of assessing and interpreting muscle strength in patients with upper motor neuron lesions, strength does correlate with performance on functional tasks (80). Fugl-Meyer et al. designed a more detailed and comprehensive motor scale in which 50 different movements and abilities were rated (81). The test evaluates strength, reflexes, and coordination, and a composite score is derived on a scale of 0 to 100. The Fugl-Meyer Scale is reliable, and repeat scores reflect motor recovery over time. It is widely used in the research setting but has not been widely adopted by clinicians because it is time-consuming to complete. A variety of other motor scales have been developed to provide a more comprehensive picture of motor function (rather than merely impairment), which are primarily utilized in a research setting, such as the Wolf Motor Function Test (82), Motor Activity Log (83), the Arm Research Action Test (84), and others.

Spasticity or Muscle tone refers to the resistance felt when the examiner passively stretches a muscle by moving a limb and generally increases if the stretching is performed more quickly (velocity dependent). Abnormal increases in muscle tone are quite commonly seen in hemiparetic limbs, and patients with more severe spasticity tend to have less well-preserved motor control. Factors such as posture and limb position can affect spasticity and must be considered when measuring muscle tone. Scales for measuring spasticity remain of limited reliability, with the most widely used being the Ashworth Scale (see Chapter 50, Spasticity) (85). Most spasticity can be managed conservatively with stretching and positioning. In cases where medical therapy is needed, treatment with oral antispasticity medications, such as baclofen or tizanidine, is often ineffective and may cause sedation or cognitive impairments. The use of injection therapy (principally botulinum toxin, but previously more commonly phenol or alcohol injection) is generally effective in reducing spasticity, although it does not ameliorate the underlying impairments in motor control.

**Therapy for Motor Weakness**

**Early Phase and Supportive Care**

In the early poststroke phase, the hemiparetic limb(s) may be completely paralyzed and are at high risk for the development of contractures or nerve pressure palsies. Therapy during this early phase should consist of proper positioning of the patient in bed and support of the arm in a wheelchair trough when sitting. Traction on the arm should be avoided when the patient is moved or transferred to a wheelchair. All joints of the affected limbs should be passively moved through a full range of motion at least once daily to prevent contractures.

If the limb(s) become quite spastic, frequent slow stretching can help to reduce tone. Spasticity usually dominates in the flexors of the upper limb and may hold the wrist and fingers in a constant position of excessive flexion. A static wrist-hand orthosis is often helpful in maintaining these joints in a functional position.

**MOTOR RECOVERY**

Motor recovery may become evident within hours to days after a stroke. A variety of approaches have been advocated to facilitate and enhance motor recovery, including traditional approaches, such as the neurodevelopmental technique advocated by Bobath (86), functionally oriented exercise training (such as practicing transfers and early ambulation), and a recent focus on repetitive task-oriented practice. Neurodevelopmental (NDT) exercises are intended to normalize muscle tone and prevent excessive spasticity using special reflex-inhibiting postures.

---

**TABLE 23.9 Brunnstrom Stages of Motor Recovery**

<table>
<thead>
<tr>
<th>Stage</th>
<th>Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage 1</td>
<td>No activation of the limb.</td>
</tr>
<tr>
<td>Stage 2</td>
<td>Spasticity appears, and weak basic flexor and extensor synergies are present.</td>
</tr>
<tr>
<td>Stage 3</td>
<td>Spasticity is prominent; the patient voluntarily moves the limb, but muscle activation is all within the synergy patterns.</td>
</tr>
<tr>
<td>Stage 4</td>
<td>The patient begins to activate muscles selectively outside the flexor and extensor synergies.</td>
</tr>
<tr>
<td>Stage 5</td>
<td>Spasticity decreases; most muscle activation is selective and independent from the limb synergies.</td>
</tr>
<tr>
<td>Stage 6</td>
<td>Isolated movements are performed in a smooth, phasic, well-coordinated manner.</td>
</tr>
</tbody>
</table>

Other traditional techniques, such as the system developed by Rood, involve superficial cutaneous stimulation using stroking, brushing, tapping, and icing or muscle stimulation with vibration, to evoke voluntary muscle activation. Brunnstrom emphasized the synergistic patterns of movement that develop during recovery from hemiplegia. She encouraged the development of flexor and extensor synergies during early recovery, hoping that synergistic activation of muscle would, with training, transition into voluntary activation. Proprioceptive neuromuscular facilitation (PNF) was developed by Kabat and colleagues and relies on quick stretching and manual resistance of muscle activation of the limbs in functional directions, which often are spiral and diagonal. Generally, there is little empirical evidence supporting the use of these techniques (NDT, Brunnstrom, Rood, PNF) which are increasingly being replaced with functional task-oriented approaches.

Newer techniques, such as Constraint-Induced Movement Training (CIMT) (87) and robot-aided exercise (88), have been increasingly incorporated into clinical protocols. CIMT utilizes an intensive short course (generally 2 weeks) of upper limb training with repetitive task-oriented practice and behavioral shaping techniques to enhance not only motor control of the upper limb but also incorporation into behavioral repertoires. The mechanism underlying this improved function remains uncertain. Taub (who developed this technique) initially hypothesized that this therapy was able to overcome a “learned nonuse” of the weak limb, in which the potential for improved motor function is present but not incorporated into actual usage (89–91). The identification of changes in cortical mapping using transcranial magnetic stimulation (61) and in activation of the primary and supplementary motor cortex using fMRI (92) suggests that cortical reorganization of the motor systems is an important underlying mechanism for the effects of this treatment. Other task-oriented practice techniques have also shown efficacy (93), and research on identifying the optimal exercise regimen is ongoing. Logistic and cost considerations have limited the availability of Constraint-Induced Movement Therapy, as has the suitability of this therapy for only a modest range of deficits. Modifications of Constraint-Induced Movement Therapy are being studied in an effort to provide a more logistically feasible treatment program (94).

Robot-aided therapeutic exercise has several potential advantages over conventional therapies as it provides consistently delivered therapy and is particularly suitable for highly-repetitive exercises. Most robotic systems provide assistance to users attempting movements and thus can be used by individuals with more severe paresis after stroke. Upper limb devices, such as the InMotion2 (Interactive Motion Technologies, Cambridge, MA), have been shown to provide benefit in both the acute (95,96) and chronic phases of stroke (88). Lower limb robotic systems have been developed for gait training, as discussed below (“Therapy for Mobility”). While these devices hold the promise of providing more efficient therapy, the current generation requires skilled assistance (from a therapist or aide) in setting the patient up and in supervision of treatment. Robotic therapy has not been proven superior to comparably dosed hands-on therapy. Robotic therapy is discussed in greater detail in Chapter 83.

Electrical stimulation of the limb (sometimes termed therapeutic electrical stimulation or functional neuromuscular stimulation) has long been used in some form since the 1960s as a treatment for hemiparesis, with some recent evidence accumulating that it may be beneficial in restoring motor use to some extent (97). Please see Chapter 71 for a more extensive discussion of this therapy.

Biofeedback using electromyogram (EMG) signals has been used in an attempt to improve motor control poststroke. Results of trials have been mixed, some showing benefit but others no better results than control therapy. While one review of clinical trials of biofeedback did find that biofeedback was somewhat effective (98), the magnitude of the benefits is small, and clinical use in the United States is relatively rare. Conventional EMG biofeedback involves recording surface EMG from the test muscle and using auditory or visual display of the EMG signal as feedback to the patient on the activity status of the muscle. The EMG signal supplements conventional reeducation given by the therapist. A combination therapy in which therapeutic electrical stimulation is triggered by the EMG signal of the target muscle is available (99). In another approach, robotic devices that relay on surface EMG signals to control or trigger the device are being studied (100).

### THERAPY FOR MOBILITY

In the early phase after stroke, some individuals will not have sufficient trunk control and hemiparetic leg strength to maintain upright posture for walking. These patients should receive initial therapy to develop gross trunk control and training in pre-gait activities such as posture, balance, and weight transfer to the hemiparetic leg. As recovery progresses, stroke survivors usually develop better gross motor skills and trunk balance and greater strength in the leg. Despite the presence of spasticity and the inability to selectively activate individual muscles, most stroke survivors will walk, although many will require an ankle-foot orthosis and cane and will walk more slowly than previously. Individuals with less severe paresis may have an asymmetric gait but not require a brace for their affected leg. Many stroke survivors will require a cane for walking, which may be a standard cane or, less commonly, four pronged, depending on their balance impairments. While hemiparesis often makes it impossible to use a walker effectively, individuals with milder paresis or other stroke-related balance impairments such as ataxia may find a walker helpful.

There have been reports that hemiplegic patients benefit from intensive gait training when therapy consists of walking on a treadmill with body weight partially supported (PBWSTT) with a harness (101,102). The harness substitutes for poor trunk control, and the motor-driven treadmill forces locomotion. During early training, the patient is assisted by
two or three therapists in controlling the trunk, pelvis, and weak leg. Treadmill training with body weight support appears superior to conventional therapy with improved gait velocity resulting (103). Progressive increase in treadmill speed as tolerated accentuates this benefit (103).

Robotic therapy has been proposed as an alternative to PWBSTT training, in part because of the need for two or even three staff members to assist the patient during a treatment session. Robotic devices, such as the Lokomat (Hocoma, Inc.) and the ReoAmbulator (Motorika, Inc.), are now commercially available for this purpose. While the feasibility of this approach has been demonstrated (104), there is evidence that therapist-assisted locomotor training with a treadmill is more effective (105). Continued enhancement of the robotic devices and treatment algorithms may result in increased efficacy of robotic therapy in the future.

**Sensory Impairment and Central Pain**

Sensory impairment may occur with or without accompanying motor weakness following a stroke, although loss of proprioception generally results in reduced motor performance even in the absence of weakness per se. Lesions of the thalamus may cause severe contralesional sensory loss and result in a central pain syndrome in some individuals. Central pain can occur, particularly with strokes involving the spino-thalamo-cortical pathway. The pain usually begins a few weeks after stroke onset and becomes intractable to conventional medications, including opiate analgesics. The pain is classically described as burning in character, although other types of pain can predominate.

With lesions of the cortex, primary sensory modalities may be preserved, but the perception of these sensations may be both qualitatively altered and quantitatively reduced. Perceptual impairment, as often seen in parietal lobe lesions, frequently is manifested as inability to perceive stimulation of the affected side during simultaneous bilateral stimulation (extinction), reduced two-point discrimination, reduced object recognition (stereognosis), and impaired recognition of digits drawn in the palm (graphesthesia).

Regrettably, few interventions have been proposed to specifically treat sensory loss after stroke, although a variety of pharmacological treatments have been studied for treatment of central pain poststroke. Tricyclic antidepressants have been best studied for this condition, although some patients do not have an adequate response, and anticholinergic side effects can be problematic (especially in the elderly). Anticonvulsants, such as gabapentin or lamotrigine, show some efficacy as well (106,107).

**OUTCOME AND PROGNOSTICATION IN STROKE REHABILITATION**

Patients, their families, and members of the health care team all benefit from an accurate prognosis to help inform decision making. The prospects for survival, the degree of recovery that may be expected, and the extent of possible residual disability following rehabilitation are all important elements of prognosis.

The challenge in developing accurate prognostic models is not surprising, given the myriad of complex functions performed by the brain, the variability of stroke lesion location and size, the limitations of our assessment techniques, and the impact of many baseline factors on stroke recovery, such as age, comorbid conditions, personality and coping abilities, and social factors.

**Predicting Survival**

Early death following a stroke is usually related to the underlying pathology, the patient’s age, and to the severity of the lesion. For patients with a first cerebral infarction aged 45 to 64, survival is 88% to 92% at 30 days; for those aged 65, the survival rate is similar. By contrast, for patients with hemorrhagic stroke (including SAH), the survival is reported to be only 62% to 63% for those aged 45 to 64, falling to 55% for those aged 65 and older (1). Coma following a stroke onset indicates a poor prognosis, presumably because coma occurs frequently in cerebral hemorrhage, and when it occurs in relation to cerebral infarction, it reflects a large lesion with cerebral edema.

**Predicting Disability and Functional Status**

The key outcomes from a rehabilitation perspective are restoration of function and community reintegration. The central purpose of the rehabilitation program is to lessen ultimate disability; therefore, considerable attention has been directed at the identification of factors that will predict the functional outcome of the patient, especially with respect to walking and activities of daily living (ADLs).

The prognosis for recovery from lacunar lesions is usually excellent, although significant persisting deficits may occur when the lesion is strategically located. With large vessel infarctions, due either to thrombosis or embolism, prognosis is related to the volume of the lesion. Outcome is poorest when the lesion involves more than 10% of intracranial volume (108).

While many stroke survivors are initially unable to walk independently, 54% to 80% achieve this milestone within the first 3 months poststroke (although frequently requiring the assistance of a cane or AFO) (59,109). Data from the Framingham cohort reported by Gresham et al. (110) indicate that long-term survivors of stroke show good recovery of functional mobility, with 80% being independent in mobility.

Most patients with significant neurologic impairment who survive a stroke are initially dependent in basic ADLs, that is, bathing, dressing, feeding, toileting, grooming, transfers, and ambulation. The capacity of individuals to perform these activities is usually scored on ADL scales, such as the Functional Independent Measure (FIM®) (111). Almost all patients show improved function in ADLs as recovery occurs. In most reports, between 47% and 76% of survivors achieve
partial or total independence in ADLs (112–114). Basic ADL scales, such as the FIM, while sensitive to changes during the early phase of rehabilitation, have limited sensitivity at the upper end of the functional range and ceiling effects. Instrumental ADLs, such as the ability to prepare meals or perform housework, are not measured by the FIM yet represent an important component of reintegration into the community for stroke survivors.

Most authors who have attempted to determine which factors predict ultimate ADL functional outcome have used multivariate analysis. The single most useful predictor of functional outcome is the initial ADL assessment (most commonly, the FIM score). Other important variables include age and sitting balance. A list of important variables predicting outcome is given in Table 23-10. Not all of these factors were shown in every study to statistically predict outcome status.

The effect of age on outcome may partly be related to more frequent comorbid diseases and functional impairments, rather than merely a result of age per se.

Intuitively, it would seem reasonable to assume that patients with more severe neurologic deficits would have worse functional outcomes, but this is not necessarily the case when isolated neurologic impairments are considered. For example, analyses of predictive variables have failed to show that patients with sensory deficits have a poorer ultimate outcome than those with preserved sensation (115). When considering functional independence as the outcome measure of interest, the severity of the initial ADL score is generally one of the more reliable predictors of ultimate ADL function. On average, patients admitted for rehabilitation with lower ADL scores do not have as good a functional outcome as patients who initially had higher admission ADL scores. Most patients with FIM scores of greater than 80 are discharged home, although many other factors influence discharge disposition (116).

It is important that rehabilitation begins as early as possible. The Post-Stroke Rehabilitation Outcomes Project demonstrated that patients have better outcomes if they are admitted into a rehabilitation program early rather than later, regardless of severity of the stroke (117). In addition to enhancing recovery, early rehabilitation and mobilization appear to reduce the likelihood of secondary complications, such as contractures and deconditioning, and help patient motivation.

**Social Variables**

Stroke survivors with severe disabilities who need maximum physical assistance in ADLs and who have bowel or bladder incontinence are the most likely to require long-term institutional care (112). While functional status is the most important issue in determining discharge destination, psychosocial factors, especially pre-stroke family interaction (118) and the presence of an able spouse, also influence whether the patient returns home. A supportive family whose members are willing and able to provide significant physical care may be able to manage a severely disabled patient at home. By contrast, a patient with much less disability but no family support may require institutional care if not fully independent.

When a clinician is confronted with the challenge of evaluating an individual patient, guidelines for predicting functional outcome are useful but are not precise because multiple variables interact. A patient who might be judged as having a good prognosis for functional outcome may do poorly because of a negative psychosocial factor. The best estimate of prognosis can be made only after a thorough and comprehensive evaluation of the patient’s medical, neurologic, functional, and psychosocial statuses.

### MEDICAL MANAGEMENT DURING REHABILITATION

There is a high incidence of comorbid medical disorders among patients recovering from stroke, reflecting the age of the patient population and the fact that cerebrovascular disease is often part of a generalized disease process. If severe, or if poorly managed, these disorders may interfere with the patient’s participation in the rehabilitation program and may adversely affect outcome. Similarly, medical complications frequently occur during the postacute phase of rehabilitation, affecting as many as 60% of patients and as many as 94% of patients with severe lesions (119). Common medical and neurologic complications are listed in Table 23-11 (110,112–114). Some of the important and more frequent disorders are discussed briefly.

**Cardiac Disease**

In a large majority of patients, a stroke is an acute event in the course of a systemic disease, for example, atherosclerosis, hypertensive vascular disease, or cardiac disease causing embolic stroke. As many as 75% of stroke patients may show evidence of coexisting cardiovascular disease, including hypertension (estimates range from 50% to 84%) (120) and coronary artery disease (as many as 65%) (120). Another group of heart diseases cause a stroke through cardiogenic cerebral
embolism. These diseases include atrial fibrillation and other arrhythmias from multiple causes, valvular disease, cardiomyopathy, endocarditis, or recent myocardial infarction.

Concomitant heart disease has a negative impact on short-term and long-term survival and probably on functional outcome of stroke patients (121,122). Acute exacerbations of heart disease occur frequently during postacute stroke rehabilitation (121). Common problems include angina, uncontrolled hypertension, hypotension, myocardial infarction, congestive heart failure, atrial fibrillation, and ventricular arrhythmias. Development of one of these complications may have minimal or no impact on the patient’s progress or outcome if the problem is promptly diagnosed and appropriately treated. However, these complications often do impact the patient’s capacity to participate fully in the therapeutic program. Congestive heart failure and angina decrease exercise tolerance and reduce capacity to roll over in bed, transfer, and walk.

**Urinary Tract**

Urinary function can be affected in several ways by stroke, including urinary infection, urinary retention, and urge incontinence. At the time of the acute stroke, some patients develop urinary retention due to altered mental status or direct effects of the stroke on the neurological control of micturition. Premorbid voiding dysfunction, often in the form of benign prostatic hypertrophy in men and stress incontinence in women, is a common comorbid condition in stroke patients and often becomes a more significant problem after a stroke due to the patient’s decreased mobility or communication abilities. Catheterization, often with an indwelling catheter, is common during the acute management of the stroke survivor. While this alleviates acute urinary retention, it may lead to urinary infection and interferes with reestablishment of a normal voiding pattern. Indwelling catheters should be removed as quickly as possible, and intermittent catheterization substituted for individuals unable to void spontaneously. Noninvasive measurement of bladder volume using ultrasound is often helpful when managing individuals with impaired bladder function after stroke.

While urinary retention generally resolves quickly in affected stroke survivors, many stroke survivors develop urinary urgency and/or incontinence. Disinhibition of the bladder detrusor is common and results in urinary frequency and urgency that may result in incontinence in some individuals. Anticholinergic medications such as oxybutynin chloride (Ditropan) or tolterodine (Detrol) are useful to inhibit bladder contraction, although these may cause anticholinergic side effects, such as dry mouth, or confusion.

**MUSCULOSKELETAL PAIN AND COMPLEX REGIONAL PAIN SYNDROME**

Shoulder and arm pain is common following a stroke in survivors. It tends to develop early, several weeks to 6 months postonset, and some studies have found as many as 72% of individuals affected, especially those with more severe hemiplegia (123). More recent studies suggest a lower incidence, with one study finding 37% of patients on a stroke rehabilitation unit reporting pain (124). This apparent reduction in incidence may reflect more intensive and earlier rehabilitation programs incorporating range-of-motion exercises.

Although some patients may have preexisting shoulder problems, such as rotator cuff tendinitis, the pain in most patients with hemiplegia results from varying combinations of glenohumeral subluxation, spasticity, and contracture. Of these, subluxation is the commonly implicated cause. The role of subluxation in generating shoulder pain has been debated, and subluxation may occur without pain symptoms in many individuals. Subluxation is evident on physical examination, and imaging studies are not generally useful unless other diagnoses are being considered. Treatment of a subluxed shoulder in poststroke patients can limit the subsequent development of shoulder pain (125). Some cases of shoulder pain appear related to spasticity of shoulder girdle muscles (e.g., the subscapularis), although it is unclear how often this is the actual cause of pain versus an unrelated finding.

Preventative measures to reduce the likelihood of developing shoulder pain should be instituted early poststroke in patients with flaccid paralysis of the upper limb and clinically evident subluxation or with any symptoms of shoulder pain. Proper positioning includes supporting the arm in an

---

**TABLE 23.11 Medical Complications During Postacute Stroke Rehabilitation**

<table>
<thead>
<tr>
<th>Complication</th>
<th>Frequency (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medical</td>
<td></td>
</tr>
<tr>
<td>Pulmonary aspiration, pneumonia</td>
<td>40</td>
</tr>
<tr>
<td>Urinary tract infection</td>
<td>40</td>
</tr>
<tr>
<td>Depression</td>
<td>30</td>
</tr>
<tr>
<td>Musculoskeletal pain, RSD</td>
<td>30</td>
</tr>
<tr>
<td>Falls</td>
<td>25</td>
</tr>
<tr>
<td>Malnutrition</td>
<td>16</td>
</tr>
<tr>
<td>Venous thromboembolism</td>
<td>6</td>
</tr>
<tr>
<td>Pressure ulcer</td>
<td>3</td>
</tr>
<tr>
<td>Neurologic</td>
<td></td>
</tr>
<tr>
<td>Toxic or metabolic encephalopathy</td>
<td>10</td>
</tr>
<tr>
<td>Stroke progression</td>
<td>5</td>
</tr>
<tr>
<td>Seizure</td>
<td>4</td>
</tr>
</tbody>
</table>

arm trough or lapboard when seated and avoiding traction on the arm when transferring. Selective use of supportive slings during ambulation or standing is advisable, although their use negatively impacts balance. Excessive use may contribute to undesirable reduction in opportunities to use the upper limb and potentially slow functional recovery and cause loss of range of motion. In cases where spasticity appears responsible for pain, the use of botulinum toxin or phenol injections may be appropriate (126).

The incidence of upper limb complex regional pain syndrome (CRPS, also known as reflex sympathetic dystrophy or shoulder-hand syndrome) remains controversial. Some studies have reported a high prevalence of 1 in 8 based on clinical criteria and up to 1 in 4 using bone scan criteria (127,128), but these figures appear inflated by inclusion of other causes of upper limb pain in hemiplegic stroke survivors, including patients with pain related to shoulder subluxation and some with central pain syndromes. Preventative measures, including frequent passive range of motion, desensitization with massage, and active incorporation of the paretic upper limb in the therapy program, are important and may have contributed to the apparent reduction in the frequency of this syndrome. Please see Chapter 49 for a discussion of treatments such as oral steroids in established cases of CRPS.

VENOUS THROMBOEMBOLISM

All patients with significant immobility related to stroke should receive DVT prophylaxis, which should consist of either low-dose subcutaneous heparin or low molecular weight heparin, with the latter somewhat more effective in high-risk individuals. Treatment with low-dose low molecular weight heparin reduced the risk of DVT (odds ratio 0.34) without an increase in the risk of major intracranial or extracranial hemorrhage (129). In patients who cannot safely receive these medications, external pneumatic compression devices are an effective alternative, particularly in the acute hospital phase. Although many rehabilitation units attempt to use pneumatic compression in addition to or instead of pharmacological treatment, their use in this setting is problematic due to the patients’ involvement in out of bed activities. The optimal duration of prophylaxis remains uncertain, although most practitioners discontinue its use once patients are walking significant distances on a frequent basis or at the time of discharge to the community.

All patients with suspected DVT should undergo prompt investigation by venous duplex ultrasound imaging. Routine screening examinations are not generally used in this population at present. Symptoms may be subtle or absent, and therefore, physical examination is not reliable enough to confirm or exclude this diagnosis (130).

Patients with newly diagnosed DVT should immediately receive full-dose anticoagulation with low molecular weight heparin (or occasionally intravenous conventional heparin) until a therapeutic INR (between 2 and 3) can be achieved with the use of warfarin for at least 2 days. Patients with acute DVT are commonly placed on bed rest for 24 hours, although the value and optimal duration of this practice remain debated, and some guidelines recommend immediate ambulation when feasible (131). In patients who cannot safely receive anticoagulation (most commonly those with a recent intracranial hemorrhage), an inferior vena cava filter should be inserted to prevent pulmonary embolism.

DEPRESSION

Depression is common following stroke and, depending on diagnostic criteria, has been reported in as many as 50% of patients (132,133). Some have reported a relationship between left frontal or bifrontal lesions and major depression (134), though this may be an artifact of the language function needed to complete the commonly used depression assessment scales that interferes with accurate measurement of depression in aphasic patients. Multiple theories exist for the high incidence of depression poststroke, which appears more common than would be expected merely as a result of the psychosocial stresses accompanying acquired disability. One hypothesis is that stroke may result in brain catecholamine depletion through lesion-induced damage to the frontal noradrenergic, dopaminergic, and serotonergic projections (135).

The diagnosis of poststroke depression may be complicated by the presence of normal sadness associated with the loss of independence, emotional lability that may result from stroke, and the high prevalence of symptoms often associated with depression (e.g., sleep disturbance, fatigue, altered appetite) that are known to result from stroke itself. Persistently depressed mood, loss of interest in socialization, and limited participation in the rehabilitation program are often more reliable indicators in this population (136).

Persisting depression correlates with delayed recovery and poorer ultimate outcome. Active treatment should be considered for all patients with significant clinical depression. Patients with poststroke depression generally respond well to standard antidepressant medications, with SSRIs commonly prescribed. Some practitioners will also treat with a stimulant medication (typically methylphenidate) in an attempt to “boost” the response and achieve more rapid results. A recent study found preventative treatment with an SSRI effective in reducing the frequency of poststroke depression, and some physicians are adopting this treatment approach for high-risk patients (137).

SEXUALITY

Stroke survivors commonly experience sexual dysfunction (138), with reductions both in libido and sexual performance predominating, although occasionally stroke can cause hyper-sexuality. The precise mechanisms for sexual dysfunction after stroke have not been determined, but a combination of psychosocial and medical issues appears to be involved. Stroke is most common in people who are older, have hypertension,
and/or diabetics—all risk factors for sexual dysfunction even without a stroke. Moreover, sexual dysfunction is common in the adult population at large, with as many as 40% to 45% of adult women and 20% to 30% of adult men having at least one manifestation (139). Thus, some of the sexual dysfunction observed in this population may not directly result from the stroke.

Regardless of the specific cause(s), the prevalence of sexual problems after stroke is estimated to be between 57% and 75% (140,141). Sexual dysfunction is also very common among the sexual partners of stroke survivors and is likely due to a combination of psychosocial stressors and the reduced sexual availability of their usual partner due to the sequelae of stroke.

Many stroke survivors and their sexual partners harbor fears (sometimes not expressed) that resumption of sexual activity might precipitate another stroke. While limited data are available to directly address this concern, most practitioners agree that resuming sexual activity is safe and appropriate after discharge from the hospital in the vast majority of cases. This is largely based on extrapolation from studies that have found a low risk of myocardial infarction from resumption of sexual activity (142).

The natural history of sexual dysfunction after stroke appears favorable in most cases. In one study of men with stroke, 82% of those who experienced erectile dysfunction after stroke improved spontaneously within a few months. The degree of dependence in ADLs is a strong predictor of decreased sexual frequency among stroke survivors (143,144).

It is important to recognize that some sexual dysfunction after stroke may be iatrogenic and a result of medications used to treat stroke-related conditions. In particular, SSRI antidepressants, such as fluoxetine and paroxetine, are well-known to cause sexual dysfunction in as many as 30% to 70% of those receiving them without concomitant medical issues such as stroke. Similarly, antihypertensives, including β-blockers in particular, are a common cause of sexual dysfunction. Angiotensin-converting enzyme inhibitors and angiotensin-receptor blockers are generally better tolerated from this perspective. Lastly, anticoagulant medications are well-established causes of reduced libido and sexual dysfunction (145).

Phosphodiesterase-5 inhibitors, such as sildenafil (Viagra) and tadalafil (Cialis), are useful for selected men with poststroke erectile dysfunction. Caution should be exercised with regard to potential medication interactions, including α-blockers and nitrates.

Generally speaking, physicians and other health care professionals caring for stroke survivors should inform stroke survivors of their willingness and availability to discuss sexuality after stroke and provide further information to patients and their sexual partners when desired. Embarrassment on the part of both the patient and the physician is often a barrier to discussing this subject, as are implicit assumptions about sexuality (e.g., a physician assuming that an older widowed individual is not sexually active). Reassurances regarding the safety of resuming sexual activity are a particularly critical aspect of this counseling.

Ongoing Care

The observations that continued improvement in motor function is possible in stroke survivors even after years of stable chronic deficits have created both opportunity and challenge for rehabilitation providers. We do not yet know what the ideal rehabilitation program consists of, particularly with regard to intensity and duration of therapy. Uncertainty regarding the maximal achievable outcomes for stroke survivors both preserves hope for our patients and creates complexity for practitioners.

Furthermore, stroke survivor’s other rehabilitation needs may extend for many years after the initial event. Late issues may include depression, spasticity, contracture, osteoporosis, weight gain, and deconditioning due to reduced activity. The gradual progression of some of these issues, such as gradual loss of range of motion resulting from spasticity and inadequate stretching, may result in functional loss without the patient seeking medical care. Patients with substantial residual disability from a stroke likely benefit from periodic physiatric assessment on an ongoing basis.

REFERENCES

31. Anonymous. Tissue plasminogen activator for acute ischemic stroke. The
32. Hacke W, Albers G, Al-Rawi Y, et al. The desmoplasate in acute isch-
24. Bederson JB, Awad IA, Wiebers DO, et al. Recommendations for the manage-
30. Anonymous. Tissue plasminogen activator for acute ischemic stroke. The
23. Anonymous. Anticoagulation and the risk of cerebrovascular events in pa-
22. Anonymous. Guidelines for the management of patients with unruptured intracra-
21. Anonymous. Guidelines for the management of patients with unruptured intracra-
18. Leonardi-Bee J, O’Carroll C, Eikelboom JW, et al. Low-dose aspirin and the risk of stroke in people aged <65 years: a systematic review and met-
17. Sacco RL. Risk factors and outcomes for ischemic stroke.
16. Ridker PM, Danielson E, Fonseca FAH. Rosuvastatin to prevent vascular
12. Foulkes MA, Wolf PA, Price TR, et al. The Stroke Data Bank: design,
11. Anonymous. Acute stroke treatment in acute cardioembolic stroke: a meta-analysis of random-
7. Diener HC, Bogousslavsky J, Brass LM, et al. Aspirin and clopidogrel compared with clopidogrel alone after recent ischaemic stroke or tran-
6. Ridker PM, Danielson E, Fonseca FAH. Rosuvastatin to prevent vascular
5. Langhorne P, Pollock A, Stroke Unit Trialists C, et al. What are the com-
4. Roth EJ. Medical complications encountered in stroke rehabilitation.
3. Mayer SA, Brun NC, Beigi A, et al. Functional recovery following reha-

PART III | MAJOR CONDITIONS

7. Mayer SA, Brun NC, Beigi A, et al. Functional recovery following reha-
6. Ridker PM, Danielson E, Fonseca FAH. Rosuvastatin to prevent vascular
5. Langhorne P, Pollock A, Stroke Unit Trialists C, et al. What are the com-
4. Roth EJ. Medical complications encountered in stroke rehabilitation.
3. Mayer SA, Brun NC, Beigi A, et al. Functional recovery following reha-


PART III | MAJOR CONDITIONS


Traumatic Brain Injury

NATURE OF TRAUMATIC BRAIN INJURY

Traumatic brain injury (TBI) is one of the major causes of disability worldwide. The systematic study of the residual effects of TBI can be traced to World War II and the work of Alexander Luria, Kurt Goldstein, and others (1,2). In this early work, much was learned about the deficits following penetrating injuries to the brain in soldiers with gunshot wounds. The pattern of residual dysfunction often corresponded to a focal lesion caused by the bullet passing through the brain. These focal deficits were similar to those observed in strokes.

Most TBIs are classified as closed head injuries, wherein the skull is not actually penetrated. The nature of injury sustained in vehicular accidents and falls (e.g., blunt impact, acceleration-deceleration) often results in multifocal lesions and diffuse brain damage with a variety of physical, cognitive, and neurobehavioral impairments that are unique to each person and pose formidable obstacles to community integration. Modern international conflicts involving the use of improvised explosive devices have led to a new injury mechanism—blast injury—the biophysics and clinical outcomes of which are not yet fully understood (3).

TBI affects all age groups. However, complex issues are posed by children with TBI because the injury interacts with the processes of biological, psychological, and social development. The special problems of children with TBI are not treated in detail here, but there are excellent sources for additional information on this topic (4,5). At the other end of the age spectrum, one finds the older adult with brain injury, whose recovery tends to be slower than the younger population and may be complicated by preexisting comorbidities and the reduced plasticity of the aged brain (see also Englander and Cifu, 1999 for a review of this topic (6)). Given the increasing percentage of older persons within the general population, it is likely that a growing number of survivors of brain injury will be in this subgroup.

The Range of Outcomes: Death to Complete Recovery

Injury type and severity vary widely. In some injuries, commonly referred to as concussions or mild traumatic brain injuries, the person may suffer only a period of confusion or a brief episode of unconsciousness. Most of those with mild TBIs return to normal activity within days. In a minority of cases, somatic, cognitive, and affective symptoms may persist for weeks, months, or longer. Mild brain injuries constitute the vast majority of TBIs. At the other extreme, death from severe brain injury is estimated at 20.7/100,000 population, which represents a 22% decline from 1979 (7). Most rehabilitation efforts are focused on survivors of moderate and severe TBI, among whom 80 to 90,000 survive each year with disability. Nearly 2% of the U.S. population requires ongoing assistance with activities of daily living (ADL) as a result of a TBI (8,9).

Measures of Injury Severity

Both depth and duration of coma have been considered to be indices of the severity of TBI. Clinical assessment of coma was made more precise and objective with the advent of the Glasgow Coma Scale (GCS) in the 1970s (10), a quantitative measure of the depth of unconsciousness (Table 24-1). Coma is defined as not opening the eyes, not obeying commands, and not uttering understandable words. A GCS score of 8 or less in the acute period is operationally defined as a comatose state (11). Ironically, however, the use of the GCS as a marker of injury severity has become more problematic with changes in emergency management, in that increasing numbers of patients are intubated and subjected to chemical paralysis in the field, making it impossible to record an accurate GCS score. The Full Outline of Unresponsiveness (FOUR) is a recently developed measure of depth of unconsciousness that adds assessment of brainstem reflexes, and does not require verbal responding, thus avoiding the confound of intubation. The FOUR is of comparable reliability to the GCS (12), but it does not avoid the problem of chemical paralysis and is too new to have accumulated a large volume of prognostic data.

Duration of coma is often defined as the time until the patient resumes following commands. Both duration of coma and initial GCS score have been reported to predict neurobehavioral outcome from TBI (13,14), varying with the outcome measures selected and the time since injury. Stein and Spettell (15) showed that when the GCS was modified slightly to include complications such as intracranial lesions, its predictive power was improved.

The duration of posttraumatic amnesia (PTA) has been used as a measure of injury severity as long ago as the 1930s (16). During PTA, patients are out of coma but remain disoriented and amnesic for day-to-day events. Duration of PTA is measured from the onset of TBI to the resumption of ongoing memory; the duration of coma, if any, is thus included. Researchers have used PTA as an index of injury severity and an important predictor of outcome (17,18). Retrospective
measurement of PTA duration from examination of medical records, or by asking the person to estimate the amnesic interval, can be unreliable. The Galveston Orientation and Amnesia Test (GOAT) provides an objective, widely used, and reliable way of measuring PTA prospectively (19). However, it has been criticized on the grounds that although it assesses orientation, it fails to capture the amnesia characteristic of PTA.

An alternative, the Westmead PTA Scale, includes a measure of day-to-day memory as well as standard orientation questions (20). Duration of PTA, assessed prospectively with the Westmead, is a strong predictor of long-term outcome variables such as employment status (21–24).

The predictive power of PTA duration is better for patients whose brain damage is caused primarily by diffuse axonal injury (DAI) than for those with primarily contusions or other focal brain injury (25). Another alternative, known as the Orientation Log (O-Log), has been used increasingly and found to be easier to administer than the GOAT. Unlike the GOAT, the O-Log allows for cueing to assist individuals with language or severe memory impairments, and it does not require the patient to verify information about events relating to the injury (26). The initial O-Log score, combined with measures such as time since injury and number of O-Log administrations, accurately predicted resolution of disorientation in 76% of a sample of 389 individuals aged 10 to 93 years (27).

### EPIDEMIOLOGY OF TBI

The epidemiology of TBI varies with injury severity and mechanism. Based on 2003 estimates from the United States Centers for Disease Control and Prevention (CDC), the overall incidence of TBI resulting at least in an emergency department visit was 538.2/100,000, meaning that approximately 1 in every 200 individuals sustains a TBI annually. The incidence of mild TBIs that do not lead to emergency department visits (i.e., are unevaluated or evaluated in a private physician’s office) is unknown. The vast majority of individuals—nearly 80%—sustain a “mild” TBI (i.e., a concussion) and are discharged directly from the emergency department. The rate of hospitalization for mild TBI has declined substantially from 130/100,000 population in the 1970s to 51/100,000 in the 1990s, perhaps due to changes in admission criteria (8).

Approximately 290,000 individuals are hospitalized and discharged alive annually (about 18.6% of the total), most of whom have moderate or severe injuries. More than 50,000 individuals (about 3.25% of the total or 15% of those with serious injuries), die prior to hospitalization or before hospital discharge (28) (Table 24-2). The current rate of hospitalization represents an approximately 50% decrease from studies conducted 20 to 25 years ago. It is considered to be a reflection of substantially fewer “mild” cases being hospitalized, rather than a drop in incidence of TBI. Precise measures of prevalence within the population are unknown; however, the CDC estimates that approximately 5.3 million Americans are living with a TBI-related disability (29), with about 125,000 new individuals annually requiring assistance with ADL at least a year after their injuries (30).

These overall figures obscure significant differences in injury severity by age, gender, ethnicity, and mechanism of injury. For example, adults and men, on average, sustain more
severe injuries than children and women (28). In addition, the rate of TBIs is elevated among African-Americans and Native Americans compared to other ethnic groups (30,31).

The most common etiologies of TBI include falls, motor vehicle accidents, and assaults, with falls now outnumbering motor vehicle injuries in many regions, as a result of the aging of the population. Motor vehicle accidents continue to have the highest fatality rates, however, and 40% to 50% of those injured were not wearing protective equipment (seat belts or helmets) at the time of injury (28). The mechanism of injury, however, also differs for different social groups, with injuries from firearms seen at the highest rate among middle-aged men (with rates on the decline), motor vehicle injuries seen most among adolescents and young adults (with rates also declining), and falls seen most among the elderly (with rates on the rise) (32). Elderly patients have higher death rates and a slower recovery process, compared with the young adult population (33).

A variety of risk factors have been identified as influential in determining who is likely to sustain a TBI. The most common factor cited is alcohol intake before the TBI (34). In the TBI Model System database as of 2001, approximately 50% of those screened for blood alcohol level (which includes about 80% of cases) were legally intoxicated at the time of injury (35). Other factors have been noted, such as preinjury personality disturbance, history of attention deficit hyperactivity disorder (ADHD) (36), family discord, or antisocial behavior, but little systematic research has been done to relate these factors to risk of injury (37). Helmet use by both motorcyclists and bicyclists reduces the severity of injuries that occur (38,39).

The economic and social impact of TBI is considered enormous but has not been extensively researched to date. In a study based on more than 300,000 individuals with TBI, Max et al. estimated that the total lifetime cost for all people who sustained TBI in 1985 in the United States was $37.8 billion (40). Charges for acute care and rehabilitation alone in the Model System database averaged about $120,000 per patient, excluding physician charges (41). Estimates of return to work vary greatly, from 15% to 100%, depending on the admission criteria of the vocational program under study (42,43). In a population-based study, conducted without regard to whether or not individuals received vocational rehabilitation services, 47% of those with severe injuries who were working prior to injury, were working again 1 year later. Comparable figures for those with moderate and mild injuries were 78% and 81%, respectively (44). TBI creates strain in intimate relationships, affects role functioning, fosters economic hardship, and creates a great burden on the family (45). These issues are discussed further later in this chapter.

### PATHOPHYSIOLOGY OF TBI

#### Primary Injury

Primary injury is defined as damage that occurs directly and immediately as a result of trauma to the brain. Cortical contusion and DAI are the two subtypes of primary injury. DAI is the distinguishing feature of TBI. Acceleration-deceleration and rotational forces that commonly result from motor vehicle accidents produce diffuse axonal disruption. The direction in which the force is applied affects the severity of injury, with lateral impact leading to poorer outcome than head-on or rear-end collisions (46). Depending on the severity of injury, such lesions may be microscopic, or they may coalesce into focal macroscopic lesions, with a preponderance in the midbrain and pons, corpus callosum, and white matter of the cerebral hemispheres.

### TABLE 24.2 Numbers and Rates of TBI-Related Emergency Department Visits, Hospitalizations, and Deaths by Age Group, United States 2003

<table>
<thead>
<tr>
<th>Age (years)</th>
<th>Number</th>
<th>Rate</th>
<th>Column</th>
<th>Number</th>
<th>Rate</th>
<th>Column</th>
<th>Number</th>
<th>Rate</th>
<th>Column</th>
<th>Total</th>
<th>Rate</th>
<th>Column</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>b</td>
<td>%</td>
<td></td>
<td>b</td>
<td>%</td>
<td></td>
<td>b</td>
<td>%</td>
<td></td>
<td>b</td>
<td>%</td>
</tr>
<tr>
<td>0–4</td>
<td>216,000</td>
<td>91.2</td>
<td>17.6</td>
<td>18,000</td>
<td>92.1</td>
<td>6.3</td>
<td>1,035</td>
<td>5.2</td>
<td>2.0</td>
<td>235,000</td>
<td>118.5</td>
<td>15.0</td>
</tr>
<tr>
<td>5–14</td>
<td>188,000</td>
<td>458.2</td>
<td>15.3</td>
<td>24,000</td>
<td>58.3</td>
<td>8.4</td>
<td>1,250</td>
<td>3.1</td>
<td>2.5</td>
<td>213,000</td>
<td>520.5</td>
<td>13.6</td>
</tr>
<tr>
<td>15–24</td>
<td>313,000</td>
<td>760.1</td>
<td>25.6</td>
<td>56,000</td>
<td>135.4</td>
<td>19.2</td>
<td>9,053</td>
<td>22.0</td>
<td>17.8</td>
<td>378,000</td>
<td>917.5</td>
<td>24.2</td>
</tr>
<tr>
<td>25–44</td>
<td>254,000</td>
<td>301.0</td>
<td>20.7</td>
<td>58,000</td>
<td>69.1</td>
<td>20.1</td>
<td>13,904</td>
<td>16.5</td>
<td>27.4</td>
<td>326,000</td>
<td>386.7</td>
<td>20.8</td>
</tr>
<tr>
<td>45–64</td>
<td>163,000</td>
<td>237.9</td>
<td>13.4</td>
<td>50,000</td>
<td>72.4</td>
<td>17.1</td>
<td>11,698</td>
<td>17.0</td>
<td>23.0</td>
<td>225,000</td>
<td>327.3</td>
<td>14.4</td>
</tr>
<tr>
<td>≥65</td>
<td>90,000</td>
<td>251.8</td>
<td>7.4</td>
<td>84,000</td>
<td>234.1</td>
<td>29.0</td>
<td>13,796</td>
<td>38.4</td>
<td>27.2</td>
<td>188,000</td>
<td>524.3</td>
<td>12.0</td>
</tr>
<tr>
<td>Total</td>
<td>1,244,000</td>
<td>420.9</td>
<td>100.0</td>
<td>290,000</td>
<td>99.9</td>
<td>100.0</td>
<td>50,757</td>
<td>17.5</td>
<td>100.0</td>
<td>1,565,000</td>
<td>558.2</td>
<td>100.0</td>
</tr>
</tbody>
</table>

*Numbers and rates may not sum to totals due to rounding.

*Rate per 100,000 population.

*Total includes 21 deaths with unknown age.

*Sample size is 30 to 59; the value of the estimate was reported but may not be stable.

Note: Emergency department visits does not include 247,000 hospitalizations, deaths, transfers, or persons with an unknown disposition. Hospitalizations do not include 21,000 persons who died in-hospital.
DAI is primarily responsible for the initial loss of consciousness (50). The precise mechanisms of axonal damage remain controversial but include direct axonal shearing and disruption of the intra-axonal cytoskeleton that may lead to axonal swelling and disconnection (51). Such injury may be a risk factor for development of Alzheimer’s dementia (52). Animal data suggest that some of the loss of axonal integrity may happen after a delay (51), allowing the possibility that preventive treatments may be developed.

Cerebral contusion is the other main type of primary injury. These cortical bruises commonly occur at the crests of the gyri and extend to variable depths, depending on severity. Contusions often occur on the undersurface of the frontal lobes, as well as the frontal and temporal tips, regardless of the site of impact, due to the internal architecture of the skull (Fig. 24-2). The lesions usually are bilateral but may be asymmetric. Coup-contrecoup injuries are more likely seen when the moving head hits a stationary object, such as with a fall. In this situation, one may see a contusion at the site of impact as well as another, often larger, contusion on the opposite cortex. Cerebral contusions may produce focal cognitive and sensory motor deficits and are risk factors for seizure disorders but are not directly responsible for loss of consciousness. In contrast to DAI, contusions may result from relatively low-velocity impact such as blows and falls. A given patient’s pattern of functional
deficits may be more focal (e.g., from contusions) or diffuse (e.g., from DAI) or may include features of both. The balance of these two pathologic features influences the nature of neurobehavioral deficits. Deficits related to DAI tend to recover gradually, with the pace of recovery inversely related to the duration of coma, whereas recovery from deficits related to cortical contusions depends more on the size and location of the focal injury (25).

**Secondary Injury and Initial Neurosurgical Management**

Secondary injury can be defined as any damage to brain tissue that takes place after the initial (primary) injury. Whereas the only intervention for primary injury is prevention (with the possible exception of regenerative treatments), secondary injuries are at least to some degree treatable and theoretically preventable. The primary injury may set in motion a variety of pathologic processes that result in more severe and widespread brain damage. These processes range from subcellular events to those involving multiple organ systems and often work in concert. Apoptosis and excitotoxicity are examples of intracellular processes that lead to secondary injury, but can be driven by factors taking place at the tissue or systemic level. Local ischemia may lead to direct neuronal necrosis or initiate excitotoxicity by increasing levels of extracellular glutamate (53). This, in turn, leads to influx of calcium into neurons, which activates a number of proteases and increases the production of free oxygen radicals, resulting in intracellular injury and cell death (54). The local ischemia may be a result of several factors including direct trauma to blood vessels, systemic hypoxia, and cerebral hypoperfusion.

Cerebral hypoperfusion is assessed by evaluating the cerebral perfusion pressure (CPP), which is defined as the difference between mean arterial pressure (MAP) and intracranial pressure (ICP). Any factor that increases ICP or decreases MAP can decrease CPP, leading to increased ischemic injury. Management of CPP requires consideration of the pressure dynamics within the cranium, which is essentially a closed structure that contains brain tissue (cells and extracellular fluid), blood, and cerebrospinal fluid (CSF). Increases in the volume of any of these components will increase ICP. Expanding extra-axial or intraparenchymal hematomas, acute hydrocephalus, and brain edema are well-known causes of increased ICP, whereas systemic hypotension reduces CPP by decreasing MAP.

There is class II evidence to support a recommendation for maintaining CPP between 50 and 70 mm Hg (normal value for adults: 70 to 100), whereas data suggest that aggressive efforts to keep CPP above 70 mm Hg increase the risk of developing adult respiratory distress syndrome (55). Current guidelines recommend ICP monitoring in patients with severe head injury (GCS of 3 to 8) and an abnormal computed tomography (CT) scan (hematomas, contusions, edema, or compressed basal cisterns); or a normal CT scan and two out of three adverse features (age >40, unilateral or bilateral motor posturing, or systolic blood pressure <90 mm Hg) (56).

Increased ICP is not treated in isolation. However, current guidelines recommend initiation of intervention when ICP is greater than 20 mm Hg. A number of options exist to decrease ICP. Bolus doses of the osmotic diuretic mannitol may be employed for increased ICP secondary to edema, providing MAP is maintained through adequate fluid replacement (57). There is increasing interest in the role of hypertonic saline as hyperosmolar therapy as well (58). Use of high-dose sedatives such as propofol and pentobarbital is also supported (55). Hyperventilation, which had been a mainstay for the management of high ICP, is now only recommended in very specific cases, while corticosteroids are not recommended at all, and are in fact contraindicated for persons with moderate and severe TBI (55). A craniectomy is another strategy used to decrease ICP, by increasing cranial volume. A recent review did not identify randomized, controlled studies to support this intervention in TBI, although some less rigorously designed trials have suggested utility. Randomized trials are ongoing (59).

Systemic factors such as anemia, hypotension, pulmonary injury, and cardiac or respiratory arrest also may contribute to secondary injury by diminishing the delivery of oxygen to the injured brain. Brain infection may occur from open skull fractures, CSF rhinorrhea, or iatrogenically from ICP monitoring. Protracted seizures also can lead to secondary injury through increased metabolic requirements, disruption of spontaneous respiration, direct injury, and aspiration.

**Neurodiagnostic Tools**

Many imaging and neurophysiological technologies are available to assist in TBI management. In the acute postinjury period, CT scanning can detect intracranial hemorrhage, brain swelling, hydrocephalus, and infarction. The raw data can be adjusted to better evaluate structures of different radiodensities, for example, bone windows to identify skull fractures. However, CT is not sensitive in identifying small contusions, white matter injury, or in the evaluation of the posterior fossa (60). In the postacute phase, CT scanning can be useful, especially when a patient’s neurological status is deteriorating or failing to progress as anticipated. In such situations, CT may identify progression of hydrocephalus or hygromas, evidence of increased ICP, or new bleeds.

MRI has some advantages over CT, including lack of x-ray exposure, greater resolution in the brainstem, better identification of isodense collections of blood, and detection of small white matter lesions (61). However, MRI takes longer to perform, and requires that the patient not have any MRI-incompatible implants or equipment. In the broadest sense, CT is of greatest use in the acute setting, while MRI may be more appropriate in the subacute and chronic setting.

It is important to note that the absence of findings on CT or MRI does not rule out a TBI (62). Newer neuroimaging techniques may be more sensitive to structural and functional changes in the brain after trauma. Diffusion-weighted MRI (DWI) has been utilized to identify cytotoxic edema and to quantify the size of lesions after TBI. Studies have been carried out to correlate this with outcome (63). DWI has also been
used to increase the sensitivity of MRI in detecting DAI (64). More recently, diffusion tensor imaging, a technique derived from DWI, has also been utilized to determine the presence of DAI as well as evaluate changes in white matter tracts over time (65,66).

It may be that functional imaging can play a role in evaluating the brain after TBI due to its ability to provide information about the metabolic activity of specific regions at rest as well as during mental activities. Single photon emission CT (SPECT), positron emission tomography (PET), and functional MRI (fMRI) can all measure regional cerebral perfusion, although only PET and perfusion fMRI can quantify blood flow in absolute terms. Since perfusion may be compromised in structurally intact brain tissue, either as a result of reduced vascular delivery, or reduced perfusion demand by inactive neural tissue, reductions in flow may identify areas of functional compromise (67). Some fMRI studies conducted later postinjury have found that individuals with TBI demonstrate more widespread cortical activation during mental tasks than uninjured controls, suggesting that increased mental effort is required to perform tasks after injury. This finding has been seen in mild TBI as well as with more severely injured persons (68,69).

Another technique that is based on MRI technology is magnetic resonance spectroscopy (MRS). Although MRS studies utilize conventional MRI systems, they require different software that detects and quantifies certain brain metabolites that may be markers of particular pathophysiologic processes. For instance, choline can be measured, and levels may be related to the degree of membrane injury from DAI and other pathologic processes (70). Glutamate may be a marker for metabolic abnormalities associated with brain injury (71). Shutter et al. have reported that MRS data obtained early in the acute hospital stay can be highly predictive of 6- and 12-month Glasgow Outcome Scale (GOS) scores, especially when combined with clinical data such as GCS motor subscore (72).

Neuroprotection, Neuronal Recovery, and Functional Recovery

Recovery from TBI often is incomplete. Reports have identified subtle but persistent deficits after even mild TBI (73,74). Yet, many people who sustain TBI make tremendous gains from coma to the reemergence of a variety of complex skills. This recovery is believed to occur at multiple levels, from alterations in biochemical processes to changes in family structure (75). A better understanding of the pathophysiologic mechanisms involved in secondary injury, as well as improved understanding of the processes that underlie functional recovery may lead to improved treatments to restore function in injured individuals.

As mentioned previously, there is an increased understanding regarding the mechanisms that contribute to secondary injury. Thus, in theory, interventions could be designed that block various ion channels, scavenge free radicals, inhibit excitatory neurotransmitters, or block the internal signals for programmed cell death, thereby salvaging a greater proportion of neural tissue that is not irreversibly damaged by primary injury. Despite the theoretical promise, however, the clinical trials that have studied such interventions have been generally disappointing (76). There are many possible reasons for the failure to identify useful neuroprotective treatments, including the heterogeneity of pathophysiologic mechanisms at work, difficulties in extrapolating from animal models, outcome measures chosen, and the possible need to combine different therapeutic mechanisms to maximize impact (77,78). An active search for additional neuroprotective agents continues, with some preliminary promising results (79–82).

Once a given degree of brain damage has occurred, a number of processes, overlapping in time and pathobiology, contribute to functional recovery. Restoration of effective perfusion and oxygenation, resolution of edema, and weaning of sedating drugs can lead to improvements in function in the first days and weeks after injury. Latent neuronal circuitry may be unmasked by injury, supporting functions that were previously carried out by neurons that are no longer functioning (83). Experience, often in the form of intense practice, has been shown to alter sensory-motor cortical representations (“maps”) in healthy humans as well as lesioned animals, suggesting that appropriately structured therapy might directly contribute to improved function after brain damage (84–86). Other interventions such as transcranial magnetic stimulation may also positively influence plasticity and cortical reorganization after TBI but await rigorous assessment (84). Resolution of diaschisis, either spontaneously, or through the actions of neurotransmitter agonists, may also play a role, and receptor antagonists may slow down recovery (85–87). Animal studies of neural transplantation are in progress and human trials in more focal neurological conditions have already appeared (88). At a more behavioral level, a person with a particular deficit may discover (or be taught) another strategy to accomplish the task that relies on remaining intact neural systems (89). Finally, as long as an individual is conscious and retains some learning capacity, it may be possible for them to learn specific tasks and activities, even though their overall capacity may remain fixed.

Although there is considerable evidence for several recovery mechanisms, particularly in the animal literature, there is as yet limited systematic application of these principles to the design of human rehabilitation efforts. Nevertheless, the growth of basic science research related to neural recovery is increasingly being translated into clinical research, with the anticipation of more theoretically based rehabilitation efforts in the future (90).

OUTCOME AFTER TBI

Types of Deficits

The majority of survivors of severe TBI emerge from coma and achieve remarkable progress toward regaining their preinjury functional abilities. Most, however, are left with a combination of physical and neurobehavioral impairments, which
Outcomes collectively result in activity limitations and reduced societal participation. Each individual’s specific pattern of deficits is a consequence of the severity of the injury, nature of brain damage, and medical complications, varying greatly from one person to the next. However, deficits in cognition, particularly affecting attention, memory and/or executive function, as well as fatigue are nearly ubiquitous after moderate or severe TBI.

Changes in behavior, mood, and personality after TBI may be the most difficult impairments to manage effectively. Behavior problems range from minor irritability or passivity to labile emotional displays or disinhibited, bizarre, or aggressive behavior. The severely brain-injured person may appear egocentric and childlike, and may display a loss of empathy and concern for others.

Limitations in activities related to self-care, toileting, mobility, basic communication, and feeding commonly occur in severe TBI. In those with less severe injuries, difficulties may not be revealed until more complex community-oriented skills are attempted, such as time and money management, community mobility, and school or vocational tasks. Restrictions in vocational, educational, and interpersonal domains are widespread. There is frequently loss or diminution of friendships and intimate relationships, growing psychological distress, and a devastating impact on the family system. The psychosocial consequences of TBI, and the cognitive and behavioral impairments that underlie them, are reviewed in greater detail later in this chapter.

Outcome Measurement
The measurement and prediction of outcomes have taken on increased importance in health care generally and rehabilitation specifically. There are several distinct uses for outcome measurement, with different implications for the types of outcomes measured and their interpretation. Outcome data for clinical program evaluation systems in brain injury rehabilitation are required by the Commission for the Accreditation of Rehabilitation Facilities (CARF), with the intent of ensuring that the types of functional outcomes achieved by each program are reasonable with respect to programmatic goals. However, the quality of functional outcomes achieved by clinical programs is a function of demographic and clinical characteristics of their client populations as well as the effectiveness of their treatment services. Consequently, comparison of the quality of different programs can only be made after appropriate “case mix adjustment”—a technique that is still in a rather primitive state.

Outcome measurement can also identify treatments that appear promising in producing outcomes superior to the prior standard of care, which can then be confirmed in controlled research. Highly accurate outcome prediction could also, in principle, guide individual treatment decisions. For example, if it could be predicted in the first few days postinjury that a patient would remain permanently vegetative, the physician and the family might choose not to pursue aggressive treatment; if independent ambulation were predicted not to be possible, the energies of physical therapy could be redirected to other goals, and a definitive wheelchair could be prescribed earlier. For outcome prediction to be used in this way, however, it would need to be highly accurate in projecting the specific outcomes of individual patients; unfortunately, this goal has not been achieved.

The outcome scale used in many studies of acute TBI is the GOS. It has been demonstrated to have a high degree of interrater reliability (91). It has been used to correlate early injury severity measures (e.g., GCS, length of PTA) and outcome at 6 months postinjury. Several major drawbacks in the utility of the scale for rehabilitation purposes have been identified:

- The categories are so broad that it is not a sensitive measure of progress during rehabilitation.
- The global categories do not provide a real indication of functional abilities.
- Cognitive and behavioral dysfunctions are poorly addressed in the outcome categories.

Despite these limitations, it continues to have widespread use for its intended purpose—to provide a quantitative, general way of describing outcome. An extension of the GOS, the GOS-extended (GOS-E) seeks to remedy some of the shortcomings of the GOS by providing a broader range of outcome scores to improve sensitivity (92), and by incorporating a comparison with preinjury status in both medical and psychosocial domains (93) (Table 24-3).

Several other outcome measurement scales have been developed to be more sensitive to functional changes that are of interest in rehabilitation. Some of these are brain-injury-specific, and others are general rehabilitation scales.

**TABLE 24.3 Glasgow Outcome Scale-Extended**

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Dead</td>
</tr>
<tr>
<td>2</td>
<td>Vegetative state (VS)</td>
</tr>
<tr>
<td></td>
<td>No cerebral cortical function that can be judged by behavior (not able to follow simple commands or communicate)</td>
</tr>
<tr>
<td>3</td>
<td>Lower severe disability (lower SD)</td>
</tr>
<tr>
<td></td>
<td>Needs full assistance in ADL throughout the day</td>
</tr>
<tr>
<td>4</td>
<td>Upper severe disability (upper SD)</td>
</tr>
<tr>
<td></td>
<td>Needs some supervision/assistance in ADL, but can be alone for &gt;8 h/d</td>
</tr>
<tr>
<td>5</td>
<td>Lower moderate disability (lower MD)</td>
</tr>
<tr>
<td></td>
<td>Independent in ADL, and can shop and travel independently on public transportation, but has not returned to previous position or lifestyle</td>
</tr>
<tr>
<td>6</td>
<td>Upper moderate disability (upper MD)</td>
</tr>
<tr>
<td></td>
<td>Able to resume previous position or lifestyle with alternative/modified duties or part-time due to injury</td>
</tr>
<tr>
<td>7</td>
<td>Lower good recovery (lower GR)</td>
</tr>
<tr>
<td></td>
<td>Able to resume previous position or lifestyle (may be modified), but reporting some problems</td>
</tr>
<tr>
<td>8</td>
<td>Upper good recovery (upper GR)</td>
</tr>
<tr>
<td></td>
<td>Complete return to previous lifestyle with no reported problems</td>
</tr>
</tbody>
</table>
The Disability Rating Scale (DRS) was developed specifically for TBI and is intended to assess changes “from coma to community” (94). This scale produces a quantitative index of disability across ten levels of severity, with high interrater reliability, and is considerably more sensitive to clinical change than is the GOS, particularly at the severe end of the spectrum, with less sensitivity to improvement within higher levels of functioning (95,96).

The Rancho Los Amigos Levels of Cognitive Functioning Scale focuses on cognitive recovery after TBI and the capacity to interact effectively with the environment. Each level is accompanied by a lengthy description of behaviors that meet the criteria for placement at that level (97). A comparison of this scale with the DRS revealed that it has slightly lower validity and reliability (98).

The Functional Independence Measure (FIM), an 18-item rating scale, is the most widely used outcome measurement scale in medical rehabilitation, and is now incorporated into the federal prospective payment system for inpatient rehabilitation. Although the FIM contains both motor and cognitive subscales, its scoring is weighted toward the physical items, and it fails to strongly capture the individual's capacity for self-direction. Because of the relative insensitivity of the FIM to cognitive and behavioral deficits, the Functional Assessment Measure (FAM) was developed to supplement it with more cognitively oriented items. Interrater agreement for the FAM is lower (67%) than for the FIM (88%), and although it extends the range of difficulty somewhat compared to the FIM alone, its items are highly redundant with existing FIM items in terms of their scoring (96).

All the outcome scales discussed until now are well suited to measuring the changes that occur during acute rehabilitation, but they focus primarily on the activity level of the World Health Organization’s International Classification of Functioning, Disability, and Health (ICF) (99) and none of them adequately addresses higher-level functions for those with mild injuries or for those with severe injuries who have been discharged into the community (100).

Measures of behavioral and social functioning in the community that have been developed specifically for TBI outcome measurement include the Neurobehavioral Functioning Inventory (NFI) (101) and the Mayo-Portland Adaptability Inventory (MPAI) (102). The NFI is a self-report measure of the frequency of neurobehavioral symptoms, completed by either the injured person or a family member, that yields subscale scores in six domains of emotional/behavioral function: depression, somatic, memory/attention, communication, aggression, and motor. Both construct validity and criterion-related validity were demonstrated to be acceptable (103). However, this measure lacks sensitivity to the more subtle long-term problems in the domains of executive function and social behavior (104). The MPAI was developed to provide an evaluation of progress during postacute brain injury rehabilitation. It is a 30-item rating scale completed by staff or survivor on two fundamental dimensions derived from a principal components analysis: Physical/Cognitive Impairment Scale and Social Participation Scale. Both person reliability (0.82) and item reliability (0.96) were found to be good to excellent (102).

The first outcome measure developed specifically to measure community functioning of persons with brain injury was the Community Integration Questionnaire (CIQ) (105). This 15-item questionnaire has been demonstrated to consist of three subscales: home integration, social integration, and productivity. It can be completed by the person with brain injury or a significant other. The CIQ has been found useful as a measure of “objective” quality of life and level of productivity, but has also been subject to criticism for its psychometric properties (106). Another measure of community functioning, originally developed for persons with spinal cord injury, is the Craig Handicap Assessment and Reporting Technique (CHART) (107). This measure focuses on measurement of handicap or participation restriction. Original subscales included physical independence, mobility, occupation, social integration, and economic self-sufficiency. A revised version of the scale, which included a dimension identified as Cognitive Independence, was developed more recently (108). This additional scale also showed high test-retest reliability (0.87) in a mixed sample of individuals with neurologic illness including TBI and was also found to be a useful tool in evaluating postacute functioning in persons with TBI (109).

Neither the CHART nor the CIQ measures change relative to preinjury. One measure that does so and has both patient-rated and relative versions is the Sydney Psychosocial Reintegration Scale (SPRS) (Form A) (110,111). It comprises 12 items covering three domains: occupational activity (OA); interpersonal relationships (IR); and independent living skills (LS), rated on a seven-point scale ranging from 0 (extreme change) to 6 (no change), giving total scores from 0 to 24 for each domain and an overall total score from 0 to 72. Other versions of the scale measure current level of functioning (Form B) and as categorized on three levels: good, limited and poor (Form C) (110,111). The SPRS has been shown to have high levels of internal consistency, with Cronbach’s alpha ranging from 0.69 to 0.89 for the three domains and reaching 0.90 for the total score (110). High levels of interrater reliability have also been reported for the total score (0.95) and for the three domains (0.86 to 0.94) (110).

Many of the scales mentioned earlier, as well as other instruments, are described in more detail on the Center for Outcome Measurement in Brain Injury (COMBI) website, located at www.tbims.org/combi.

**Prediction of Outcome**

Prediction of outcome raises questions about both precision and values. While relatively imprecise predictions may still be useful for service planning, modeling reimbursement, or understanding factors associated with recovery, greater precision is required to use outcome prediction for individual clinical decision making and family counseling. Moreover, although many of the outcome measurement scales have levels labeled “good” or “poor,” this should not be assumed to
represent what injured individuals or their families believe are lives worth living. The GCS is the most widely used measure of injury severity and is a primary basis for most early predictions of outcome. The total coma score when taken at 2 to 3 or 4 to 7 days postinjury is highly predictive of outcome at 6 months, as measured by the GOS (91). Scores less than 8 are usually predictive of poor outcome. Duration of PTA also is highly correlated with ultimate outcome, with PTA greater than 14 days associated with greater likelihood of moderate or severe disability (91). More recent outcome studies have suggested that PTA duration is a stronger predictor of long-term outcome than GCS (22,112,113). The rate of early recovery, as reflected in serial DRS scores, is also predictive of final outcome (114).

Multimodality evoked potentials (MEP), a combination of brainstem auditory evoked responses (BAERs), visual evoked responses (VERs), and somatosensory evoked responses (SEPs), also have been used as an early means of assessing neurologic status and predicting outcome. Greenberg et al. found that maximal recovery occurs in approximately 3 months for patients with minimal EP abnormalities, whereas severe EP abnormalities suggest that maximal recovery may extend to 12 months (115). Somatosensory evoked potentials (SEP) can predict the gross outcome of acutely injured patients (116). The bilateral absence of the median nerve SEP suggests a very poor outcome in patients that have sustained severe TBI, whereas preserved SEPs do not guarantee a favorable one (117). While there are those who may advocate the use of evoked potentials to allocate rehabilitation resources to patients with a potential to benefit, studies have shown that evoked potentials should not be the sole basis for prediction of outcome in cerebral injury, because of inadequate sensitivity and specificity (118).

Reactive pupils are associated with better outcomes than nonreactive pupils; 50% of those with reactive pupils achieve the moderate disability or good recovery range, as opposed to 4% with nonreactive pupils (119). An absent oculovestibular response, elicited by injecting ice water into the ear of a comatose patient, is an indication of severe brainstem dysfunction and poorer outcome (11). The presence of intracranial hemorrhage (120), and high levels of creatine kinase (BB fraction, reflecting destruction of brain tissue) measured early after TBI (121) also suggest poorer outcomes. Hyperglycemia and low levels of thyroid hormones have negative prognostic significance, presumably reflecting the severity of the stress response (122,123).

The research from Glasgow and other major TBI centers strongly suggests that the bulk of neurologic recovery from acute brain injury occurs within the first 6 months. The maximal duration of the recovery period is more controversial with some researchers affirming that recovery is virtually complete by 1 year, whereas others assert that recovery can extend 2 years or more postinjury (124). It is clear, however, that certain areas of dysfunction recover more quickly than others. For example, recovery of physical abilities and functional skills such as mobility occurs rapidly, often within 3 months after injury (125). Recovery of more complex mental abilities, as assessed by neuropsychological measures, appears more variable. This aspect of recovery has been studied in large samples in the Traumatic Coma Data Bank study (126) and in the TBI Model Systems Project (127). Impairments of attention, information processing speed, memory, and executive function have been shown to persist up to 10 years postinjury (104). Preinjury medical and psychological factors also may affect the prognosis. For example, the presence of a prior TBI or neurologic deficit is likely to slow the recovery process. Also, if cognitive or behavioral abnormalities existed before the injury, there is a greater likelihood of a slower and less complete recovery. Acquired brain damage is thought to exacerbate preexisting behavior disorders (128). Demographic factors, including education, age, and preinjury employment status have also been shown to influence outcome (22,129,130).

Multivariate models combining PTA with age, preinjury occupational status, and early physical and cognitive disability have accounted for 60% or more of the variance in employment outcome at 1 and 6 years postinjury (24,131).

There probably is no final endpoint to the recovery process; rather, the pace of recovery slows, and its scope narrows, though occasional cases of impressive late improvement occur. Even those with permanent cognitive and physical impairments can continue to learn new skills for solving particular functional problems, albeit slowly. Thus, neurologic and cognitive recovery merge imperceptibly into ongoing learning and adaptation.

THE ASSESSMENT AND TREATMENT PLANNING PROCESS

The rehabilitation needs of the survivor of severe TBI often begin at the emergency department but are not likely to end for many years. Although many medical conditions and physical deficits stabilize within 1 year after injury, the presence of long-term psychosocial disorders often necessitates ongoing intervention (132). One year after injury, a quarter to a third of those with severe injuries report unmet service needs, including need for assistance coping with cognitive impairments, emotional distress, and managing financial matters (133). Survivors of TBI have a continuing need for treatment and support services across the lifespan, particularly to address psychosocial difficulties (134).

Treatment and Support Options

Many hospital and community-based services may be of help to survivors of TBI. Changes in private and governmental funding streams stimulated much growth of specialized TBI rehabilitation services in the 1980s and early 1990s, but subsequent funding cutbacks and managed care practices have reduced their availability.

Initial management takes the form of aggressive neurosurgical intervention to minimize secondary injury. The physiatrist may be called in as a consultant even when the patient is still in the ICU, to assist the acute care team in
preventing complications such as contractures, pressure ulcers, heterotopic ossification (HO), and bowel and bladder problems that may impede later rehabilitation of surviving patients, and to assist in choosing medications that minimize sedation.

A tremendous variety of treatment programs may be appropriate for a given individual at various points in the recovery process. Acute inpatient rehabilitation is typically provided to individuals who are able to participate actively in treatment, and whose array of medical, physical, and cognitive deficits precludes a safe community placement. Less intensive subacute rehabilitation may be most appropriate for those who are vegetative, slow to recover, or cannot tolerate intensive therapy. However, many such patients, even though their rehabilitation participation is limited, have intensive medical needs that are not adequately addressed in subacute facilities. A day treatment program may be provided to individuals who can be managed at home but continue to display a variety of physical or neurobehavioral problems. This type of treatment often combines extensive cognitive rehabilitation, behavior management, daily life skills training, community activities, and prevocational activities (135,136). Residential programs may be chosen for individuals who display disinhibited, aggressive, or self-abusive behavior that cannot be managed at home. Such programs aim to reduce inappropriate behaviors and teach more effective means of communication and social interaction, through contingency management and/or pharmacological intervention (137). Transitional living programs assist individuals who have mastered basic ADLs and social interaction skills to live increasingly independently in the community. Typically, the brain-injured person lives in a supervised group home or apartment setting and is given instruction and increasing responsibility in skills needed to live independently (e.g., cooking, cleaning, money management, community mobility, job seeking), with fading of supervision over time. The final goal for many individuals is return to work, most typically through job coaching and supported work, funded by state vocational rehabilitation services.

Many clients with TBI have difficulty applying knowledge gained in one setting to related problems or different environments. In such instances, it is particularly important that the independent living and vocational training be given in the actual community or on the actual job that the individual will be attempting. Some postacute programs deliver community-based services, actively involving the client and family in goal setting and delivering most therapy in the relevant community setting in order to maximize the client’s motivation and engagement in the rehabilitation process and the durability of outcomes (138).

Teletreatment is an emerging facet of service provision that may allow for more accurate assessment and treatment planning in the homecare context, and may allow rural patients to receive services from specialized TBI rehabilitation centers (139,140). As one example, a randomized trial comparing the effects of scheduled telephone intervention after discharge from acute inpatient rehabilitation to usual postdischarge care found a positive impact on 1-year outcomes of this relatively low-cost intervention (141).

During the long rehabilitation process, a case manager may follow patients across multiple treatment services, providing crucial coordination and social support. This type of case manager serves as a liaison among the patient, family, and service providers and gathers medical records, arranges for medical visits, screens programs and facilities, and helps coordinate admission and discharge. With the advent of managed care, intensive case management may be even more crucial although, as noted earlier, it is less often provided (142).

Unfortunately, in many regions, these and other needed services are either unavailable or difficult to access due, in part, to the rapidly evolving changes in health care financing noted previously (143). This calls for creative financing strategies, assisted by case managers, along with advocacy efforts to expand the range of available services. Some states have developed Medicaid waiver programs that extend the array of available services, through the use of public funds (144,145).

**Evidence for the Effectiveness of TBI-Specific Services**

As in most areas of rehabilitation, there is currently insufficient evidence for the efficacy of the services discussed earlier (146,147). However, observational data and randomized pilot studies do provide some evidence to support the impact of comprehensive systems of care for TBI, particularly in the postacute period (136,148,149). One of the few randomized trials comparing an organized program of cognitive rehabilitation to a limited home program failed to find an overall treatment effect, but did find superior outcomes for the cognitive rehabilitation program, for the more severely impaired subgroup (150).

**Patient Assessment and Treatment Planning**

Standards of care and practice for the management of acute and postacute care of persons with TBI have been promulgated by organizations such as the American Association of Neurological Surgeons (56), the Commission on the Accreditation of Rehabilitation Facilities (CARF) (151), and the American Congress of Rehabilitation Medicine (152). Rehabilitation standards emphasize a team model to promote coordination and information sharing across therapeutic disciplines, since a specific cognitive impairment may interfere with the performance and retention of a mobility, communication, or ADL skill. Similarly, disruptive behaviors and lack of initiation cut across all therapy domains. Each of these problems requires the development of a unified team view of the patient’s deficits and needs and the creation of a unified treatment plan.

When a patient enters a rehabilitation service, an initial assessment is needed to guide treatment planning. The goals of this assessment vary with time since injury. Early clinical assessment usually is aimed at defining broad functional areas in need of treatment (e.g., impaired mobility, memory deficits). Ideally, these areas are guided by an assessment of...
the injured person’s lifestyle, needs, and personal and family goals. As the pace of recovery slows, and discharge to home or another facility approaches, therapy focuses more intensely on the specific skills and behaviors that will be prerequisites in the new environment (e.g., toilet transfers, learning a daily schedule). Assessment also depends on injury severity. It is focused more on physical function and basic sensory processing in the severely impaired and on cognitive, social, and vocational function in the mildly impaired.

Discipline-specific assessments must be melded into a patient-oriented assessment for treatment-planning purposes. Typically, severely injured patients have a large number of medical, physical, cognitive, and behavioral impairments. Therefore, the team must develop priorities for treatment, framed in terms of long- and short-term goals, based on the needs and priorities of the patient and family, as well as such factors as estimated length of stay, functional importance of each impairment, age and developmental stage, and prognosis for improvement. As time passes, the goals become more specific and functionally oriented, such that “the patient will have increased ROM in all joints” becomes “the patient will have adequate hip flexion for erect sitting throughout the day.” Identification of the patient’s physical and cognitive strengths can also be very helpful in determining the best way to circumvent specific deficits. Goals being addressed by a single discipline must ultimately be shared by the team. For example, a communication strategy designed by a speech pathologist should be carried out by nurses, other therapists, and family to promote generalization of the skill.

Treatment planning also must consider the ICF (99), and its hierarchy of body structure and function, activity, and participation, and must choose at which of these levels to intervene. For example, it may be concluded that a patient’s ambulation activity limitation is related to limitations in body structure and function in the form of limited range of motion, abnormal tone, weakness, impaired balance, reduced proprioception, and disordered attention. Clinicians then must choose whether to try to address each of these impairments to improve ambulation, or whether to assess the patient’s mobility skills in a motorized wheelchair instead, which would be an activity-level intervention.

A neuropsychological assessment may be important in identifying intact cognitive skills and clarifying cognitive mechanisms responsible for various behavior and skill deficits. In the severely impaired patient, formal testing may be impossible; however, the neuropsychologist’s observation of the patient may still be helpful. Higher-level patients should receive formal testing of such core cognitive areas as attention, learning and remembering, language comprehension and production, visual perception, planning, reasoning, and organization. It should be kept in mind that the results of formal neuropsychological testing have limited predictive ability for real-world function (153). Thus, patients’ skills also should be assessed in naturalistic settings. Moreover, it must be understood that results obtained soon after injury only give an indication of current strengths and weaknesses. Reassessment after longer periods is important to provide a clearer indication of cognitive difficulties that may be more lasting.

**SPECIAL SUBPOPULATIONS**

**The Vegetative State**

About 20% of survivors of severe TBI remain unresponsive 1 month after injury. After 2 to 4 weeks of unconsciousness, coma evolves into the vegetative state, a condition of wakeful unresponsiveness that is characterized by the presence of spontaneous sleep-wake cycles but absence of cortical activity as judged behaviorally (154).

Patients who are vegetative 1 month postinjury still may experience substantial recovery, but their chances of doing so diminish over time. Of patients who are vegetative at 1 month, there is approximately a 50% chance of regaining some degree of consciousness within a year and approximately a 28% chance of improving to a level of independence (155).

Patients continue to emerge from the vegetative state following trauma for at least a year, with rare individuals showing recovery of consciousness even later (little research is available for follow-up periods greater than a year), but nearly all are severely disabled if their emergence is this delayed (156). The term persistent vegetative state has been used extensively in the literature, but without consensus on its definition. It has been suggested that this term be abandoned because it confuses diagnosis (vegetative) with prognosis (persistent). Emergence from the vegetative state following nontraumatic injuries (such as cardiac arrest) is far less likely overall, and very few patients emerge beyond 3 months postinjury (155). This suggests that patients with traumatic injuries complicated by substantial secondary anoxic injury are also likely to have a poorer prognosis than those with uncomplicated trauma.

The following factors have positive prognostic significance for emergence from unresponsiveness: young age, reactive pupils and conjugate eye movements, decorticate posturing rather than decerebrate or flaccid states, early spontaneous eye opening, absence of ventilator dependence or hydrocephalus, shorter time between injury and rehabilitation admission, better scores (within the vegetative range) on the DRS, and more rapid early functional improvement (157–159). Unfortunately, no set of prognostic variables is precise enough to guide early clinical decision making. The life expectancy of those who remain permanently vegetative is not precisely known, but one study of patients in vegetative states of mixed etiologies revealed that almost 75% had died within 5 years (160). Similarly, TBI survivors with severe mobility impairments (though not necessarily vegetative) have a reduced life expectancy due primarily to cardiovascular disease, respiratory disease, choking, and seizures (161).

The minimally conscious state (MCS) refers to individuals who show some evidence of awareness in the form of visual tracking and/or motor behavior that is nonreflexive and contingent on environmental events (e.g., intermittent following of simple commands, replicable pulling out of tubes) but who...
do not consistently follow commands or communicate intelligibly (162). The MCS, like the vegetative state, can be a transitional state on the way to greater recovery or it can be the permanent functional plateau (163). Misdiagnosing patients who are in the MCS as vegetative may occur more than 40% of the time, highlighting the need for rigorous diagnostic assessment (162,164).

Initial coma in TBI probably reflects disruption of brainstem-alerting mechanisms, often with relative preservation of higher brain structures. Brainstem-alerting mechanisms tend to recover with time, however, resulting in return of the sleep-wake cycle. Thus, a vegetative state of long duration generally includes extensive damage to subcortical white matter in higher brain regions, including the thalamus (165,166). Akinetic mutism and the locked-in syndrome may be confused with the vegetative state, but akinetic mutism generally involves damage to the medial frontal lobes, and hypertonia and posturing are absent. The locked-in syndrome generally results from a bilateral pontine stroke, and there is evidence of preserved consciousness and communication through eye movements (167,168).

Establishing emergence from the vegetative state can be done clinically, by observing for volitional responses to the environment, such as following commands, orienting visually to salient objects, attempting to remove tubes and restraints, and the like. Essentially, any behavior that is nonstereotyped and that indicates some evaluation of environmental stimuli is evidence of emerging consciousness. A formal and quantitative assessment strategy is an absolute requirement in working with vegetative and minimally conscious patients; without this, team and family members will disagree about whether or not evidence of consciousness is present and whether improvement is occurring. Several standardized scales are available for objectively grading responsiveness in vegetative and minimally conscious patients (169–172), although only the Coma Recovery Scale-Revised (169) incorporates items that explicitly mark the boundary between VS and MCS. In addition, the principles of single-subject experimental design can be used to answer important clinical questions in individual patients, such as whether they can see (173), follow commands (174), or reliably use a yes/no signaling system; or whether they respond to therapeutic medications (175). Assessment should take place repeatedly and at various times of day because patients may respond inconsistently, particularly as they are first emerging from the vegetative state. When family members report observing volitional behavior that has not been witnessed by staff, individualized assessments of the relevant behaviors can be conducted. Family members may overinterpret reflexive or coincidental behaviors, but staff members may fail to elicit the patient’s best performance. Definitive assessment of the vegetative state must await withdrawal of potentially sedating drugs and ruling out peripheral sensory and motor deficits (e.g., blindness, deafness, extensive polyneuropathy) (164,176).

The possibility that patients who appear vegetative may have some degree of consciousness that is not reflected in observable behavior has been explored recently using functional imaging and event-related potential (ERP) methods. These studies suggest that there exist patients who are capable of following commands to engage in specific mental activities that can be detected physiologically, even though they cannot produce observable behavioral responses (177,178). It is not yet clear how many patients classified as vegetative may be conscious as assessed by more subtle methods, or whether such patients have a better prognosis for the emergence of functionally useful behavior.

Many treatments have been attempted for patients who are in the vegetative state, but none has been subjected to an adequate controlled clinical trial. The pathologic heterogeneity of unresponsive states makes it unlikely that one treatment will help all affected people. Deep brain electrical stimulation of the mesencephalic reticular formation or nonspecific thalamic activating system has been reported to improve the clinical status in some vegetative patients, but small heterogeneous samples and early treatment make it difficult to rule out spontaneous recovery (179–182). A more recent report of late recovery of a patient in MCS after thalamic electrical stimulation does not appear attributable to spontaneous recovery (183). Dopaminergic pharmacologic treatments, including l-dopa, bromocriptine, and amantadine, also have been reported to be of help (184–187). Since all these treatments primarily augment ascending arousing influences, they would seem unlikely to benefit patients with extensive cerebral lesions. Coma stimulation (involving the systematic and frequent provision of sensory stimulation to all sensory modalities) has been widely advocated (75), but studies on it, like those on the pharmacologic agents, suffer from serious methodologic flaws, which have been summarized in review articles (188,189). Large, multicenter clinical trials will be needed to definitively assess the value of early treatments designed to improve recovery, because of the confounding and variable effects of spontaneous recovery. However, such trials are fraught with practical and ethical challenges, particularly when they involve a placebo treatment (190). A number of case reports have been published in which patients who had been vegetative for several years after traumatic or anoxic injuries paradoxically regained consciousness following the administration of a single dose of zolpidem, and whose consciousness could be maintained by repeated dosing of the drug (191). The mechanism of zolpidem response is currently under investigation. The proportion of VS patients capable of responding to zolpidem is unknown but appears to be small (192).

In the absence of definitive treatments to alter the prognosis in the vegetative state, the main goals for rehabilitation are to optimize medical stability, preserve bodily integrity, and objectively define the patient’s current sensory and cognitive capacities with measures that can be monitored for change over time (156,176). This includes screening for adverse medical events such as undiagnosed seizures, hydrocephalus, and endocrine disorders. Attempts to optimize pulmonary hygiene and maintain skin integrity are also essential. Aggressive treatment of hypertonia and contractures is warranted early because they predispose to skin breakdown, interfere with positioning, and
are costly and time consuming to manage in the chronic stage. Finally, sedating medications should be avoided until the prognosis has declared itself. This includes a variety of antispasmodic, anticonvulsant, antihypertensive, anticholinergic, and antihistaminic agents that may have subtle cognitive effects in susceptible patients.

Regular evaluation with a quantitative assessment scale should be carried out. This will reveal subtle improvement that should lead to updated treatment plans, deterioration that should lead to further diagnostic evaluation, or no change, which should lead to family counseling about prognosis and planning toward home or chronic care placement. When a patient remains permanently vegetative, or when a patient remains minimally conscious but has left a specific advance directive, the possibility of forgoing further medical treatments and even life-sustaining fluid and nutrition may be discussed with the family after careful review of local legal guidelines relevant to this area and any institutional policies and ethical guidelines relevant to end-of-life determinations. However, as mentioned, a small number of vegetative or minimally conscious patients may continue to improve over several years, making the determination of permanence more difficult (193).

**Mild TBI**

At the other end of the spectrum from unresponsive states is the patient with mild, or minor, TBI (MTBI), which may occur with or without impact to the head. This is generally defined as a TBI with the following characteristics:

- Loss of consciousness, if any, 30 minutes or less
- PTA 24 hours or less
- Initial GCS 13 to 15
- No focal neurologic deficit
- Negative CT and/or MRI

The initial symptoms of the cerebral injury may be difficult to disentangle from those of the common coincident injuries to the scalp, neck, and peripheral vestibular apparatus (194). Acute complaints after MTBI typically fall into three symptom clusters (195):

- Cognitive: difficulties with attention, speed of information processing, and memory
- Affective: irritability, depression, anxiety
- Somatic: headache, dizziness, insomnia, fatigue, sensory impairments

Headaches are the most common and persistent symptom and the symptom that most strongly differentiates those with mild TBI from trauma controls (196). These symptoms clear within the first few weeks or months postinjury for the majority of patients. Group studies have generally revealed impaired speed of information processing in the early hours or days after injury, but no decrement, or only modest or transient decrement, on neuropsychological measures for individuals with MTBI compared to uninjured controls beyond the first 2 weeks after injury (197). For a proportion of individuals (15% to 25%), however, difficulties persist and are associated with social and vocational failure seemingly out of proportion to the severity of the neurologic insult (197). The etiology of these persistent complaints (often termed postconcussion syndrome, PCS) has been elusive and remains controversial. Premorbid factors such as substance abuse, psychiatric disorder, and age have been implicated but do not explain all cases of persistent disability. The presence of other stressors, preinjury psychological problems, being a student or being in a demanding occupation, being injured in a motor vehicle accident and/or having neck or back injury have also been associated with poorer outcomes (196,198). The idea that pending litigation or financial gain accounts for PCS has not been confirmed (195). In one SPECT study, MTBI patients with unusually persistent disability showed hypoperfusion in the anterior mesial regions of the temporal lobes (199). There is probably no one cause of PCS; premorbid variables, idiosyncratic neurologic vulnerability, and psychological reactions to acute symptoms may all be found to play roles.

The treatment of MTBI should include patient and family education about the typical symptoms and their time frame, and guidance on how and when to resume preinjury activities. Patients with persistent symptoms may benefit from psychotherapy (200), pain management protocols (201), or holistic community reentry programs offering these components along with education, vocational counseling, and group support. Many of the somatic symptoms are responsive to interventions that can be provided by an experienced physical therapist in conjunction with judicious use of medications. Therapeutic interventions may include vestibular habituation exercises (202), Rocabado exercises (203), myofascial release, trigger point injections, nonsteroidal anti-inflammatory medications, and muscle relaxants. For more extensive information, the reader is referred to more detailed publications on this aspect of brain injury (197,204–206).

Prompt diagnosis and treatment of “concussion” in sports related injuries has received increased public awareness. Practice parameters have been established by a subcommittee of the American Academy of Neurology and recommendations were also derived from the Second International Conference on Concussion in Sport (207,208). A Standardized Assessment of Concussion was developed to provide a sideline evaluation that could be performed by nonphysicians (209). Computerized neuropsychological tests have been developed for more standardized evaluation and to provide baseline data to which subsequent postinjury data can be compared (210,211).

**TBI in Children and Adolescents**

Children and adolescents are a high-risk group for TBI, with etiologies varying by age group. Falls and child abuse are overrepresented among infants and toddlers, accidental injuries and motor vehicle accidents in school-age children, and injuries due to violence and risk-taking behavior among adolescents.

The management and sequelae of TBI in children and adolescents are very similar to those of adults, but there are a number of issues that are specific to children. On the
one hand, it may be challenging to disentangle neurological recovery from normal growth and maturation. On the other hand, young children may appear to have minimal deficits after TBI, since the skills expected of them are minimal, only to reveal increasing developmental lags as they mature. Because “independence” is not expected of healthy children, many children with TBI are sent home without rehabilitation services, with their parents providing the assistance needed to function safely (212).

Although it has been estimated that 20,000 children and adolescents reenter school each year with significant disabilities related to TBI, surveys of special education services estimate that only about 12,000 of the more than 5,000,000 U.S. special education students receive services due to a TBI (213). This discrepancy may have several sources. Children already receiving special education services, particularly those with ADHD, are overrepresented among those injured, and they may retain their original etiology’s label; children may “grow into” their educational disability without the educational system recognizing their prior injury; and some school systems keep records by educational needs rather than etiology (212,213). Policy recommendations for improving special education services for this population have been published (213).

Problems with attention and concentration appear to be among the most frequent sequelae of TBI in children, as they are in adults, occurring in up to 50% of survivors, although as many as 19% of those injured carried a premorbid label of ADHD (5,36). Learning and memory difficulties are also relatively common in those with severe injuries (5). Language impairments have also been reported in children injured in the preschool years (214). In children with severe TBI, injury to the frontal lobes may result in executive dysfunction (215). In general, many children with mild injuries make a good recovery. Children with severe injuries show the most persistent cognitive impairments, with those injured prior to age 8 showing less recovery than those aged 8 to 12 years (5). This is not consistent with the premise that younger children show increased cerebral plasticity in the face of injury, but rather suggests there are periods of increased vulnerability (216). TBI may also have important effects on emotional adjustment in children and adolescents, with depression, dysthymia, and anxiety disorders among the most common new problems (36). In a study of college students with a history of mild TBI, intellectual impairments were not found, although the students with prior TBI had increased levels of subjective distress in several subscales of the Symptom Checklist-90—Revised (217).

Family function also has a significant influence on outcome following pediatric TBI (4), and there is also reason to be concerned about the impact of the child’s injury on the family. In particular, it has been suggested that siblings of children with TBI-related disability may experience adverse impact themselves, including greater psychopathology, depression, or reduced emotional well-being and more negative relationships (218).

The preponderance of data suggests that pediatric TBI is underreported and underrecognized in its importance. A high proportion of children with a history of TBI are likely to present with educational special needs, cognitive problems, and psychological/behavioral problems. The fact that some of these children had such problems premorbidly, coupled with the delay between injury and appearance of the problems, may contribute to their underrecognition. Thus, it appears crucial that children with significant TBI be identified and tracked so that the relevance of their TBI to future problems and needs does not go unnoticed.

**TBI in the Elderly**

As discussed previously, the incidence of TBI is high among the elderly, with up to 86% of TBIs in those over 65 due to falls (219). Although mortality at a given level of severity is higher among the elderly than among younger individuals, there is also some evidence to suggest cognitive impairments are more severe and functional outcomes poorer, in those who are older (127,220). Considerable controversy exists about the prognosis for functional recovery among elderly survivors, due to the fact that many older patients with TBI were already suffering from functional decline at the time of the injury, and to the fact that those elderly patients admitted to rehabilitation facilities are a highly selected group. Research that avoided this selection bias by enrolling elderly patients in acute care, found that GOS scores of moderate disability, or good recovery never occurred in those with GCS scores less than 11 (221). This was true even among the subgroup with mild original injuries and subsequent deterioration to lower GCS scores. Though it appears likely that the functional prognosis for elderly patients is, indeed, worse than for younger ones, there is also a concern that this may contribute to a self-fulfilling prophecy, in which elderly survivors of TBI are not given the same rehabilitation opportunities as their younger counterparts.

Understanding the rehabilitation problems and needs of the elderly presents some of the same challenges as those seen in the pediatric population. In a retired population, what are the appropriate standards for community integration? If the normal developmental sequence is for an increasing prevalence of disability and need for assistance with aging, how does one assess the outcome of the rehabilitation process against this moving functional target? But, unlike children with TBI, whose parents are already in the role of caregiver, elderly survivors of TBI may have few support options. A spouse may already have died or be too frail to provide the necessary care, and adult children, if any, may be unwilling or unable to provide sufficient assistance. In the absence of family and community supports, therefore, a nursing home destination may seem to be a foregone conclusion, further eroding optimism about the cost effectiveness of aggressive rehabilitation interventions.

**MEDICAL PROBLEMS AFTER TBI**

**Fractures**

For patients with TBI who also have long bone fractures, open reduction and internal fixation can promote early mobilization,
simplify patient care, and improve predictability of fracture outcome. Early surgical intervention may reduce health care costs and delay in mobility, but there has been concern regarding risks for potential secondary brain injury due to surgery, including the risks of general anesthesia (222). A recent article compared outcomes of patients with TBI who underwent early surgery (<24 hours after injury) or late surgery (>24 hours), including both functional and neuropsychological measures. Early surgery was not associated with poorer outcomes (223).

While most fractures are diagnosed in the acute injury period, some are not detected until transfer to rehabilitation, because of increased movement, awareness, and communicative ability as the patient improves. Details regarding the initial trauma may increase the level of suspicion for a missed fracture. The incidence of missed fractures has been reported as high as 11% in this population (223). Any unexplained swelling, deformity, or pain response should prompt evaluation for occult fracture. Fractures in TBI predispose to HO. For example, despite appropriate operative management of 23 acetabular fractures, 61% had poor outcomes because of HO (224).

**Seizures**

Posttraumatic seizures and epilepsy are known complications of TBI. The risk of seizures is related to injury severity, presence of skull fractures, cortical contusion, subdural hematoma, and age (225). The risk of new seizure development is greatest in the first 2 years postinjury and gradually declines. In a population of patients who developed seizures in the first 2 years after injury, 33% of the seizures first occurred within the first month and 80% occurred within the first year (226). In another study, 86% of patients that had one late posttraumatic seizure (>7 days after injury) had one or more additional seizures within 2 years after injury, so the risk of recurrence can be high (227).

Most seizures are diagnosed clinically on the basis of focal or generalized motor activity. Patients with muscle spasms or tremors may present diagnostic dilemmas. In such cases, a routine or sleep-deprived EEG may reveal epileptiform activity. More definitive is a 24-hour EEG correlated with observations of the suspicious activity. Seizures in limbic and association areas may lead only to altered behavior or states of consciousness, presenting further diagnostic challenges.

Treatment during the first week postinjury with phenytoin or valproic acid can be effective for reducing the incidence of early-onset seizures (228,229). Long-term seizure prophylaxis is not recommended because it has not been shown to be beneficial in preventing late-onset posttraumatic seizures (230).

Carbamazepine and valproic acid have been found to be relatively free of adverse cognitive effects; their superiority to phenytoin in this regard is debated (231,232). None of these drugs is free of cognitive and physical adverse effects, however (233). These older antiepileptic drugs (AEDs) are all hepatically metabolized. Some newer AEDs are renally metabolized, which may be an advantage in certain clinical situations. It is hoped that some of the newer AEDs such as levetiracetam may not have as many deleterious side effects, but they have not been as well studied as some of the older agents. Studies are ongoing to examine utility and safety in the TBI population (234).

There are no standard recommendations on duration of AED treatment. In view of the expense and the potential toxicity of anticonvulsants, most clinicians withdraw medications after 1 to 2 seizure-free years. Of patients with previous seizure disorders of mixed etiologies who had been seizure-free for 2 years, 35% relapsed after tapering of their anticonvulsants (235).

**Hydrocephalus**

Ventricular dilation occurs in up to 40% of patients with severe TBI and usually begins to appear within 2 weeks of injury. In most instances, ventriculomegaly results from diffuse atrophy or focal infarction of brain tissue (i.e., hydrocephalus ex vacuo) and, thus, is a sign of primary and secondary injury but not a syndrome requiring treatment. The less common communicating hydrocephalus is generally associated with abnormal CSF pressure dynamics, causes neural dysfunction, and warrants treatment.

Unfortunately, the classic symptom triad for the diagnosis of hydrocephalus—incontinence, gait disorder, and dementia—is of little help in severely disabled patients. Failure to improve or deterioration of cognitive or behavioral function should prompt assessment with a CT scan. Flattening of the cortical sulci and periventricular lucency may support the diagnosis of clinically important hydrocephalus, but this can be a challenging determination (236). At present, there is no single test that is felt to be the gold standard in determining whether hydrocephalus is clinically significant. A recent review of the literature looked at the value of prognostic tests for idiopathic normal pressure hydrocephalus (NPH). The tap test, where 40 to 50 mL of CSF is drained and the patient is subsequently assessed to evaluate clinical improvement, was considered specific but not sensitive. A CSF infusion test to determine outflow resistance was more sensitive and specific, while prolonged drainage of CSF from a lumbar catheter may have the highest predictive value (237). The more sensitive tests may carry a higher risk of complications, especially in less medically stable or cognitively compliant individuals, which would include many patients with TBI. Even with an accurate diagnosis of clinical hydrocephalus, the prognosis for improvement from shunting is uncertain. This may be partly because the patient has other cognitive and motor deficits unrelated to hydrocephalus. Mazzini et al. suggest that patients that demonstrate a clinical deterioration may be more likely to benefit from shunting (238).

The etiology of posttraumatic hydrocephalus appears to be multifactorial. One proposed cause is the change in hemodynamics, CSF hydrodynamics, and brain metabolism caused by the presence of a craniectomy defect (239). This is of interest because there has been much discussion in the literature regarding the optimal timing of cranioplasty for patients that have undergone craniectomies, and reported improvements in function after cranioplasty have been attributed to this mechanism (240).
Patients with ventricular shunts may experience complications due to shunt failure, infection, or over/under drainage. The latter complication has been lessened through use of programmable shunt valves. Though not problem free, they have improved the ability to make minor changes in CSF flow non-invasively (241). The changes associated with hydrocephalus and shunt failure may be subtle and it is critical that the entire rehabilitation team’s assessments of cognitive and behavioral fluctuations be taken into account in this evaluation.

Cardiac Abnormalities
Cardiac complications may appear acutely, particularly in the setting of multiple trauma, and may involve direct injury to cardiac muscle, vessels, or valves, as well as the generation of arrhythmias, any of which may lead to impaired perfusion and increased secondary brain injury. Because of the multitude of potential injuries, there is no gold standard for cardiac evaluation in the setting of blunt trauma (242).

Hypertension, tachycardia, and increased cardiac output in the acute postinjury period may result from the increased release of epinephrine and norepinephrine (243). Central sympathetic hyperactivity can lead to ongoing myocardial necrosis (244). Although β-blockers may be considered, some medications in this class, such as propranolol, have been shown to cause cognitive impairments in hypertensive patients (245). It is unclear whether highly polar β-blockers such as atenolol or nadolol have fewer side effects in this population due to decreased ability to cross the blood-brain barrier.

Autonomic disturbances after brain injury may lead to dramatic increases in blood pressure and heart rate. These disturbances have been known by a number of names. More recently, the term paroxysmal autonomic instability with dystonia (PAID) has been coined (246). It is suggested that the diagnosis of this syndrome should be made when at least five of seven symptoms are present (fever, hypertension, hyperhidrosis, tachypnea, tachycardia, posturing, and dystonia). The etiology and treatment of this syndrome remains controversial. It has been suggested that environmental factors such as noxious stimuli may play a role in triggering the dysautonomia (247).

Pulmonary Disorders
Multiple trauma often causes pneumothoraces, pulmonary contusions, and lacerations. In addition, adult respiratory distress syndrome, excessive fluid administration during resuscitation, and intense α-adrenergic outflow may lead to noncardiogenic pulmonary edema (248). These and other problems can further compromise cerebral oxygenation.

Many brain-injured patients require tracheostomies for ventilation and suctioning. Humidified air may be delivered through a tracheostomy collar to maintain moist secretions and prevent tracheitis, and frequent suctioning may be required. Patients who initially require supplemental oxygen usually can be weaned from it by assessing pulse oximetry on oxygen and again after it has been stopped. Patients who require long-term tracheostomies but have begun to vocalize may benefit from one of the varieties of tubes that permit vocalization.

Most TBI patients eventually can be decannulated. Indirect laryngoscopy screening can be used to check for adequate vocal cord abduction and to rule out subglottic stenosis (249). In some instances, tracheal or subglottic stenosis will prevent decannulation and will require dilation or surgical management. If there is no sign of anatomic obstruction, a small-caliber tracheostomy tube will allow air to bypass the tube while it is plugged for progressive intervals. The patient is checked at the end of each interval or at any sign of distress with pulse oximetry. Once the patient has tolerated 24 hours of plugging without incident, decannulation can take place, and the tracheostomy stoma can be covered by a gauze pad or occlusive dressing until it heals. Some patients will have difficulty with decannulation because of the quantity of secretions and inability to cough them up into the pharynx. Thus, persistent suctioning is required. The tracheostomy tube itself may be an irritant that evokes secretions, however. To evaluate this possibility, a tracheal button (i.e., a small plug that keeps the stoma open) can replace the tube temporarily to see if the elimination of the tube allows patients to manage their own secretions; if not, the tube can be replaced.

Long-term survivors of severe TBI have been reported to show decreased lung capacity, vital capacity, and forced expiratory volume. The etiology of these abnormalities is not entirely clear but appears to be a combination of muscle weakness and incoordination, decreased chest compliance, and deconditioning (250). Guidelines for cardiopulmonary conditioning specifically for patients with brain injury have not been clearly established.

Pulmonary embolism is a potentially fatal end result of venous thromboembolism. Because persons with TBI who are hospitalized have many risk factors, it is not surprising that the incidence of deep vein thrombosis (DVT) is elevated in this population (251). The problem is compounded by a reluctance to use pharmacological prophylaxis because of the risk of intracerebral bleeding. The literature is too limited to allow development of a treatment standard (252). A recent survey of a group of TBI rehabilitation centers failed to identify a consensus regarding prophylaxis or screening (253).

Hypothalamic and Endocrine Dysfunction
Hypothalamic-pituitary disorders after TBI may be more common than previously thought and can lead to serious medical complications (254). Any or all of the hormones that make up the hypothalamopituitary axis may be affected. Severity of TBI, as measured by the GCS, is an important risk factor. The presence of DAI and basal skull fracture is also a risk factor for pituitary dysfunction (255). Injury may result from direct trauma, hemorrhage, ischemia, or interruption of blood supply due to increased ICP or edema (255,256). Elevation of cortisol is related to increased ICP but only in the presence of an intact brainstem, and is independent of ACTH secretion (257). Onset of hypopituitarism is usually fairly acute, with one study failing to identify any new cases beyond 6 months posttrauma (258).
Cranial Nerve Dysfunction

Olfactory nerve injuries (CN I) may accompany significant head trauma because of the delicate anatomy of the fibers exiting the cribriform plate. They are particularly likely in association with CSF rhinorrhea or after trauma to the frontal or occipital regions. The diagnosis may be missed by incomplete sensory testing including the inability of the patient to participate in testing. Objective testing of olfaction using a scratch- and-sniff odor panel is possible. Olfactory deficits may play a role in altered feeding behavior. People with anosmia are at increased risk of injury, as the sense of smell can warn one of potential dangers such as hazardous chemicals or fire.

Visual impairment may occur as spotty scotomata that differ in the two eyes, as homonymous hemianopsia, or as complete blindness. Optic nerve lesions (CN II) must be distinguished from hemisattention, cortical blindness, and visual agnosia. Early assessment can be done in gross terms through funduscopic examination, visual evoked response studies, and pupillary assessment. Vision and visual attention can be evaluated in severely impaired patients, using repeated presentations of colorful photographs and cards in each hemifield (173). When cooperation allows, more precise visual field and acuity testing should be performed. Neuropsychological assessment may clarify visual perceptual disorders. Information about field cuts and acuity should be shared with all team members to ensure that therapy materials are of appropriate size and placement.

Extraocular movements may be affected by damage to cranial nerves III, IV, and VI, their brainstem nuclei, or by impairment of coordinative structures in the midbrain and cerebellum (269). In addition, orbital fractures or damage to the extraocular muscles may produce concomitant gaze. Neuroophthalmologic examination may help clarify the pathophysiology of the deficit. Alternating or unilateral eye patching can eliminate diplopia. Prisms and strabismus surgery require that the patient be capable of ocular convergence to maintain binocular vision, but surgery sometimes may be appropriate solely for cosmetic purposes. CN III also contains parasympathetic fibers that control the pupillary response. This is particularly important as loss of the pupillary response (“blown pupil”) may signify uncal herniation.

Temporal bone fractures may disrupt facial nerve function (CN VII) temporarily or permanently. This must be differentiated from cortical or subcortical central facial weakness. Facial weakness, when combined with corneal insensitivity as a result of trigeminal lesions, can lead to corneal ulceration; corneal lubricants or tarsorrhaphy may be indicated.

Skull fractures also may disrupt the auditory or vestibular pathways (including CN VIII). Ossicular dislocation can lead to conductive hearing loss. BAERs can be helpful in the assessment of auditory function in unconscious or uncooperative patients; however, if no response is seen at normal intensity, hearing threshold should be evaluated with increasing

---

Chronic endocrine alterations also occur after TBI. An autopsy study of 100 brain-injured patients revealed a 62% incidence of pituitary injury (259); hypothalamic injuries coexisted in some patients. In addition, there are prominent axonal connections between orbitofrontal cortex and the hypothalamus, providing a potential mechanism for endocrine derangement without overt damage to the hypothalamus or pituitary (260). The few studies of hormone levels in survivors of TBI have found abnormalities in at least one hormone, ranging from 36% to 69% after severe TBI (261,262). Research to date has used laboratory definitions of endocrine abnormalities. Thus, the actual health impact of these derangements is unclear.

Deficiencies in growth hormone (GH) and the gonadotropins appear to be the most common anterior pituitary problems after TBI, followed by adrenocorticotropic hormone (ACTH) and thyroid-stimulating hormone (TSH) (254). The most significant posterior pituitary disorder, diabetes insipidus, has been reported to have a prevalence of 26% in acute care (258). It should be noted that the estimated prevalence of hormonal deficiency depends on which laboratory assessment method is used. More important is the determination of when to initiate hormone replacement when a deficiency is identified. Treatment of diabetes insipidus and ACTH deficiency should be started as soon as a deficiency is identified because of the potentially severe clinical consequences (263). Loss of adequate antidiuretic hormone production can lead to life-threatening hypernatremia and hypovolemia, and deficiency of ACTH may lead to life-threatening adrenal crisis, including hyponatremia, myopathy, hypotension, and hypoglycemia. In the context of panhypopituitarism, the only hormone for which immediate replacement is not recommended is GH. This is because GH levels may normalize after other hormone deficiencies are corrected (264).

Though less dramatic, deficiencies of other hormones may lead to significant morbidity. Of particular difficulty in the TBI population is the fact that clinical signs and symptoms of deficiencies of many of the pituitary hormones mimic neurological problems often seen as a result of brain trauma. These include weakness, fatigue, cognitive deficits, decreased energy and depression (265,266). More specific consequences of hormone deficiency include: weight gain, changes in skin and hair (TSH), and sexual dysfunction and amenorrhea (gonadotropins). While amenorrhea may be transient and related to the stress of trauma, early hormone replacement therapy for hypogonadism is suggested at least for patients with multiple endocrine deficits. Persistent amenorrhea in the absence of other endocrine abnormalities may require further evaluation and, as appropriate, hormonal therapy.

Although most persons with TBI who require hormone replacement therapy have life-long needs, improvement over time is sometimes seen, so there is a need for periodic reassessment (267,268). Needs may also change during periods of medical illness or stress. In addition to traditional clinical testing, determinations regarding dosing of hormones and evaluation of efficacy of treatment can be aided by measuring clinical elements such as fatigue, depression, quality of life, and cognitive function (263).
intensities. Caloric testing can provide information about vestibular function. Later, standard audiologic evaluation, including speech perception, should be performed. Hearing aids or ossicular repair can be helpful in selected patients. The patient's auditory capabilities should be shared with the entire treatment team so that care can be taken to communicate in the environment and style that optimize comprehension.

Movements of the tongue, pharynx, and larynx often are impaired by severe TBI at various levels in the central nervous system (270). Evaluation of oral motor function including cough and gag reflexes and indirect laryngoscopy may help clarify the location and extent of the problem. Such disturbances will be relevant for safe oral feeding as well as vocal communication.

**Sensory Deficits**

TBI can produce disturbances of any of the sensory modalities. Depending on the location and severity of the damage, these may be disturbances of basic sensation (e.g., decreased visual acuity) or of perceptual processing of that sensation (e.g., impaired visual spatial perception). Both forms of disturbance can be disabling.

Disorders of somesthetic sensation can result from damage to a variety of brain structures and may distort touch, pain, temperature, and position information. Pain syndromes may result from central injury to the thalamus, other cerebral structures, or from deranged neurotransmission between the thalamus and cerebral cortex (271).

Basic sensation can be assessed in each modality as soon as the patient is able to cooperate. Visual, auditory, and somesthetic sensory pathways can be assessed grossly earlier with evoked potential studies. The presence of normal basic sensation, however, does not ensure that the patient can form the complex perceptions needed to recognize a visually presented object, or identify a letter through stereognosis.

When patients fail at a task for which basic sensory input is intact, other causes of inability to perform can be explored. A patient recovering from cortical blindness, for example, may have full visual fields and functional acuity but may be unable to name visually presented objects. Neuropsychological assessment can help to differentiate among object-naming disorders, visual agnosias, scanning disorders, and disorders of complex visual perception (272).

There is no uniform strategy for coping with sensory deficits. Rather, treatment is planned around the severity of the deficit, strength of remaining sensory capabilities, and cognitive status of the patient. All team members should be involved in this planning because the patient's sensory, motor, and cognitive capacities all are relevant to the process.

**Heterotopic Ossification**

Clinically significant HO occurs in 11% to 20% of severely injured patients, primarily in proximal joints of the upper and the lower extremities (273). The etiology of HO in brain injury is unknown. Risk factors include prolonged coma, increased muscle tone and limited movement in the involved extremity, and associated fractures (274). A recent article also identified the presence of autonomic dysregulation as predictive of the development of HO (275).

HO may present with pain, warmth, swelling, and contracture formation but may be occult.Earliest diagnosis is possible with a bone scan, but ossification subsequently is visible on plain radiographs. Diphosphonates and nonsteroidal anti-inflammatory agents, particularly indomethacin, and radiation therapy have been used in an attempt to arrest early HO or to prevent its postoperative recurrence (273). A study of patients with early and intermediate HO found that calcium deposition occurred in the early phase, inflammation in the intermediate phase, and appearance of osteoclasts and osteoblasts in the late phase (276). This suggests that diphosphonates may be beneficial early and anti-inflammatory medication at an intermediate time, but no controlled treatment comparisons have been performed to date. A recent Cochrane review of articles that addressed acute pharmacological intervention failed to identify enough compelling evidence to support the use of medications to treat acute HO (277). Radiation has been used for prophylaxis, to treat existing HO (278) and after surgery to prevent recurrence (273). Long-term adverse effects have not yet been studied, particularly in young patients. Range-of-motion exercises are indicated to prevent ankylosis. If ankylosis seems inevitable despite exercises, it should be encouraged to occur in the most functional position.

Surgery for removal of ectopic bone should be undertaken only for clear functional goals, such as improved standing posture or ambulation or independent dressing and feeding. In general, surgery is not undertaken earlier than 18 months after injury (279).

**Increased Muscle Tone and Contractures**

Increased muscle tone following TBI may reflect true spasticity (i.e., increased phasic stretch reflexes). Dystonia, posturing in response to head position and cutaneous stimulation, and extrapyramidal syndromes also are common. The mere existence of abnormal tone is not an indication for treatment. Treatment should be based on functional considerations.

**Motor Disturbances**

A variety of other motor disturbances result from TBI. The diffuse nature of many injuries makes it difficult to tie patterns of motor dysfunction to specific anatomical structures. Deficits may include paralysis or paresis involving isolated muscle groups or a single limb, combinations of limbs, or the whole body, and may be caused by cortical lesions, disruption of subcortical white matter tracts, or by lower motor neuron lesions. Critical illness polyneuropathy and myopathy may be seen in more severely injured patients (280).

Disorders of balance and coordination may result from damage to the cerebellum or its connections. Patients with good muscle strength may be unable to ambulate or even sit independently because of profound ataxia. Similarly, limb ataxia may preclude self-feeding, writing, and independent ADLs.
Tremors are reported to be the most common movement disorder after severe TBI, with 5.4% of patients in one study having tremors considered to be disabling (281). Parkinsonism may be seen after a single TBI (usually severe or involving direct injury to the substantia nigra) or after repeated brain injuries as in boxers. While multiple medications have been used to treat these movement disorders, the degree of response is often disappointing.

The patient’s movement abilities should be assessed by the physiatrist, physical therapist, and occupational therapist. Formal gait analysis may help reveal the specific impairments affecting gait (282), and may suggest orthotic alterations or facilitate selection of muscles for targeting with antispasticity or other treatments. Weakness can be addressed through active assistive range-of-motion and progressive resistive exercises. Ataxia is notoriously difficult to treat. A weighted walker or wrist cuffs may be of modest benefit. If ataxia varies markedly from proximal to distal joints, selective splinting or strategies to stabilize the extremity may be useful.

An attempt should be made to diagnose tremors specifically with reference to their frequency, amplitude, and occurrence during action, rest, and sleep (283). L-dopa and pramipexole have been used in movement disorders such as tremors and ataxia, but functional tests should be performed on and off medication to assess efficacy in individual patients (284). Some improvement in tremor-related disability may be seen with botulinum toxin injections (285). Stereotactic surgery with radiofrequency ablation may be an alternative, but is recommended only for those with severe, disabling tremor (286). The significant risk of adverse side effects in the TBI population must be taken into account when considering this intervention. There are case reports of benefit from deep brain stimulation for posttraumatic movement disorders (e.g., (287)), but there is uncertainty about the anatomic target for stimulation, when to consider this intervention and who may be most likely to benefit. Orthoses and adaptive devices may assist weak or ataxic patients in performing functional tasks. Dexterity exercises may be used to increase manual speed and coordination.

Slowed motor responses are among the most common deficits associated with TBI and may limit self-care and employment productivity. It appears, however, that much of this motor slowness can be attributed to central processing delays rather than a motor deficit per se (288). There is no consensus on how to treat this generalized slowness, although there is evidence the methylphenidate improves processing speed (55,289).

Dysarthria sometimes is noted after brain injury and may be the result of either peripheral or central nervous system damage affecting motor control of the speech mechanism. The dysarthric patient may receive oral motor strengthening and breath support training to improve sustained volume. Kinematic articulography may provide more objective data regarding the nature of dysarthria and better target the area of focus for treatment (290). A recent Cochrane review did not identify any studies of sufficient rigor to support or refute the utility of therapy for dysarthria (291).

A growing body of literature supports early mobilization to aid recovery both peripherally and centrally (292). A number of automated systems for guiding motor-related therapies are in development. These computerized devices deliver movement exercises via various forms of virtual reality (VR), and make use of displacement and/or force sensors to monitor performance (293).

Nutrition and Feeding

TBI can produce dramatic increases in basal metabolism, with catabolism, weight loss, and low serum albumin. Thus, the acutely injured patient must have increased caloric and protein intake, but the need for invasive procedures requiring an empty stomach, and the difficulty in achieving high rates of nasogastric tube feeding may interfere. While it is recommended that full nutritional replacement be instituted within the first week of injury (294), there is no consensus on how soon to consider gastrostomy or jejunostomy placement in patients who remain obtunded or lack oral motor capability (295,296). The availability of access for nutrition is not the only challenge. The gastrointestinal system must be able to move the food through the tract and absorb nutrients. In the TBI population, concern about delayed gastric emptying may be overstated (297). Gastroparesis may be a problem, especially acutely, but the medication most frequently used (metoclopramide) has the potential of slowing recovery (298), although it may be useful in decreasing reflux by increasing gastroesophageal sphincter tone (299).

Obstacles to oral feeding include both cognitive and oral motor deficits. Neuroleptic and anticholinergic medications may contribute to poor oral motor control. Attempts at oral feeding generally begin with assessment of the reflexive and voluntary components of oral motor function by a speech, occupational, or physical therapist experienced in dysphagia treatment. This is followed, if necessary, by an oral motor facilitation treatment program to decrease the latency of the swallowing reflex and improve oral motor strength and coordination.

Patients who appear to be candidates for oral feeding but who have frequent coughing or an episode of clinical aspiration may benefit from videofluoroscopy performed collaboratively by an experienced feeding therapist and a radiologist. By using different food textures and placements and head positions, it may be possible to determine a strategy to normalize swallowing. Fiber-optic endoscopy has also been found to be useful in evaluating patients with dysphagia after TBI (300). When the patient has made maximal progress in terms of food quantity and texture, the patient or family and nursing staff should be trained in the optimal oral feeding methods. Patients need to be monitored closely to determine that they are meeting their nutritional needs, especially when transitioning from tube feeding to oral intake. Altered textures, thickened liquids and injuries to the sensation of taste and smell may all affect appetite. Physical and cognitive limitations, side effects of medications, and depression may also decrease intake. Serum electrolytes and prealbumin are objective measures of hydration...
and nutrition. When necessary, pharmacological interventions may be appropriate to increase appetite (e.g., megestrol acetate, oxandrolone, and dronabinol) (301).

**Bowel and Bladder Dysfunction**

Frontal lobe lesions can impair inhibitory control over bowel and bladder evacuation, leading to urgency and incontinence. Detrusor hyperreflexia has been reported to be the predominant type of bladder dysfunction (302). Impaired mobility or dexterity, inability to communicate toileting needs, impaired initiation, cognition or behavioral deficits also contribute indirectly to incontinence, making this issue a significant one in persons with brain injury.

**Pain**

Pain is a complex phenomenon that can complicate the rehabilitation process regardless of injury severity, and can interfere with physical activity, cognitive performance, and sleep, as well as contribute to behavioral disturbances. There are numerous etiologies for pain in this population. When TBI occurs in the context of multisystem trauma, it may be related to fractures and traumatic neuropathies and plexopathies. Complex regional pain syndrome Types I and II are also seen in this population. Other neuro-orthopedic complications such as contractures, hypertonia, and HO may generate noxious stimuli (303). Pain may also result from a host of other coexisting medical conditions commonly seen after TBI, involving gastrointestinal, genitourinary, pulmonary, renal and cardiac sources. The most appropriate treatment depends on an assessment of the source and nature of the pain. In patients who are able to describe their symptoms, assessment is similar to such assessment in non–brain-injured individuals. In minimally conscious patients, however, it may be difficult to distinguish pain syndromes from spontaneous moaning or grimacing that may have no painful etiology. Moreover, more interactive patients with memory problems may have difficulty reporting the frequency, pattern, and precipitating causes of their pain. In such instances, it may be helpful to record the behavioral manifestations (e.g., grimacing) and their relationship to potential provoking factors (e.g., ranging of a particular limb).

The range of treatment modalities available to the general population can also be applied to patients with TBI who are experiencing pain. In the setting of multiple trauma such as orthopedic injuries or in the postoperative period, narcotic analgesics may be most appropriate. Ideally, these medications should be given in standing doses or prior to therapies that incite pain rather than “as needed” since this provides better pain relief and avoids inadvertently reinforcing pain behaviors. In all cases, but especially for medications with more sedating properties, the need for appropriate pain relief must be balanced by consideration of possible slowing of the rehabilitative process (304). Nonsteroidal anti-inflammatory medications may decrease the need for narcotic medications, but have their own side effects that need to be considered. Alternative drug delivery systems, such as intrathecal infusion, may provide pain relief while decreasing the side effects compared with enteral delivery, particularly when the patient is already being treated for spasticity with medication via an intrathecal pump.

A number of medications have been used to treat neuropathic pain, but most formal studies have involved patients with diabetic neuropathy or complex regional pain syndrome. Outcomes and side effects have not been well studied in the TBI population, so use may be governed by evaluating potential side effects in the specific clinical setting and determining efficacy once a medication is chosen. Antiepileptic medications, tricyclic antidepressants, and venlafaxine have demonstrated efficacy, whereas serotonergic antidepressants have been insufficiently studied (305,306). In addition to medications, there may be a role for physical modalities or sympathetic nerve blocks.

**Headache**

Headache is perhaps the most common pain complaint after TBI. Interestingly, there is some evidence to suggest that it is less common among persons with more severe TBI (307). Most posttraumatic headaches (PTH) resolve within a few months although 18% to 33% of headaches may persist beyond 1 year (308). Although most headaches are not due to serious intracranial pathology, one should at least consider problems such as intracranial hemorrhage, hydrocephalus, CNS infection, or another head injury. The identification of other new neurological findings may be helpful in identifying such complications, especially in a patient with cognitive or language deficits who may not be able to provide an accurate history or description.

There are many types of PTH, and complete review is beyond the scope of this chapter. It is helpful to be aware of some broad categories of PTH, as an accurate diagnosis should lead to greater success in determining an effective treatment strategy. Musculoskeletal headaches can be broadly classified as headaches that are related to irritation of muscles, ligaments, tendons and joints such as facet joints. Interventions for such headaches may include trigger point injections, muscle stretching and strengthening, physical modalities, and medications, although it is not clear how efficacious these strategies are (309–311).

Headaches may also result from injury to nerves in the scalp damaged by direct trauma or surgery. These headaches are often described as lancinating or shooting, and may include allodynia. Headache due to occipital neuralgia is often reproducible when the nerve is palpated, leading to a characteristic referred frontal or retro-orbital pain. Diagnosis and treatment may be provided by an occipital nerve block (312). Pharmacological interventions may include medications employed for other neuropathic pain disorders. Interventions for posttraumatic migraines and posttraumatic tension headaches also generally parallel treatments seen in their corresponding nontraumatic conditions. However, one must take into account the potential cognitive effects of these medications.

**Sleep Disturbance**

Sleep disturbance after TBI is quite common and can interfere with rehabilitation and adversely affect cognition (313).
Incidence of sleep disturbance has ranged from 30% to 70%, depending on time postinjury and defining criteria (314), with highest incidence acutely but a substantial frequency of chronic sleep problems. The initial trauma may lead to injury to neuronal structures known to be important in the regulation of sleep, such as the reticular activating system (RAS). Acutely, the environment, medications, pain, stressors, cognitive deficits, and behavioral problems may all contribute to the problem, but maladaptive behaviors and beliefs may develop and lead to persistence of the sleep problems (314). Specific sleep disorders such as obstructive sleep apnea and restless leg syndrome may have existed prior to the TBI, whether known or unknown, and may also require treatment.

Nonpharmacological management of insomnia includes environmental manipulation, relaxation techniques, and behavioral therapy. A number of medications have also been utilized to treat insomnia such as benzodiazepines, antidepressants (especially tricyclic antidepressants), and nonbenzodiazepine sedatives. All these medications have potential cognitive side effects, which must be considered in the setting of brain injury rehabilitation. Whenever sleep disturbances are suspected and treatment initiated, it is important to have objective measures to assess efficacy of the intervention(s) selected, since self-reporting may be inaccurate.

Fatigue is a closely related and among the most common complications of TBI (315,316). Like insomnia, causes of fatigue are multifactorial. There may be primary injury to neural structures such as the ascending RAS, which is very important in maintaining arousal (317). Endocrine dysfunction (266) and depression (318) can also contribute, but many patients report fatigue without an identifiable cause. Treatment is primarily focused on the identification of precipitating factors, and use of compensatory strategies such as pacing and scheduling of daily activities. Medications, especially stimulants such as methylphenidate and dextroamphetamine, have been used with mixed results. A recent double-blind, randomized, placebo-controlled trial of modafinil, a newer wakefulness-promoting medication, failed to detect a clinically significant effect (319).

**Impairments of Arousal and Attention**

Deficits in arousal and attention are among the most widespread after TBI (288). The comatose patient suffers from profoundly impaired arousal, and many of the cognitive and emotional complaints in minor TBI are hypothesized to be attentional in nature (288). Arousal may be defined as the general state of responsiveness to environmental stimuli. Normally, arousal undergoes slow fluctuations in relation to diurnal rhythm, food intake, and activity level (i.e., tonic arousal). Arousal also can be modulated over brief intervals by demands in the environment (i.e., phasic arousal) (321). The RAS has a primary role in control of arousal, exerting its influence over diffuse cortical regions and receiving cortical inputs in return (322). Damage to the RAS plays a critical part in coma onset (50). In principle, a deficit in tonic arousal could lead to generalized impairments in responsiveness and profound slowing of information processing. Impaired phasic arousal could impair performance in cognitively demanding situations.

Attention may be considered to be the selective channeling of arousal. Attention (associated with conscious awareness) is directed to a particular set of internal or external stimuli out of the infinite set of possible targets of awareness. Attention is not a unitary phenomenon either in psychological or neurophysiological terms, and there is no precise agreement about how to divide up its component processes; at the least, the following phenomena can be distinguished (323):

- **Arousal:** the state of receptivity to sensory information and readiness to respond.
- **Selection** (sometimes referred to as **focused attention**): the ability to focus attention on particular stimuli or responses. Selection is often made on spatial grounds (i.e., selecting a stimulus in a certain location), and damage, particularly to right parietal or frontal lobes, can interfere with normal spatial selective attention (hemispatial neglect).

**NEUROCOGNITIVE AND NEUROBEHAVIORAL DEFICITS**

There is great variability in the patterns of cognitive and behavioral impairments seen following TBI. Despite this, there are certain commonalities across patients, most likely related to the typical areas of injury to the gray matter (i.e., frontal and temporal poles) and white matter (i.e., midbrain and corpus callosum). Thus, difficulty with alertness, attention, memory, planning and problem solving and abstract reasoning is very common, as are behavioral disturbances related to impulsivity, impaired initiation, and other problems of behavioral control. In the sections that follow, we review some of the most common cognitive and behavioral impairments affecting persons with TBI. Although we discuss problems separately, difficulties labeled by clinicians and family members as “behavior problems” are often highly influenced by cognitive and perceptual disturbances. For example, comprehension deficits or slowed information processing can induce brain-injured persons to act out in frustration when their capacities are overtaxed. Variations in behavioral response may also be related to preinjury factors including age, socioeconomic status, substance use history, and personality differences. The totality of subjective experiences and personal history must be considered when evaluating problems and developing treatments for them.

Neurocognitive and neurobehavioral status may be assessed with standardized tests, as in a neuropsychological evaluation, and/ or by structured observations in real-world situations. The latter are more useful for clarifying the impact of cognitive impairments in persons who are low functioning or when status is likely to fluctuate (e.g., in the early stages of recovery). Formal neuropsychological assessment is beyond the scope of this chapter but is discussed extensively in the textbook by Lezak et al. (320).
• Strategic (executive) control: the ability to sustain attention over time; inhibit disruption by distracting influences; shift attention in line with changing goals and priorities; manipulate information currently held in mind (referred to as working memory); and divide attention between two or more task demands.

• Processing speed: the speed at which information is transmitted within the nervous system to allow cognitive processing to occur.

Numerous brain regions appear to play a role in attentional processes. Research has particularly highlighted the roles of the right hemisphere, bilateral prefrontal and parietal cortices, the anterior cingulate, and portions of the thalamus and basal ganglia in attentional function (324). Disorders of attention may impair learning and performance in several ways. Affected patients may have difficulty focusing on any task, may be easily distracted, or may show hemispatial neglect. Attention deficits can lead to secondary decreases in language comprehension or visuospatial function because the patient’s information processing is interrupted and disorganized. Inability to shift attention might be manifested, for example, in the inability to attend to task instructions and also monitor one’s own performance. It has been suggested that PTA is most appropriately considered a confusional state related to attention deficits rather than a true memory disturbance (325). Similarly, frontal lobe disorders such as impulsivity and perseveration may be interpreted in attentional terms, as loss of goal-directed control over attention such that it is easily pulled or dysfunctionally fixed on irrelevant aspects of a task (326).

Exactly which of these aspects of attention are disrupted in TBI remains a subject of some controversy. There is general agreement that processing speed is reduced in TBI, but the location of this slowing in the stream of information processing is still under debate. Laboratory research shows that patients with TBI have difficulties with sustained attention (327) but that phasic arousal to auditory stimuli is preserved even in severe injury (321). Tasks requiring divided attention appear to be particularly impaired by TBI, when both tasks require conscious control (328,329). When performing independent work in distracting environments, individuals with TBI have more off-task behavior than uninjured controls both in the presence of distractions and in their absence (330). On the basis of the many attentional studies in TBI, it appears that, other than slowed processing, most attentional complaints can be linked to strategic control of attention rather than basic arousal or selection. Strategic control, in turn, is believed to be dependent on prefrontal and limbic goals and motivations being linked to basic attentional mechanisms.

Arousal and attention are assessed best by a combination of formal neuropsychological tests and behavioral observation by all disciplines (323). It is critical to arrive at more precise diagnoses than “impaired attention” because impairments of different components of attention may have different therapeutic implications. Unfortunately, there is no consensus on what set of neuropsychological tests are most informative about attention impairments and, in particular, most useful in differential assessment. Digit Symbol and the Symbol Digit Modalities tests are sensitive to impaired information processing speed (104). Digit span forward and backward and other tests of working memory and mental control may also be useful (320,331). In less impaired individuals, tests such as the PASAT (322) and Test of Everyday Attention (333) may provide additional insight. The Sustained Attention to Response Task (SART) (334) appears to be sensitive to impairments in sustained strategic control over attention (104). In general, treatment of these disorders can be grouped in terms of pharmacologic, behavioral, and compensatory strategies (Table 24-4). Attention can also be assessed with observational rating scales such as the Moss Attention Rating Scale (MARS), which is particularly useful for patients who are too impaired for formal testing (335).

### TABLE 24.4 Treatment of Disorders of Arousal and Attention

<table>
<thead>
<tr>
<th>Component</th>
<th>Pharmacologic</th>
<th>Behavioral</th>
<th>Compensatory Strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tonic arousal</td>
<td>Methylphenidate&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Naps</td>
<td>Engage in tasks when most alert</td>
</tr>
<tr>
<td>Phasic arousal</td>
<td>Amphetamine</td>
<td>Upright position</td>
<td></td>
</tr>
<tr>
<td>Selective attention</td>
<td>Bromocriptine</td>
<td>Frequent task changes</td>
<td>Give alerting clues</td>
</tr>
<tr>
<td>Hemispatial neglect</td>
<td>Nondistirating tricyclics</td>
<td>Graded training to attend left; prism adaptation training</td>
<td>Nondisturbing environment</td>
</tr>
<tr>
<td>Strategic control</td>
<td>Amantadine</td>
<td>Training in hierarchic attention skills</td>
<td>Position tasks to the patient’s right</td>
</tr>
<tr>
<td>Processing speed</td>
<td>Methylphenidate</td>
<td>Reinforce rapid performance</td>
<td>Simplify decision making, provide supervision, train solutions to specific problems</td>
</tr>
</tbody>
</table>

<sup>a</sup>Note that most of these treatments are investigational.

<sup>b</sup>All these medications may have uses beyond regulation of arousal and act on other components of the attentional network. Their respective roles are under investigation.
Many medications used for other purposes have negative effects on arousal, attention, and general cognitive function. Anticonvulsants such as phenytoin and phenobarbital, antihypertensives such as methyldopa and propranolol, and antispasmodic drugs such as diazepam, baclofen, and dantrolene all may impair cognitive performance (336–339). Therefore, attempts should be made to withdraw such medications or replace them with less sedating alternatives (e.g., carbamazepine, ACE inhibitors, phenol or botulinum toxin nerve/muscle blocks) (340).

The role of pharmacologic treatment remains controversial. Some studies have suggested attentional benefits from psychostimulants such as methylphenidate but are subject to methodologic criticisms (small numbers of subjects, inadequate control for spontaneous recovery, and failure to specify which attentional components are being assessed) (341). A randomized controlled study suggested that methylphenidate improves cognitive processing speed, and some aspects of sustained attention, and caregiver and clinician ratings of attentiveness, while having little effect on overt orienting to extraneous distracting events (289,342). Although methylphenidate has been reported to increase seizure risk (e.g., (343)) this is not well documented, and research suggests that, on the contrary, it may have some anticonvulsant effect (344). Other dopaminergic drugs such as amantadine and low-dose bromocriptine have been reported to improve overall attentiveness, increase the ability to divide attention (345), or diminish hemispatial neglect (346). Higher-dose bromocriptine, however, does not appear to have positive effects, as judged by a randomized placebo controlled pilot study (90). Of the many drugs that have been proposed as beneficial for attention deficits, a structured evidence-based review identified methylphenidate as the agent with the strongest evidence (347). Because of the heterogeneity of attention deficits, it appears unlikely that all patients will benefit from the same drug, but selection of appropriate patients remains to be clarified. Therefore, it is recommended that, when a patient is being considered for pharmacologic therapy of attention deficits, individualized measures of the behaviors of interest be assessed both on and off the medication (175).

Behavioral retraining of attention has been advocated but is of uncertain efficacy. Some studies have suggested benefit from retraining of patients with neglect or distractibility (348–352), whereas others have failed to document changes (353–355). Moreover, few of the studies that demonstrate improvement have assessed generalization of benefits to the everyday environment. A meta-analysis of behavioral treatments for attention deficits suggests that studies involving a prepost design without an untreated control group show far more improvement than those involving a comparison to a control group, indicating that much of the apparent “benefit” of attention retraining programs results simply from repeated administration of the outcome measures themselves, combined with residual spontaneous recovery. The studies that do demonstrate greater improvement in the treatment than the control group, tend to involve attention coaching in complex tasks such as driving and ADLs rather than retraining of specific attentional capacities in isolation (356). Together, these results suggest that therapists’ energy would be better spent giving feedback about attentiveness in important real-world tasks, than supervising repetitive drills of core attentional capacities. In addition, it is important to ensure that strategies learned in a therapy environment are carried over into a variety of functional activities. Therefore, if a team member is training spatial attention in, for example, reading, interdisciplinary plans should be made to transfer the same protocol to other settings such as ambulation and ADL tasks.

In uninjured people, extensive practice of virtually any task leads to the ability to perform it with minimal attention (i.e., automatically). This phenomenon has received little study in brain-damaged patients; however, it would seem possible to deal with some attentional impairments by practicing activities to such a degree that they no longer place much demand on disordered attentional processes. Recent work by Schmitter-Edgecombe and Beglinger supports this (357).

Impairments of Learning and Memory
Memory impairments are among the most common, persistent, and disabling of the cognitive complaints after TBI. All patients with moderate-to-severe injuries, and most with mild injuries, experience permanent amnesia for the events immediately preceding and following the injury. These intervals are referred to as retrograde and PTA, respectively. As patients recover from the acute confusional state postinjury, retrograde amnesia typically “shrinks” toward the present (358). The final interval unaccounted for ranges from minutes to weeks or months and, in very severe cases, may “eliminate” highly salient preinjury memories, such as the birth of a child or the death of a parent. PTA is the interval of permanently lost memory following the injury; it covers both coma and the acute confusional state and is nearly always longer than retrograde amnesia. As noted earlier, its duration is commonly used as a measure of injury severity.

These permanent gaps in memory are often emotionally disturbing to patients who must reconstruct the missing interval by obtaining reports from others. Retrograde and PTA also have legal ramifications, particularly when injuries are caused by assaults or reckless driving. Medical professionals may need to intervene with the legal system when persons with TBI are asked to testify about events they are unable to remember.

Most individuals with moderate and severe TBI suffer from ongoing deficits in anterograde memory (difficulty in storing and retrieving new information) well beyond the resolution of PTA. Because of the diffuse and multifocal nature of TBI, the learning and memory impairments typically affect multiple modalities (e.g., visual/spatial as well as verbal) and have a broad impact on “everyday memory” (359). The neural substrate of memory deficits after TBI is not fully understood. However, memory test scores have been shown to correlate with the degree of posttraumatic atrophy of the hippocampus, a limbic structure thought to play an important role in new learning (360). Given the prevalence of attentional deficits in TBI, it is also plausible that some “memory” impairment can
be attributed to defective processing of information at the time of presentation, although one study showed that challenging attentional resources at the time of stimulus did not disproportionately affect the later recall of TBI patients (361). There is evidence that when persons with TBI are given extra learning trials, their later retrieval of the learned material can be comparable to that of uninjured controls (362). However, other research suggests that even with extra learning experience, persons with TBI fail to take spontaneous advantage of retrieval cues (e.g., categories in a word list) which boost the retrieval scores of control subjects (363). These findings imply that at least some memory deficits may be due to problems with initial acquisition, which might be overcome with extra training or rehearsal of information; and that persons with TBI might need help to notice, and use, aspects of the material that will facilitate later retrieval.

The bulk of the research on memory impairment has concerned memory for past events or past learning, that is, retrospective memory. However, a very important aspect of real-world function is prospective memory, which is the ability to remember to do something in the future. Prospective memory may depend partly on frontal/executive control processes that help to maintain an active goal state over time and provide “triggers” for action at the appropriate moment (364). Prospective memory is significantly impaired in persons with TBI, whether the task is framed as something that must be done at a specific time, or done in association with another activity or event (365). Real-world prospective memory problems appear as a wide variety of “failures of intention” such as forgetting to run errands or take medications, missing appointments, and neglecting to provide information or messages to others. Of the many commercially available tests and test batteries for the assessment of memory, only a few include a measure of prospective memory (e.g., the Rivermead Behavioural Memory Test) (366). More recently, the Comprehensive Assessment of Prospective Memory (CAPM) has been developed (367).

In the clinical setting, evaluation of memory impairment should not rely exclusively on test results. A functional assessment by the interdisciplinary team and careful interviewing of the family or caretakers will help to identify the environmental situations in which the disorder disrupts function, the frequency with which real-world memory problems occur, and most importantly, the cues or conditions that facilitate memory performance. As noted above, it is also important to remember that other cognitive difficulties such as in attention, language or executive function, will affect memory performance and may be interpreted or described by the affected person and family as a “memory” problem. In such cases, the memory system may actually be relatively spared, and other difficulties must be sought as the root of poor or inconsistent performance.

An influential model of memory processing that has relevance to TBI rehabilitation highlights the distinction between implicit and explicit memory. Explicit memory refers to consciously remembered material. For example, remembering yesterday’s events, drawing a picture from memory, and recounting a conversation all involve explicit memory. In contrast, implicit memory operates outside conscious awareness and cannot be precisely verbalized; it is expressed via improved performance, for example, when new motor skills or other forms of procedural learning are acquired. These two memory systems, which interact constantly in normal learning, are thought to be subserved by different neuronal circuitry. Implicit memory may be diffusely represented in cortical networks, making it less vulnerable to cerebral injury than the phylogenetically newer explicit memory system, which is thought to be dependent on focal regions of the hippocampus and medial temporal lobes (368). There is evidence that in TBI, as in other organic conditions affecting memory, implicit learning is spared relative to explicit memory (369,370). A training method known as errorless learning has been studied in persons with TBI and other forms of organic memory impairment. Errorless learning (in which the trainer minimizes the errors committed by the trainee during the learning process) helps to bypass the need for explicit memory, which is thought to be critical for benefiting from error feedback. Without explicit memory, errors may inadvertently be learned by the more intact implicit memory system, interfering with performance. Errorless learning and related methods have been shown to be effective for helping people with TBI to learn new computer tasks, face-name associations, procedures for using memory notebooks, and other materials (371,372). The full range of tasks amenable to errorless learning, and its clinical utility in rehabilitation environments, require further exploration. Another widely used learning enhancement method, the spaced retrieval technique, involves prompting recall at gradually expanding intervals, and has been used with success, even when delivered over the telephone, to teach individuals with TBI to use external memory aids for prospective memory tasks (373).

Many other methods have been used in an attempt to remediate the memory problems experienced by persons with TBI. As there is little convincing evidence that memory can be enhanced pharmacologically (55), the most promising methods involve the teaching of compensatory strategies. For example, in a group of people with TBI who were living independently, training in the real-world application of memory enhancement strategies improved memory on objective tests as well as subjective ratings, and the improvement persisted at 4-month follow-up (374). For those with more severe memory deficits, a more commonly used approach is to train the use of reminder systems such as notebooks, planners, electronic organizers, or hand-held computers. The success of any reminder system depends partly on an individual’s awareness of memory deficits and acceptance of the need to use external strategies, both of which may be problematic. It will also be more successful if consistent with methods used prior to injury. In addition, persons with TBI may have difficulty with the executive demands of these strategies, that is, knowing what to record when, remembering to do it consistently, and using recorded information prospectively. Structured training programs in how to use compensatory memory systems may be worth the effort, however. For example, a comprehensive training regimen...
to teach individuals with TBI how to use memory notebooks showed that the training had beneficial effects on measures of everyday memory function (375). Similarly, improvement in prospective memory function, as measured by diary use, was found in a program focusing on self-awareness of memory difficulties, use of a customized compensatory tool, a cueing system, and organizational strategies (376).

For the rehabilitation setting, the following summary points should be considered:

- Many persons with TBI benefit more from hands-on training of procedural skills than from verbally based or didactic instruction, although a combination of modalities may be most effective.
- The learning of persons with TBI may need to be judged by changes in performance in addition to verbal self-report, as much learning can take place of which patients are unaware.
- Errorless learning can be effective for patients with severe explicit memory impairments. It may be necessary to provide cues at each step and hand-over-hand assistance (e.g., during the training of ADL) to prevent the patient from making errors. This assistance can be gradually faded as the patient begins to initiate more of the actions correctly.
- Persons with TBI may not spontaneously recognize or take advantage of cues or aspects of the material to be remembered that will assist later recall, so therapists may need to provide explicit help in this regard.
- Memory notebooks and other organizing systems are useful for many persons with TBI. Reminder systems are most effective when they are consistent with methods used prior to injury and when the entire rehabilitation team focuses on teaching only one or two applications (e.g., scheduling appointments, recording the contents of phone calls). This allows the team to achieve more consistency in teaching “how and when to record what.” New applications can be added as the user achieves mastery of the original functions.

Impairments of Frontal Executive Function

The mark of intelligent behavior is the ability to adapt to change, solve unexpected problems, anticipate outcomes, and otherwise cope with situations that fall outside of one’s routine. These abilities in turn require cognitive flexibility, self-monitoring, and self-adjustment of performance as well as simultaneous consideration of multiple alternatives and their probable consequences. From these alternatives, an appropriate plan of action must be selected and the intention carried to completion. Complex abilities such as these, which are necessary for adaptive human behavior, have been termed executive functions. There is considerable evidence linking these functions to the prefrontal cortex (364,377). TBI, which often involves damage to the frontal poles and orbitomesial cortex or their connections (378), commonly leads to disorders of reasoning, planning, and goal-directed behavior attributable to breakdowns in executive function. Depending on the nature and severity of the injury as well as premorbid factors, these impairments may appear clinically as disorganized or tangential speech and action, a tendency to be “pulled” by irrelevant stimuli, and inappropriate interpersonal behavior (379). Lack of insight into these and other deficits and poor ability to profit from feedback are usually part of the clinical picture.

Most conceptualizations of executive function include the notion of purposeful, controlled deployment of attention and other cognitive/perceptual functions in the service of goal-directed behavior unfolding over time (380). Normal behavior falls under external control at times, as when one attends “involuntarily” to a loud noise or moving object. Behavior must be modulated, however, in the service of one’s intentions. For example, if the goal is to get to an appointment in an unfamiliar place on time, one will be more likely to attend to street signs and less likely to attend to shop windows. For such modulation to occur, emotional and motivational states must be linked to perceptual and motor systems so that the appropriate aspects of the environment may be attended to and acted on, and the irrelevant ones screened out. The breakdown of this linkage contributes to the distractibility and aimlessness of many TBI patients. Even when performing so-called overlearned routine actions such as making a sandwich or cup of coffee, persons with TBI omit steps and make errors of planning, sequencing, and object use that can be attributed to failures in this attentional control of action (381). Compared to uninjured control subjects, those with frontal lobe damage due to TBI have difficulty generating “scripts” for action plans for familiar activities such as grocery shopping, and even more difficulty implementing such plans in real time (382). Deficits in planning and goal attainment have been emphasized as central to frontal executive function in the influential research of Duncan, who conceptualizes such impairment as “goal neglect” (383).

Another important aspect of executive function is control over one’s responses to people and situations, such that some behaviors are inhibited in favor of others that are more appropriate to the context. For example, to meet certain interpersonal needs, one must judge, then select, which of many possible behaviors will achieve the goal and fit with the social milieu, and use available social feedback to check the appropriateness of one’s behavior. Individuals with TBI often lack these judgments, behaving in ways that appear inappropriate and short sighted to others, and that seem to entail “personality changes.” Thus, a male patient may behave in contradiction to his preinjury interpersonal style by exhibiting a sexually aggressive manner with a female therapist, oblivious both to her reaction and to strong cues and admonitions from members of his own family.

Remaining questions to be addressed about executive function include questions about what and how many specific cognitive processes underlie these critical functions. For example, both working memory (384) and the formation of effective plans (383) have been proposed as central operations, deficits in which may explain a variety of dysexecutive behaviors. Related lines of research have examined the relationships among seemingly disparate behavioral symptoms often
categorized as executive function disorders, such as behavioral dyscontrol, personality change, and various higher cognitive impairments. These symptoms do appear related to one another, and it may turn out that executive functions are comprised of a relatively small number of correlated factors (385).

Clinical assessment of executive function deficits is difficult for several reasons. Partly because of the uncertainty in how to conceptualize these functions, there is simply less agreement as to how they should be measured compared to (for example) learning and memory abilities. In addition, by virtue of late phylogenetic and ontological development, frontal executive functions may show more interindividual variation among healthy controls than relatively hard-wired functions such as perception and language. Thus, two people could have identical prefrontal injuries yet display different patterns of executive dysfunction (379). Because executive disorders relate to the ability to deal with novelty and change, they are difficult to see within structured tasks but must be assessed in challenging or open-ended situations. Persons with TBI may do fairly well on traditional neuropsychological tests yet show profound life disruption from executive dysfunction (364). More recently developed test procedures now provide standardized analogues to the kinds of tasks that are sensitive to executive dysfunction. For example, the Behavioural Assessment of the Dysexecutive Syndrome (BADS) is a validated battery that includes tests of planning and prioritizing as well as a rating scale to capture personality and behavioral changes affecting daily life (371,386). However, recent studies have shown that the reliability of the individual subs tests and the total BADS profile score is poor (386,387). A measure which better captures everyday executive difficulties is the Multiple Errands Task, originally developed by Shallice and Burgess (388) and further developed by Alderman et al., into the MET-SV (389), with a second version adapted for use within hospital settings (MET-HV) (390). In addition, Schwartz et al. have developed the Naturalistic Action Test (NAT), a brief standardized assessment of the ability to plan and carry out everyday actions in real time (391).

Attempts to treat executive impairments have been based on one of three general perspectives (392). At one extreme, the deficits are considered to affect behavior in such a global way that they cannot be remediated. Therefore, the physical and social environment will need to be fully structured to act as “prosthetic frontal lobes” for guiding the person’s adaptive behavior. In this view, the treatment consists of determining and implementing the specific environmental conditions that will elicit and support the desired behaviors. An alternative view seeks to retrain or reinforce the separate abilities and tasks that are disrupted by executive dysfunction, for example, socially acceptable behavior or step-by-step approaches to specific situations. The third approach attempts to remediate one or more cognitive processes that affect executive function in a variety of contexts. These processes include “meta-strategies” such as generic problem-solving approaches (393) or techniques of verbal self-instruction (394). Some of these methods have shown promise. For example, in a study of 30 persons with TBI, Levine et al. provided empirical support for a strategy training method known as Goal Management Training (395). A short-term trial of this method, which teaches the steps involved in defining goals, breaking them into steps, and checking one’s performance, was associated with fewer errors on cognitively demanding tasks compared to a sham-training method (practice on cognitive tasks without strategic training). However, the extent to which this method can achieve gains that are able to be generalized to improve the day-to-day executive function of individuals with TBI has yet to be established. Other methods for addressing executive dysfunction have been discussed in the clinical literature, notably by Mateer (396).

Executive impairments are of profound functional significance for severely brain-injured persons and are major obstacles to independent living, employment, and successful relationships. Until more information is available as to the effectiveness of general neuropsychological tests, particularly its impact on real-life problems, it would seem prudent not to focus exclusively on retraining abstract capacities, but to include functional activities such as planning a meal, organizing transportation, and the like.

All team members are likely to be involved in addressing these problems. The neuropsychologist may be able to suggest treatment strategies applicable to all disciplines. In addition, each therapist can assist the patient in solving the specific problems faced within tasks. For example, the physical therapist may help the patient plan how to navigate an unfamiliar environment, while the occupational therapist may teach strategies for organizing a menu. Team members should be reminded that executive functions are least challenged in highly structured and routine environments such as hospitals and clinics. Thus, meaningful evaluation and treatment of these difficulties requires clinicians to avoid overstructuring patients, and to make use of real-world activities rather than limiting treatment to simulations or paper-and-pencil tasks within the clinical environment.

Impairments of Language and Communication

Because of the diffuse nature of TBI, the classic language disorders found in stroke (i.e., aphasia) are relatively uncommon unless there is a focal dominant hemisphere lesion. More typical problems involving language and communication after TBI are related to the everyday uses of language, that is, conversation, narrative speech, and “pragmatics,” or social rules of communication. Impairments in these areas have an impact on community function and contribute to social isolation. In research on discourse and conversation, people with TBI have been shown to have difficulty selecting the amount and type of data to provide to others during exchanges of information: they may say too much, not enough, or repeat information unnecessarily (397). Ideas may be expressed tangentially, with difficulty staying on topic and reduced coherence in the narrative stream (398). Other social communication problems include overtalkativeness, difficulty initiating and taking turns in conversation, poor use of social cues emitted by the listener and impaired social perception—specifically impaired perception of
emotion in others (399). Collectively, these behaviors require others to assume most of the burden of communication during an exchange (398). Impaired conversational skills have been shown to persist for at least several years postinjury (397) and to correlate significantly with limitations in social integration (400). The La Trobe Communication Questionnaire (401) has been developed and validated as a measure of perceived communication ability, and the Awareness of Social Inference Test (TASIT) (402) as a measure of social perception following TBI.

In the clinical setting, apparent disorders of language and communication may be related to other cognitive impairments. For example, a patient may respond inappropriately to questions asked in a crowded, noisy room because of problems in auditory comprehension, abstract reasoning, and/or divided attention. For the majority of people with TBI, basic language skills will appear functionally normal to the casual observer. In a more challenging environment such as school or work, however, or in grappling with the demands of intimate relationships, communication difficulties may come to the fore and pose major obstacles to successful community integration. For instance, humor is a critically important mode of communication and social learning in adolescence. Brain-injured adolescents show impairments in interpreting linguistic humor relative to their age-matched peers, even when their scores on standard language tests are within normal limits (403).

The treatment of communication difficulties is frequently coordinated by the speech/language pathologist on the rehabilitation team. It is important that all team members, and especially family members, learn to support and reinforce the same communication methods. Patients with motor speech deficits (dysarthrias) may receive training in the use of augmentative communication devices. These may be low-tech devices such as books of pictures or letters of the alphabet to which the patient points, or computer devices and software for conversion of text to speech and vice versa (404). Pragmatic disorders of communication, such as tangentiality and wordiness, can be treated within structured groups which provide audiocassette, videotape, and listener feedback. Persons with TBI can also benefit from modeling of good communication skills by other group members and group leaders. Peer group training for pragmatic communication skills has been validated empirically in a small sample by Wiseman-Hakes et al. (405). A recent study by Bornhofen and McDonald (406) suggests that aspects of emotion perception and social inference can be successfully retrained using errorless learning and self-instructional methods.

**Agitated, Aggressive, and Disinhibited Behavior**

Problems with behavioral and affective disorders are common after TBI and present in different forms along the continuum of recovery. In the acute stages of injury, agitation may appear as patients emerge from coma and PTA. It has been estimated that one third to one half of patients display agitation following coma (407). In a consecutive series of 100 patients admitted with severe TBI, Brooke et al. documented restlessness in 35 and agitation (defined as aggressive or threatening behaviors) in 11 (408). The factors that predispose to posttraumatic agitation are not fully understood. One study showed that acutely aggressive patients were likely to be older and more likely to be disoriented to both place and time (409). The possible role of pain (which the patient is unable to verbalize) as well as disturbances in the sleep-wake cycle must be considered. Posttraumatic agitation typically resolves within 2 to 3 weeks of its appearance (410). Its resolution is generally paralleled by improvement in orientation and other cognitive skills as the patient emerges from PTA, lending credence to the clinical view that agitation is a "stage of recovery" for some patients (411). Acute agitation predicts longer hospital stays and lower levels of independence at hospital discharge, although these relationships are confounded to some degree by the fact that agitated patients tend to be more cognitively impaired than those without agitation (412).

Agitation may include a spectrum of behaviors, not all of which occur in all patients. Several quantitative scales have been developed for the objective assessment of these behaviors; probably the most thoroughly studied during acute rehabilitation hospitalization is the 14-item Agitated Behavior Scale (ABS) developed by Corrigan, Bogner, and colleagues (413). The ABS displays good interrater reliability (414), and rating scale analysis suggests that it measures a unitary construct despite including disparate items pertaining to hyperactivity and repetitive behavior, emotionality, aggression, and cognitive impairment (415). The use of an objective, quantitative scale such as the ABS to track recovery and monitor response to treatment is recommended for several reasons. A single patient may display a variety of behaviors at different times, such that positive or negative changes might be missed if the focus of observation is too narrow. In addition, subjective assessments of behavioral change are unreliable, and may be skewed by a few dramatic incidents or by staff members' own emotional reactions to agitated behavior.

Management of acute agitation is a high priority in rehabilitation centers because agitation can place the patient in physical jeopardy, interfere with treatment, and cause general disruption in the therapy environment. It is frequently treated with psychoactive medication, but there are few controlled studies of the effectiveness of different medications. One such study showed that propranolol reduced the intensity of acute TBI agitation, although the number of agitated episodes did not change (410). Definitive evaluation of specific pharmacologic treatments will require a large parallel group design to control for the frequent but variable spontaneous improvement in agitation. A 1997 survey of physicians practicing in TBI rehabilitation facilities (416) contrasted the medications most commonly prescribed for agitation by "experts" in TBI rehabilitation versus by "nonexperts." Experts stated that they prescribed carbamazepine, tricyclic antidepressants, trazodone, amantadine, and β-blockers, in that order. Nonexperts used some of the same medications but also listed neuroleptics (e.g., haloperidol) and benzodiazepines, two classes of drugs that the experts reported avoiding unless other medications had failed.
Benzodiazepines and other sedatives have been reported to produce paradoxical agitation (417), perhaps by making the patient less able to process events in the environment and by further impairing inhibitory mechanisms. Neuroleptics may produce akathisia that can mimic agitation (418), may produce paradoxic delusions (419), and have been associated with delayed neurologic recovery in animal models (86). Atypical antipsychotic agents have been added to the treatment armamentarium since the conduct of this survey, although studies of their comparative efficacy in TBI are lacking. Theoretical advantages include a lower incidence of extrapyramidal reactions and less complete blockade of dopamine activity that may already be subnormal due to the brain injury. The advantages, disadvantages, and indications of various classes of drugs for acute agitation have been reviewed by Mysiw and Sandel (420).

In our experience, optimal treatment of posttraumatic agitation during acute rehabilitation requires a thoughtful combination of repeated observations and assessments of behavior, carefully monitored trials of selected medications (175), and ongoing staff education and support. Since recently injured patients may move rapidly through a stage of combative and agitated behavior (408), therapy and nursing staff may need encouragement to adopt an attitude of “watchful waiting” prior to initiating medications or heavy restraints. Less experienced staff may also need education to realize that there is no medication that specifically targets single, troublesome behaviors (such as screaming) and that drugs effective for such behaviors may also inhibit other, more adaptive behaviors. For patients who may be moving through temporary agitation, treatment may simply consist of providing more intensive staffing in the short term, and removing dangerous objects from the environment. Formal therapies need to be suspended or moved to environments with low levels of stimulation. The patient may need to be moved to a mattress on the floor to prevent injury. Environmental conditions that promote calm behavior should be observed and implemented as much as possible. A survey of nurses working on brain injury rehabilitation units identified excessive stimulation, excessively rigid rules, and interactions with staff, other patients, and family members, as the most frequent environmental triggers for agitation (421).

Patients with persistent agitation or ongoing problematic behavior, such as aggression, require more active interventions that may include medications, behavior modification techniques, or both. A behaviorally trained psychologist is often helpful in leading the treatment team in collecting and analyzing behavioral information prior to initiating treatment. Baseline data should be collected by all disciplines using an agreed-on operational definition of the problem behavior(s) or a scale such as the ABS, the Overt Aggression Scale (422), or the more recently developed Overt Behavior Scale, which covers a broader range of behaviors (423). Data should document frequency, duration, or intensity of the behavior, possible environmental factors that may have contributed to it (e.g., a difficult task, a late-arriving meal), and the consequences that followed it (e.g., a need was met, the patient was scolded). The team should meet periodically to pool data, so it may be analyzed for patterns and factors contributing to the problem. A mood disturbance may be revealed by occurrence of a behavior at times of family contact or by episodes of crying. Cognitive impairments may be implicated if the behavior occurs primarily when the patient is working on one content area or frustrated by attempts at communication. Such hypotheses generated by the baseline data may be tested by, for example, altering the difficulty of tasks or the method of giving feedback (424). Generally, in treating persons with acute TBI, working to prevent unwanted behaviors is more effective than trying to design “consequences” that are supposed to affect the occurrence of the behavior in the future. Therapy staff may also experience much success by using redirection techniques that help the cognitively impaired patient shift his or her attention away from the source of the behavior (425). Whatever approach is taken by the team, it should be consistent across team members until reevaluation of the data can take place. Continued data collection may reveal dramatic improvement, partial response, or no improvement. Reanalysis may clarify the reasons and allow for modification of the treatment plan.

As discussed in earlier sections of this chapter, some patients with TBI do not fully recover control over volatile or disinhibited behavior; others who did not display agitation acutely may develop chronic aggression. Indeed, approximately 20% of individuals who received acute rehabilitation for a serious TBI display aggression directed toward others (compared to 4% of controls) with no significant differences in cross-sectional prevalence between 6 months, 1 year, and 5 years after injury, although individual patients showed reduction of aggression or new onset of aggression within this interval (198). Interestingly, aggression toward self, others, and objects distinguished survivors of TBI from controls, while verbal aggression did not. Studies of prison populations that reveal high rates of prior TBI provide indirect evidence of ongoing aggression that may result in encounters with the criminal justice system (426).

Chronic aggression and irritability generally have a different character from acute agitation, in that the individual is typically oriented and functioning in the community, but experiences social friction with family, friends, and coworkers that overlaps with the types of interpersonal problems displayed by uninjured individuals. For this type of problem, the Overt Aggression Scale may be the most appropriate standardized measure.

Frontal lobe lesions are associated with chronic aggression, as one might expect given the role of frontal executive systems in inhibiting behavioral impulses that are inconsistent with ongoing behavioral goals, as well as in problem solving relevant to complex social situations. Disruption of white matter connections to prefrontal cortex may contribute to similar behavioral dysfunction in individuals without focal frontal lesions. However, one should not overlook the importance of other predictors of ongoing aggression, including depression, poor premorbid social functioning, and a history of alcohol and substance abuse (427). Together, these suggest treatment
approaches that target executive function, mood, substance abuse, and social skills, may be the most logical approach to the problem.

Traditional behavior management techniques may be applied or adapted to this population. For instance, Slifer et al. reported dramatic increases in prosocial behavior and decreases in uncooperative behavior in acutely brain-injured adolescents using differential reinforcement of appropriate behavior (428). In this simple paradigm, social reinforcement (i.e., praise and attention) and other rewards are given for cooperative behaviors, while disruptive behaviors are socially ignored. Another technique, differential reinforcement of low rates of responding, rewards people for reducing the rates of unwanted behaviors such as shouting or throwing objects (429). Such behavior modification techniques do not appear to require explicit memory and thus can be used effectively with cognitively impaired patients (430, 431), although it has been shown that persons with brain injury are less sensitive to response consequences than their uninjured counterparts (432). Rothwell et al. (433) and Ducharme (434) provide thoughtful discussions on how to adapt behavioral modification techniques to the particular needs of persons with acquired brain injury. Although highly structured behavior programs may be effective, there has been increasing interest in more collaborative approaches to behavioral rehabilitation that have been argued to lead to less conflict between patients and clinicians, and to be more easily integrated with the real-world consequences that might help to maintain behavioral progress (435). Indeed, a randomized controlled trial of “Natural Setting Behavior Management” (which provides behavior management training to community caregivers with maximal collaboration of the individual with TBI) found it to be significantly superior to both education and no-treatment controls in reducing challenging behaviors after acquired brain injury (436). Medd and Tate (437) also describe an empirically tested 6-week program for helping people with TBI to learn anger management techniques, following a cognitive-behavioral treatment model.

The role of medication treatment for chronic irritability and aggression after TBI remains controversial because of the dearth of well-controlled studies and the use of unvalidated measures of the problem and the response to treatment (347). Beta-blockers have the largest body of data supporting their efficacy (347, 438), but positive responses to other drug classes including anticonvulsants, psychostimulants, lithium carbonate, serotonergic antidepressants, and typical and atypical antipsychotics have been reported in uncontrolled trials. It should be kept in mind, however, that the efficacy of typical and atypical antipsychotics in reducing aggression in the absence of psychosis has not been clearly established in other populations such as those with developmental disabilities (439), so it should not be assumed that these drugs are effective as a “last resort” in TBI.

**Reduced Initiation**

Although challenging behaviors after TBI receive a great deal of attention, more individuals are disabled by an overall shortage of spontaneous behavior than by problems of behavioral excess. For example, some brain-injured patients simply fail to act without extensive cueing or structure imposed from without. Many survivors after TBI may be described by clinicians and family members as “lacking in initiative,” meaning simply that when left to their own devices, they are likely to occupy themselves primarily with passive pursuits such as television, or with nothing at all. In extreme cases, patients may be able to describe a course of action verbally, and express an intention to carry it out, but still do nothing. These surface descriptions, however, may mask a wide range of underlying deficits appearing in isolation or in combination. For example, some individuals may be primarily lacking in drive and, hence, at any particular moment, there is nothing that they wish to do. Others may be motivated to establish a particular goal, but that goal decays quickly and the behavior is not carried out. Still others may establish a goal and plans for carrying on the goal, but are derailed by environmental distractions.

The ability to form an intention, link that intention to an action plan, and monitor that action plan to its resolution, appears to depend most strongly on mesial frontal regions, for example, anterior cingulate cortex (440), operating in conjunction with other aspects (such as working memory and response inhibition) of the frontal executive system, as discussed previously. While it is uncommon to see patients with lesions circumscribed to this region, various forms of initiation deficits are not at all uncommon in TBI with DAI. The concept of “apathy” is strongly related to the motivation to form an intention and has been reported as a component of many different neurologic disorders. Indeed, apathy, as measured by the Apathy Evaluation Scale, was found to be most strongly correlated with measures of executive function, and learning and memory (441). However, scores on the AES were insufficiently sensitive and specific to identify patients considered “apathetic” by global clinical impression, suggesting either problems in the psychometric properties of the scale, or lack of agreement between the constructs measured by the scale and the way the concept is interpreted by clinicians (442).

Rehabilitation of impaired initiation is challenging for a number of reasons. First, as suggested above, the failure of an individual to spontaneously carry out behavioral routines can be attributed to a wide range of underlying impairments, which may require different approaches to treatment. In addition, however, the widespread neuropathology typical of TBI may directly impact systems regulating mood and energy, drive, working memory, and executive function, such that it is difficult to attribute the behavioral deficit to a single neuropsychological cause. In planning treatment for this type of problem, team members should identify the types of real-world behaviors that fail to be carried out, and should attempt to examine, step-by-step, the processes that succeed and fail between the establishment of a goal and the failure of an action plan. Doing so may reveal specific failures in goal setting, working memory, or sustained attention that can be targeted for compensatory treatment. Effective treatment should not only lead to success on the step in question, but achievement of the larger
behavioral goal as well, or investigation of additional causes of failure should continue.

Although there are no rigorous controlled trials, the neuropharmacology of the relevant neural networks, coupled with findings in related animal and human models, suggests that noradrenergic agonists such as desipramine and amitriptyline, dopaminergic agonists such as L-dopa, bromocriptine, and pergolide, and psychostimulants may be useful in energizing behavior and supporting executive systems that help maintain it (438–440).

When the problem arises from failure to maintain an intention, or failure of working memory to keep track of progress and next-steps of a task, treatment may include physical, verbal, written, or pictorial cues for the steps involved in a task, with behavioral reinforcement for proceeding from one step to the next rather than for completion of the previous step. Structured time frames for task completion, enforced by timers, may also be of help. We have successfully used tape-recorded step-by-step instructions, in the patient’s own voice, to help several patients with large right frontolimbic lesions to initiate and complete the steps of showering, dressing, and grooming without having to rely on a caretaker for these cues. Alphanumeric pagers, cell phones, and other portable electronic devices that can give automated prompts at specific times have been shown to help people with significant initiation problems to begin their daily tasks and routines in a more timely fashion (443,444). The importance of goal setting in motivating behavior should not be overlooked (445). Patients whose articulated goals promptly “decay” may benefit from a neutral alerting cue that arrives randomly and has been taught to signify that the patient should “stop and think about what you’re doing” (446).

While these emerging technologies show promise in helping people with TBI become more independent in managing their routines, electronic devices are variable in their degree of “user-friendliness,” and most are not specifically geared to persons with cognitive impairment. Treatment planning with assistive devices should therefore include careful assessment of the cognitive demands of operating them, as well on-going staff and/or family help with training, programming, problem solving, and technical support.

Whatever method is used, generalization across tasks should not be expected without explicit training. That is, the patient may learn to use a specific cueing system to initiate a morning ADL sequence but show no spontaneous improvement in starting homework independently. Development of flexible prompting strategies that can be used across tasks is an ideal goal. For example, if a person learns to use a hand-held computer to prompt a task sequence on the job, it may be possible to use the same device for other tasks with minimal new learning required. However, any application of a common strategy across tasks requires careful interdisciplinary treatment planning to ensure that strategies are trained, supported, and reinforced similarly across domains. Another treatment priority for initiation problems is education for both families and therapy staff. Whereas the need for physical assistance is obvious in patients with severe motor deficits, an on-going need for cognitive cues and prompts is easily misunderstood and can lead to frustration on the part of the patient who feels “nagged,” and on the part of the family or staff who suspect that the patient is “lazy” or simply uncooperative.

Awareness Deficits

It has been observed for over 100 years that some persons with acquired brain injury, including TBI, lack insight into deficits that may be obvious to others. Persons with TBI frequently seem to be less aware of cognitive and behavioral limitations than they are of physical deficits (447,448). Compared with uninjured control subjects, persons with TBI are also less able to detect and correct their errors on everyday tasks (449). For most patients, unawareness of deficit resolves to some degree over time with neurologic recovery and real-world experience of deficit areas (450). Reactive depression has been shown to increase along with awareness of one’s limitations, as mentioned previously, and is associated with higher symptom reporting (451).

Unawareness of deficit may be hard to distinguish from psychological denial, which refers to conscious or unconscious refusal to admit to a problem of which one is, at some level, aware. The patient’s response to direct feedback may help to sort out relative contributions of neurologic and psychological aspects (452). Regardless of its root cause, the failure to appreciate deficits creates a significant obstacle to rehabilitation efforts, and predicts poor performance in rehabilitation and poor vocational outcome (448). This may be because persons who are unaware of deficits are unmotivated to practice therapy tasks and unwilling to consider changes in employment or educational plans. In extreme cases they may see no purpose in, or need for, the entire rehabilitation effort.

The literature contains several reports of methods used to counteract unawareness of deficit. Treatment strategies may generally be classed as educational (based on teaching patients about deficits) (453) or experiential (e.g., providing systematic feedback or “planned failure” experiences) (454). Ownsworth et al. showed, in a single case with very severe TBI, that error awareness resulting in reduced errors could be enhanced in specific contexts but did not enhance awareness in other contexts or global self-knowledge of deficits (455). Ideally, feedback about deficits should be embedded in a “therapeutic milieu” that offers peer and staff support, psychotherapy, family counseling, and an emphasis on the affected person’s capabilities as well as limitations (448). In our experience, it is crucial to address awareness issues at all phases of the rehabilitation process in order to avoid patient drop-out and frustration on the part of both patients and staff. We begin the education process for patients very early, so that matter-of-fact references to brain injury and accompanying deficits will gradually be accepted as part of everyday parlance. An emphasis on tasks and content areas that are both familiar and important to the patient is also beneficial in promoting recognition of the need for further rehabilitation. Staff should be encouraged to remember that unawareness is often partially an organic deficit and that
acceptance of the permanent limitations caused by TBI may take a great deal of time, patience, and support. Therapists should also avoid overemphasizing the negative in their zeal to help patients achieve “insight,” as this can sabotage the rapport needed to help the patient develop realistic awareness of deficits and their implications.

Cognitive Rehabilitation
Given that cognitive deficits are among the most common and serious of the effects of TBI, systematic attempts to remedi- ate cognitive functions are included as an essential compo- nent in rehabilitation programs in both acute and postacute phases of injury. Cognitive rehabilitation may be defined as “a systematic, functionally oriented service of therapeutic cogni- tive activities… directed to achieve functional changes by… reinforcing… previously learned patterns of behavior, or… establishing new patterns of cognitive activity or compensa- torv mechanisms for impaired neurological systems” (456). Several important concepts are embedded within this defini- tion, including the idea that cognitive rehabilitation may be conceptualized as restorative, that is, targeted at regaining the previous skill; or compensatory, aimed at helping the person to deal with limitations by using new skills. Early attempts to develop restorative techniques used repetitive practice, such as memory drills, in the hope that this would improve cognition much as physical repetition can strengthen muscles. Such models have not proved very effective, and most programs now emphasize the development of compensatory strategies that are tailored to the individual and to the situations he/she will face in everyday life (457,458). Specific examples of intervention strategies have been discussed earlier.

With the growing emphasis on evidence-based practice in rehabilitation as in other branches of medicine, several major reviews of the literature have sought to determine whether, and under what circumstances, cognitive rehabilitation is effective (152,459). Evidence has been sought for both training programs targeted to specific cognitive skills, and more comprehen- sive or “holistic” programs in which cognitive training is combined with individual, group, and family therapies and vocational counseling/training within a therapeutic milieu (460–462). However, controlled research on cognitive rehabilita- tion methods is difficult to achieve for multiple reasons, and there have been very few well-controlled studies examining gen- eralization of gains to daily life. The bulk of available evidence suggests that, for specific cognitive functions such as attention, training in skill areas that require that function, or training in compensatory strategies as noted above, is more effective than training in “generic” cognitive activities (152,356,459). The success of a given approach will also depend on the injured individual’s level of cognitive function, self-awareness, motiva- tion, and capacity for behavioral self-regulation. Future efforts in cognitive rehabilitation may benefit in several ways from the rapid advances in information technology in the early 21st century. In addition to the potential for portable “per- sonal reminding technology” mentioned in previous sections, researchers have begun to explore the applications of VR to cognitive therapies for TBI with promising results (463,464). As technologies for designing and programming virtual envi- ronments become cheaper and more accessible, VR could have enormous potential for assisting with the efficient retraining of the cognitive components of complex skills such as route finding and driving. In the future, VR could even be used to create realistic social scenarios, to enable practice of interpersonal and communication skills.

PSYCHIATRIC DISORDERS
Depression
Depression is very common after TBI and has a major impact on functional and psychosocial outcome (465). A multicenter study of 666 patients involved in 17 TBI Model Systems found that 27% of patients, evaluated at 10 to 126 months postinjury could be considered as clinically depressed (466). A prospective study of 91 admissions for acute TBI revealed that 33% met the criteria for major depressive disorder at some time during the first year (467). Studies conducted over longer periods of time suggest that the frequency of depression increases over time, reaching its peak between one and 3 years after injury, but the frequency remains high over time even up to 30 postinjury (465,468,469). Posttraumatic depression is often accompanied by other disturbances of emotional func- tion, particularly anxiety, which is comorbid with depression in over 73% of cases (465,467). Concomitant depression and anxiety in patients’ spouses is also reported (470).

Depression probably results from both neurologic and psychosocial factors. The DAI of TBI induces acute disruption of neurotransmitter systems. It is plausible that neurotransmit- ter depletion, particularly in noradrenergic and serotonergic systems, could contribute to acute depressive symptomatol- ogy (471). Jorge et al. (472) found a subgroup of “transiently depressed” patients whose symptoms cleared within 3 months. Transient depression was most often associated with left frontal or subcortical lesions. Major depression was associated with reduced gray matter volume in the lateral aspects of the left prefrontal cortex and with impaired performance on execu- tive tasks (467), reminiscent of the association between left hemisphere lesions and depression in stroke patients (473). However, other investigators have failed to find correlations between lesion site and depression after TBI, nor does severity of the injury appear to be a good predictor of mood disorder (474).

The development of depression may be influenced by pre- as well as postinjury psychosocial factors, although these remain poorly understood. Prior psychiatric disorder is present at a higher rate than that in the general population, and those with preexisting psychiatric disorders are likely to develop the same disorder postinjury (475). Nevertheless, up to two thirds of psychiatric disorders develop for the first time after injury (465). Possibly for this reason, some studies have shown neg- ligible effects of both personal and familial psychiatric history (476).
The diagnosis of TBI-related depression can be complicated. Silver and Yudofsky (477) enumerate the issues that must be considered:

- Medications including anticonvulsants, narcotics, and benzodiazepines can cause or exacerbate depression.
- Depression may predate the injury even if it has never been diagnosed. Premorbid alcohol abuse and injury circumstances that hint at self-destructive behavior are common indicators.
- The vegetative signs of depression, such as insomnia or hypertriglyceremia, or decreased appetite, may be present for other reasons after TBI in patients who are not depressed. However, some vegetative signs (e.g., subjective anergia) do appear more frequently in TBI patients who also complain of depressed mood (478). The usual cognitive signs of depression (e.g., difficulty in concentrating) are useless for making the diagnosis following TBI.
- Self-report measures such as the Beck Depression Inventory (BDI) (479) and the NFI (466) may be useful screening tools for assessing the presence of varying levels of depression in persons with TBI.

The incidence of suicide after TBI has not been studied systematically; unfortunately, experienced clinicians can often cite at least a few cases. Of 111 patients followed by Klonoff and Lage (480), two committed suicide, and another two were hospitalized to prevent suicide—a similar suicide rate to that seen in penetrating wartime brain injuries (481). In the Klonoff and Lage study, 14 patients reported suicidal ideation, but half of these had sustained their TBI via self-destructive acts. Simpson and Tate (482) studied 172 outpatients with TBI using the BDI, the Beck Scale of Suicidal Ideation, and the Beck Hopelessness Scale. Twenty-three percent (23%) of their participants reported suicidal ideation and 18% had made at least one suicide attempt postinjury.

The treatment of depression following TBI usually requires persistent application of more than one therapeutic modality. Medication, psychotherapy, and community reentry programs may all be effective, especially in combination. Group and individual psychotherapy can help survivors of TBI reestablish a sense of identity and self-worth. Comprehensive rehabilitation programs help by reestablishing active involvement in work, recreation, and social activities.

A recent evidence-based review found insufficient evidence to support development of standards for pharmacologic treatment of TBI-related depression (347). Tricyclic antidepressants, specifically amitriptyline and desipramine, and the serotonin specific reuptake inhibitor sertraline were recommended as treatment options. Silver and Yudofsky (477) provide guidelines for the use of tricyclic and serotoninergic antidepressants in TBI. One reason to use caution with tricyclics, however, is their effects on seizure threshold (483). Although serotoninergic medications may be safer with respect to seizures and anticholinergic side effects, they may also lack some of the positive cognitive effects reported with TCAs (484,485). Dopaminergic agonists and other psychostimulants should also be kept in mind for their antidepressant effects, particularly if the clinical picture includes problems with attention and arousal and if the depression is not severe (486).

**Anxiety Disorders**

There has been much less investigation of anxiety disorders following TBI, but these appear to occur in a significant proportion of cases (468,469,476,487). A recent study by Goodinson et al. (465) found that 38% of a group of 100 TBI patients followed up 1 to 5 years postinjury had developed an anxiety disorder postinjury. Generalized anxiety disorder was the most common diagnosis (17%), followed by posttraumatic stress disorder (PTSD) (14%). Specific phobias (7%), panic disorder (with or without agoraphobia) (6%), and social phobia (6%) occurred with similar frequency. Only one person was diagnosed with obsessive compulsive disorder (1%) and one with agoraphobia (1%). Those with preinjury anxiety (13%) were very likely to develop postinjury anxiety, but 74% of the cases developed for the first time after injury.

Like depression, presence of anxiety disorders also appears to be associated with poorer functional outcome postinjury (465). Despite this, there has been little research on the management of anxiety disorders post-TBI. Cognitive Behavior Therapy (CBT) is recommended as the first choice of treatment for anxiety disorders in the general population (488). In individuals with TBI and very severe anxiety, medication such as an SSRI may be required as an interim measure to alleviate symptoms in order to facilitate the person’s response to CBT (487). However, considering the neurological vulnerabilities of the TBI population, wherever possible, psychotherapy should be considered prior to pharmacological interventions. One randomized controlled trial reported successful treatment of anxiety in six brain-injured individuals of mixed etiology relative to six wait-list controls, using cognitive-behavior therapy delivered individually, with gains on anxiety and depression measures maintained at one-month follow-up (489).

**Substance Abuse**

Substance abuse is a significant problem in TBI. Even without a TBI, long-term abuse of alcohol and other drugs can affect cognitive function and lead to a host of social problems. In the presence of TBI, substance abuse not only complicates the process of recovery and rehabilitation but also has a negative impact on the ultimate outcomes of injury. The relationship between alcohol and TBI is illustrated by the fact that 35% to 50% of persons with TBI are intoxicated at the time of injury (490), and a somewhat higher proportion have a previous history of alcohol or other drug abuse. While estimates vary, a recent review of published studies reported a substance abuse history in 37% to 64% of persons enrolled in TBI rehabilitation (414). Clinical screening for alcohol abuse may be accomplished bedside using brief, relatively nonconfrontational instruments such as the 4-question CAGE (491). For research purposes or in-depth assessment, it is more appropriate to use quantitative measures of the extent and consequences of substance use, such as the Michigan Alcohol Screening Test (MAST) and the Quantity-Frequency-Variability Index (492). A composite measure using the short form of the MAST,
information on quantity and frequency of consumption, and blood alcohol level at the time of injury has reportedly been useful in identifying problematic drinkers with TBI (493). Other methods to assess intake of alcohol and other drugs include questions developed for national/international health surveys, such as the National Household Survey on Drug Abuse (494), the Behavioral Risk Factors Surveillance System (495), and the AUDIT and ASSIST questionnaires developed by the World Health Organization (496,497). Such instruments have the advantage of being referenced to a large population-based sample, which can help in determining whether intake is statistically abnormal. However it is measured, a history of substance abuse predicts poorer TBI outcomes in terms of vocational success and overall quality of life (490,498). After injury, continued substance abuse can exacerbate neurobehavioral deficits, contribute to disruption of family relationships, and further reduce the probability of return to work. Several studies have shown a decline in drinking and drug use in the first year or so after TBI (499,500), but it also appears that people tend to return to preinjury drinking and drug use patterns by 2 years postinjury (500). These are mostly young males who were heavy substance users preinjury, whereas few develop substance use problems for the first time after injury (500). There has been less research on the impact of drugs other than alcohol on TBI; there is overlap among samples, as many drug abusers also consume alcohol (414). A history of street drug use was acknowledged by one-third of outpatient TBI participants in one study; marijuana was the most popular, followed by cocaine (501). Substance abuse remains a serious problem for a significant number of persons following TBI, and rehabilitation programs must be ready to help patients address the issue either with trained staff or appropriate referral mechanisms.

The implementation and success of treatment approaches commonly used with nondisabled populations, such as the 12-step programs of Alcoholics Anonymous and Narcotics Anonymous, may be complicated by the memory, communication and reasoning deficits caused by TBI. There are published recommendations for how to adapt these programs for TBI (502), and specialized substance abuse treatment approaches have also been developed for this population. For example, Corrigan et al. (414) describe comprehensive models for integrating substance abuse treatment into community-based TBI rehabilitation. These models emphasize patient, family and staff education, in-depth assessment, careful coordination of services, and specialized therapy techniques such as "motivational interviewing," to help the person internalize the motivation to change substance use habits. Other researchers have emphasized the need to assess each person’s receptivity to substance abuse counseling by using questionnaires specialized for this purpose, such as the “Readiness to Change Questionnaire.” This brief questionnaire has been validated for use with persons with TBI (503) and may be especially useful for revealing the “windows of opportunity” for drug and alcohol counseling that occur relatively soon after injury (504).

In a small study involving 12 TBI participants, Bombardier and Rimmele (505) found that 75% of those who received a brief motivational interview reported drinking less than one drink during a typical week at 1 year follow-up, compared to 55% of individuals who did not receive any intervention.

**PARTICIPATION IN COMMUNITY LIFE**

**Concept of Task Analysis**

As already discussed, many different impairments may contribute to the defective performance of important skills and, in interaction with the physical and social environment, may limit participation in social and vocational activities. To understand which impairments are responsible for specific limitations in activities and participation, a task analysis must be conducted. This involves considering the physical, cognitive, and social demands that different tasks place on the individual to clarify where task breakdown occurs. Ultimately, a clinician specializing in TBI rehabilitation must be as perceptive and knowledgeable about tasks as about patients. In Table 24-5, a

<table>
<thead>
<tr>
<th>Component</th>
<th>Role in Mobility</th>
</tr>
</thead>
<tbody>
<tr>
<td>Range of motion</td>
<td>Must allow for required movements</td>
</tr>
<tr>
<td>Strength</td>
<td>Necessary for ambulation, wheelchair propulsion, or switch operation</td>
</tr>
<tr>
<td>Balance and postural reflexes</td>
<td>Necessary for safe ambulation and transfer and adjustment to sudden perturbations</td>
</tr>
<tr>
<td>Muscle tone</td>
<td>Must allow for effective use of strength</td>
</tr>
<tr>
<td>Visuospatial perception</td>
<td>Necessary for environmental navigation</td>
</tr>
<tr>
<td>Spatial attention</td>
<td>Necessary for awareness of both sides of space</td>
</tr>
<tr>
<td>Concentration</td>
<td>Necessary for maintenance of locomotion in presence of distractions</td>
</tr>
<tr>
<td>Memory</td>
<td>Necessary for using previous experience of routes and locations</td>
</tr>
<tr>
<td>Planning, organization, and reasoning skills</td>
<td>Necessary for mobility in unfamiliar environments and using public transportation</td>
</tr>
<tr>
<td>Initiation</td>
<td>Necessary for turning plans into action</td>
</tr>
</tbody>
</table>
simple task analysis has been done for a mobility skill, demonstrating that an ability that is often thought of as physical has important cognitive requirements as well. Once the task demands have been understood, there is a choice about whether to modify the task to be more accommodating to the individual, or to try to modify the individual to be more adept at the task. This choice will depend on time since injury, learning ability of the individual, availability of effective treatments for specific impairments, flexibility of the task requirements, perspective of the patient, and many other factors.

Community Mobility
As shown in Table 24-5, community mobility may be limited by a complex combination of physical, cognitive, and neurobehavioral impairments. These interact with availability of transportation services, where the individual lives, presence of curb cuts, and so forth. Aggressive attempts should be made in the early months postinjury to improve these impairments. Purchase of permanent assistive devices should be postponed until the patient's ultimate level of physical function is relatively certain.

Many survivors of TBI lose the ability to drive because of epilepsy or visual, cognitive, or motor impairments (506). This severely limits community mobility, particularly in rural areas. Some may regain the ability to drive, but the criteria on which to base this judgment are controversial. Certain perceptual and motor tests may reveal impairments that preclude driving, but good performance in basic sensory and motor functions does not guarantee safe driving. Attention-demanding driving simulation tasks (507,508) and VR road tests (509) show promise in helping to evaluate driver safety after TBI, but further validation is necessary before their sensitivity and specificity is known. Actual road tests can also be used, though the range of traffic occurrences that will occur during such tests is difficult to control. Preliminary research suggests that some poor drivers may benefit from specific retraining (510). A small study of survivors of TBI who had completed a postinjury driving evaluation program and had returned to driving, did not reveal an elevated risk of accidents, providing some reassurance that such programs do not return drivers to the road who should not be driving (511); whether there are individuals who could be driving safely but are not able to clear the hurdles to return to the road is unclear.

Self-Care Skills
As with mobility, ADL independence may be limited by a complex combination of physical and cognitive impairments. A similar process of interdisciplinary physical and cognitive assessment may lead to identification of particular component processes that are especially significant. Basic ADL skills such as dressing, bathing, and feeding may be improved with a treatment program aimed at three components:

- Improving salient physical or cognitive deficits such as serial casting for contractures or stimulant medications for arousal and attention deficits
- Compensating for salient physical or cognitive deficits through provision of assistive devices for physical deficits, or written or pictorial cue cards for cognitive deficits
- Retraining tasks without specific regard for the contributing deficits using behavioral training methods (e.g., breaking down into small steps, backwards chaining, errorless learning, reinforcement) (512).

Similar treatment principles are involved in addressing instrumental ADL, which generally place greater demands on executive function. Training provided in a hospital or clinic setting will need to be generalized to the home environment through family training and/or a process of home visits and home therapy.

Social and Family Impact
Following TBI, the most obvious and disabling impairments are in the sphere of social behavior. The person with brain injury may thus have great difficulty maintaining preinjury relationships or establishing new ones. These behaviors are ultimately attributable to the interplay among neurobehavioral deficits, unique premorbid personality traits, and postinjury coping responses. Deficits in frontal executive functions, communication, and behavioral control (described in previous sections of this chapter) may particularly predispose patients to long-standing social problems. Poor performance on tests of cognitive flexibility is predictive of social withdrawal (513), as is impaired initiation (514). Reviewing the literature on social disability after TBI, Morton and Wehman (515) pointed to four recurring themes:

- Decline in friendships and social supports, leading to isolation that does not improve spontaneously
- Lack of opportunity to make new social contacts because of a restricted range of activities (e.g., unemployment, inability to drive)
- Inability to engage in preinjury leisure activities
- Depression, which further reduces social initiative and increases isolation

Unfortunately, treatment of social behavior problems is seldom completely effective. Group therapy with modeling of appropriate behavior can be beneficial and is part of most comprehensive community reentry programs for TBI. Success has been reported using structured social skills training, with videotaped feedback, in postacute patients (516). A study of a structured peer support program for persons with brain injury and their families indicated that it improved participants' overall quality of life, general outlook, coping skills and knowledge of TBI, but did not enhance level of social support received from family and friends (517). The internet has also been used as a means of providing information and social support, but its efficacy has not yet been validated (518).

Only a few investigators have systematically examined the effects of TBI on leisure and recreational participation, but they concur that TBI significantly disrupts premorbid leisure pursuits, due to residual physical and neurobehavioral...
impairments (519,520). Because many do not return to work either, loss of leisure activity may be psychologically devastat-
ing. A recreational therapist can be an integral team member, performing leisure skill assessments and developing both indi-
vidual and group treatments to teach new ways of occupying leisure time or adaptive methods for performing favorite pre-
morbid activities.

The Clubhouse model is an innovative approach to the social isolation of persons with TBI (521). Clubhouse pro-
grams are organized around a work-ordered day (i.e., up to full-
time participation), and are run by clients who have suf-
tered TBI, with staff support. A productive social milieu with peer support and naturalistic learning and opportunities for both recreational and vocational activity are key features of such programs. Client strengths and contributions are stressed, in contrast to the “deficit” emphasis found in many therapy programs.

Numerous studies have attested to the extreme stress placed by TBI on the family system (522). An early Israeli investigation of marital adjustment 1 year after war-related brain injury found interpersonal tension within the family; feelings of loneliness, depression, and isolation in the wives; lack of sexual contact with husbands because of loss of feelings of attractiveness and personality changes; and role changes within the family (523). More recent studies from the USA and New Zealand have amply replicated these findings (470,524–526). Furthermore, the stress and family disruption experienced by caretaking spouses may be worse than that experienced by parents in a comparable role (525). Studies of the sources of family stress invariably find greater impact of cognitive and behavioral deficits than physical care needs (525,527,528).

Kosciulek (529) identified five distinct and effective coping strategies used by families after TBI:

- Positive appraisal, which is commitment to seeking the posi-
tive while accepting things as they are
- Resource acquisition, or the tendency to seek help and guid-
ance
- Family tension management, which involves open expres-
sion of feelings
- Head injury demand reduction, for example, seeking sup-
port groups or becoming involved in advocacy
- Acquiring social support from friends and relatives

Inclusion of the family in rehabilitation programs has become standard practice. Families of TBI patients feel strong needs for regular communication with staff, specific information about the injury, and honest answers to their questions (530). Perhaps in response to this, a great emphasis has been placed on individual or group family education, and, as needed, family counseling, family therapy, and support groups (531). The use of a telephone-based, social work liaison follow-up has been demonstrated to reduce caregiver burden, increase quality-of-
life ratings, and enhance feelings of caregiver mastery, com-
pared to a historical control group. Though limitations were cited in the methodology by the investigators, this may prove to be “a low-cost, nonintensive intervention [which] may offer substantial benefits to families caring for people with brain injury” (532).

**Sexual Relationships**

Disturbances in sexual function after TBI often have been noted by clinicians but have not been well studied. Sexual dysfunc-
tions may include hypersexuality, hyposexuality, impotence, loss of feelings of attractiveness, inability to find appropriate partners, and incapacity to engage in intimate interpersonal relationship requiring the interpretation and expression of complex emotions (533).

Staff members and relatives often are dismayed by the sexual dysfunctions after TBI and have difficulty in managing these problems effectively. Rehabilitation facilities have developed sexual reeducation programs for brain-injured patients in which specific information is provided about sexual func-
tion and basic social skills in interpersonal intimacy are taught (533).

Sexual behavior most often is perceived by staff as a prob-
lem when the patient engages in overly inappropriate behav-
ior, such as public masturbation or continual advances toward uninterested partners. In these cases, team members often are asked to intervene to help reeducate and redirect the patient toward a more appropriate expression of sexual drive. Behavioral interventions may assist in reducing these problems.

Inadequate sexual function too often is neglected because it creates fewer problems for family and staff. However, it is a much more common problem. In a study of 208 individuals with moderate-severe TBI (69% males) 1 to 5 years postinjury, Ponsford (528) found that the majority of participants reported significant negative postinjury sexual changes, with 36% to 54% reporting a decrease in the importance of sexuality, sex drive, sexual opportunities, frequency and enjoyment of sexual activity, in addition to their ability to give their partner sexual satisfaction, to engage in sexual intercourse, to stay aroused, and to climax. The frequencies of such negative changes were significantly higher than those reported by a demographically matched control group and far outweighed the frequency of increases on these dimensions. A significant proportion of TBI participants also reported decreased self-confidence, sex appeal, higher levels of depression, and decreased communica-
tion levels and relationship quality with their sexual partner. It is likely that these changes occur as a result of both organic factors, including limbic system injury and/or hormonal disrup-
tion, and other factors including fatigue, physical disability, pain, reduced self-esteem and social skills, relationship quality, and affective change. Medications may also impact on sexual function.

It is extremely important that all potential causes of sexual difficulties be investigated and attempts made to address them. Assessing sexuality should be a routine part of the assessment and rehabilitation process, because these are not easy issues for clients to raise. Rehabilitation staff need to be willing to talk with their clients about sexual issues and have access to specialist skills for assessing and dealing with them, in the context of an understanding of the injured person’s cog-
tive limitations. The “You and Me” program, developed by Simpson (534), represents a valuable resource for staff, clients, and their families.

It is also critical for the physician to provide information regarding safe sex practices and contraception to all patients. These include mutual disease screening if indicated and use of condoms and spermacides. However, education alone is often inadequate for individuals who have difficulties with impulse control and executive function and who may be vulnerable to manipulation by others either through personal contact or through the internet. Therefore, sexual counseling should occur in the context of a full appreciation of the individual’s neuropsychological and emotional status. Communication and intimacy may be enhanced for some patients and their partners through the use of educational materials and individual and couples counseling. Some patients, particularly those with specific sexual disorders, may require the intervention of a sex therapist.

**Educational, Prevocational, and Vocational Function**

TBI may produce significant educational and vocational disability. Numerous studies of return to work following TBI suggest that both premorbid and injury-related factors play strong roles in vocational success. Previous education, work history, psychiatric history, and substance abuse are all strong predictors of return to work (535). In addition, injury severity, cognitive impairment, and behavioral disturbances are related to reemployment (535–537). Social isolation after injury has also been negatively associated with return to work (535), but whether this reflects the coexistence of depression, lack of initiative, or poor social skills that interfere with work, or whether the social isolation is the result of loss of work-related social outlets is unclear. More than half of a group of 172 of those with moderate-to-severe TBI who had returned to work reported that their duties had changed and more than a third reported continuing fatigue, problems keeping up, and a tendency to make mistakes even 5 years postinjury (23).

Many of the same factors adversely affect school performance among individuals suffering TBI during their youth. Secondary school and college students returning to study after TBI often report difficulty keeping up with the workload, with slowed learning and increased fatigue, such that 89% report that it takes a moderate or extreme degree of effort to pass courses and many require special educational accommodations (23). However, since education, unlike employment, is an entitlement, the impact of TBI may be less obvious. Although Public Law 94-142 guarantees educational opportunities for brain-injured children, as well as other disability groups, up to the age of 21 years, the implementation of special accommodations for them varies greatly, depending on the local school district’s educational philosophy, available resources, and understanding of the problem. For the elementary and secondary school child, transition back into school can be traumatic. For this reason, it is advisable for school personnel, rehabilitation team members, and the family to help the school develop an individualized education plan and to ensure that all teachers understand the impact of the brain injury. Once the child leaves school, further assistance will generally be required in order to facilitate exploration of vocational options.

A growing number of community colleges and universities offer educational programs tailored to individuals with brain injury. At these schools, remedial courses in basic skill areas such as reading or mathematics are provided. In other cases, supplemental tutoring may be made available. To assist with daily living needs, a group-living situation with assistance from a residential counselor sometimes is available. These modifications allow a higher-level brain-injured person to progress at a slower pace with a great deal of support and a far greater likelihood of academic success than in a typical college environment.

Vocational goals are important to many brain-injured people once they complete their acute rehabilitation. The major problem is that their vocational goals may be based on their premorbid abilities or vocation and may be unrealistic in the context of current disabilities. Some comprehensive rehabilitation programs include prevocational assessment and vocational training (538). After the patient completes a period of extensive cognitive rehabilitation, a vocational counselor designs a vocational program that usually involves work trials within the rehabilitation setting, progressing into the workplace. Another model of vocational rehabilitation after brain injury has been the supported employment approach (539,540). In this approach, the client may be placed directly into a competitive employment situation, where a job coach provides daily on-the-job instruction and guidance, advises the employer as to potential job modifications needed to maximize the client’s performance, and assists with other job-related issues such as transportation, housing, financial management, and substance abuse. The supported work model has been viewed as the most effective in returning brain-injured people to competitive employment, although there have been very few rigorous studies of efficacy. The conclusion of efficacy is based mainly on successful return to work after participation in such programs of individuals who have been previously unsuccessful (102,541).

Every U.S. state has a Department of Vocational Rehabilitation, providing any of a wide variety of services, including prevocational assessment, vocational assessment, work adjustment training, sheltered workshop training, or a supported work program. A neuropsychological assessment usually is requested by vocational counselors as an essential component on which to base the vocational plan. However, it is generally agreed that direct observation of clients in work settings or simulations is a more useful guide to vocational success and appropriate vocational goals than standardized neuropsychological testing (542). It is very important to educate the employer and fellow employees regarding the injured person’s difficulties and ways of supporting him or her. A graduated approach to return to work is generally necessary, with follow-up contact to deal with problems that may arise as time passes and work demands or fellow employees change.
A variety of medicolegal and ethical issues can complicate the rehabilitation process. There is frequently the need to determine the competency of the brain-injured person to manage financial affairs, engage in vocational pursuits, make medical decisions (543), or provide informed consent for research participation (544). Obtaining informed consent in the rehabilitation setting for people with TBI poses its own unique difficulties (545,546) as related to the use of behavioral or psychopharmacologic restrictive procedures (547), decision making about surgical interventions or transfer to a different type of treatment program, and rights to privacy (548). Competency assessments after TBI can be complex, involving the contributions of many disciplines (549). If the court deems a brain-injured person to be incompetent, a conservator may need to be appointed to manage financial affairs or a guardian to make decisions for the person. To complicate matters, family members are often reluctant to move toward guardianship early, in hopes that the individual will recover sufficiently to remain in charge of his/her own affairs. This often leaves an ethical gray area when clinicians understand that an individual’s ability to make decisions is compromised, but there has been no official determination of incompetence or appointment of a guardian. Most often, clinicians seek to involve the individual and their caregivers in such decisions, thus minimizing the importance of the exact decision-making status. However, when the individual with TBI and his/her caregivers disagree about important decisions, or when there is disagreement within a group of caregivers, formal determination of the legally authorized decision maker may become necessary. Legal determinations of competency are binary, whereas clinicians recognize the fact that capacity for decision making recovers gradually (550), and that the legal dichotomy is in tension with principles of empowerment that encourage impaired individuals to participate actively, with support, in making important decisions (551).

Another issue for many is personal injury litigation, which often results from accident cases. In this circumstance, the person with brain injury tries to recoup financially for past medical expenses, pain and suffering, and past and future lost wages. Although a large settlement can be very beneficial to the individual, it often takes 5 years or longer for a case to be settled in the courts. This produces a severe stress on the survivor and the family. Furthermore, individuals with brain injury can be their own worst witnesses because of their memory impairments and normal physical appearance, and many achieve minimal financial gain (552).

Perhaps the most dramatic ethical issues in brain injury have been raised about the treatment of patients in the vegetative or MCS. Clinicians, lawyers, and theologians have debated the rights of family members to withdraw life-sustaining nutrition for people in the vegetative state (156,553,554). Although ethical viewpoints vary, terminating artificial hydration and nutrition for an individual judged to be permanently in the vegetative state is legal in most jurisdictions. Matters
physical, cognitive, and behavioral sequelae of TBI have been more clearly identified and classified.

Treatment advances also have occurred. The acute management of patients with TBI has benefited from advances in neurosurgical intensive care. Medical management of the complications of TBI also has improved. The physical sequelae of TBI such as paralysis, contractures, and HO have been lessened by the same treatments used in the rehabilitation of other physical and neurologic disabilities. Yet, the neurobehavioral deficits of many brain-injured patients represent their most significant obstacles to community reintegration. It is in these cognitive and behavioral realms that our ability to improve function is least clearly defined.

Much research is in progress to refine our classification of these deficits, to clarify the extent to which they are remediable, and to identify the most effective remediation strategies. Because each brain-injured patient’s pattern of deficits is unique, it will always remain a challenge to apply the knowledge gathered from group studies to the management of the individual patient. For this reason, a thoughtful interdisciplinary treatment-planning process that considers the complex interactions of the many physical, cognitive, and behavioral deficits is essential.

REFERENCES


324. Fernandez-Duque D, Posner MI. Brain imaging of attentional network


Multiple sclerosis (MS) is a complex, multifactorial disease that requires the expertise of neurologists, physiatrists, occupational therapists, and physical therapists, as well as countless other participants in the patient’s care. Rehabilitation physicians have an integral role during the disease course including the assessment and management of paresis, spasticity, gait, activities of daily living, bowel and bladder function, wellness issues, and pain. Since MS is a long-term and often progressive disorder, the focus is not only on improving function where necessary but also on maintaining function where possible. A firm understanding of the underlying pathophysiology of MS as well as the role and rationale of treatment is the key to the rehabilitative treatment of MS.

MS is the most common cause of nontraumatic disability affecting young adults in the Northern Hemisphere (1). Onset is usually in the 20s to 40s, a time when they are in the core of their working and childbearing years (2). The disease varies from patient to patient and the care needs change from year to year. Thus, it is crucial to have a team approach to maintain and improve function. There are presently six food and drug administration (FDA) approved medications that alter the disease course of MS, and there are more in the pipeline; however, none provides a cure or an alternative to rehabilitative approaches to MS care.

**DEMOGRAPHICS AND EPIDEMIOLOGY**

MS is a common disorder in North America. Prevalence varies in the United States and Canada and ranges from approximately 40 to 220 per 100,000 population (3). With some exceptions, the prevalence of MS increases further north and south of the equator one lives (4). Various theories have been proposed to explain this distribution, as well as the notable exceptions to this geographic observation in some populations (Inuit, Laplanders, etc.) (5). Theories include vitamin D deficiency due to reduced sunlight, dairy products, genetic populations, and exposures to a variety of environmental factors (6). MS is preponderantly a disease of persons of Northern European ancestry (7). Infections have been linked to MS, but no one infection has emerged as a specific cause or precipitant of disease activity (8). In pediatric MS, late infection with mononucleosis has been associated with MS (9). Exacerbations of MS are more common after various infections, but the type of infection appears to be immaterial (10).

**ETIOLOGY**

MS is considered to be an “autoimmune” diseases (11). Like all autoimmune disorders, it is more common in women (12), occurring about twice as often in women as in men (13). MS also appears to involve genetic factors with HLA-DR2 in DR-positive families—having a greater chance of developing the disease (14). The risk of concordant MS is 30% with monozygotic twins, 5% with dizygotic twins, and between 2% and 4% for first-degree family members of people with MS (15). Multiple genetic linkage studies have confirmed a linkage with the major histocompatibility region, as well as less well defined linkages to other zones that code for interleukins (16). MS is not increased in children adopted into families with MS (17), which indicates that family aggregation of MS is due to genetic factors rather than environmental (18). The specific cause of MS remains unknown.

**PATHOLOGY, PATHOGENESIS, AND PATHOPHYSIOLOGY**

The hallmark of MS pathology is the presence of multifocal plaques (lesions) of demyelination in the cerebral hemispheres, optic nerves, brain stem, and spinal cord (19). The early plaque has a demarcated area of demyelination with incomplete axonal injury, inflammatory infiltrates composed of lymphocytes and macrophages, and evidence of astrocytic proliferation and gliosis (20). Ultrastructural analysis of axons has shown an early reduction in axonal fibers and axonal transections in new demyelinating regions (21).

In acute active lesions, gadolinium leaks into tissue parenchyma due to blood-brain barrier (BBB) interruption that accompanies the inflammatory response (22). Perivascular lymphocytic infiltrates are evident in areas of demyelination (23). Macrophages are the most prominent inflammatory cells in the lesion and many are filled with myelin debris (24). Proinflammatory CD4+ T cells appear to be critical components in the demyelinating process (25). B cells from peripheral blood are recruited to the active plaque and differentiate into plasma cells that synthesize and release immunoglobulin (26). Plasma cells are more commonly encountered in chronic lesions (27).

With progression of the pathological process, chronic inactive lesions are less inflammatory and become hypocellular,
with relatively quiescent oligodendrocyte precursor cells (28). Degeneration occurs, as oligodendrocytes are destroyed and astrocytes proliferate. As demyelination occurs, there can also be some degree of remyelination in the lesion (29). However, as the disease progresses, demyelination at the plaque margin takes place in the newly remyelinated areas and leads to expansion of the lesion. This eventually results in permanent scarring.

PATHOPHYSIOLOGY OF MS

The clinical symptoms of MS are often due to loss of axonal conduction. Demyelination of segments of conducting axons causes conduction block which varies depending on the extent of demyelination and whether compensatory mechanisms have intervened. Conduction block in experimental demyelination occurs at sites of demyelination and does not occur in otherwise unaffected segments (30). The block appears to be most severe in the first few days after experimental demyelination (31). Acutely demyelinated axolemma has a relatively low sodium channel density which may be insufficient for the action potential to be propagated effectively (32).

Another factor influencing conduction is inflammation (33). Cytokines have been shown to play a role in conduction block, particularly proinflammatory cytokines such as tumor necrosis factor-α and interferon (IFN)-γ (34). Both of these cytokines induce nitric oxide formation (35). Nitric oxide production is increased in MS (36) and has been shown to cause a dose-dependent conduction block (37).

Safety factor is a measure of the excess current allowing conduction divided by the minimum current necessary to depolarize an axon (38). In normal myelinated axons, this is usually a factor of three to five times. In demyelinated axons this is reduced, often measured at just above 1. Small changes in environmental factors can thus cause axonal block in such fibers (reduced safety factor). This may be the basis for worsening MS symptoms with fever and with exercise, both of which may reduce safety factor to the point where conduction block occurs.

Restoration of conduction occurs in demyelinated axons after a few days or weeks. This is likely related to the appearance of sodium channels along the demyelinated portion of the axon to allow microsaltatory conduction along these demyelinated segments (39). Remyelination occurs with improved conduction in previously demyelinated segments (40). In addition, the resolution of inflammation may also improve conduction block possibly due to normalization of nitric oxide levels (37).

Patients frequently note that with repeated activity they develop weakness, especially with walking. In experimental demyelination, a train of stimuli over time will elicit intermittent conduction block (30). This correlates with hyperpolarization of the membrane (41).

**Restoration, Remyelination, and Neuroprotection**

Recovery of function (remission) may be due to resolution of inflammation or the pressure of edema, removal of humoral factors, reattachment of paranodal myelin, or rerouting of nerve transmission through alternative pathways (brain plasticity) (42). Remyelination may also be a key component in the restoration phase, which usually requires days to weeks (43). Early studies of spinal cord MS lesions suggested that regeneration of central myelin is accompanied by Schwann cell invasion of the central nervous system (CNS) and the presence of peripheral myelin within the demyelinated region (29).

A classic neuropathological finding with remyelination is “shadow plaques.” These represent partial reduction of myelin staining as remyelinated axons have thinner myelin sheaths. Such remyelination may be extensive in some patients (44). Remyelination may be seen on magnetic resonance imaging (MRI) as intermediate lesions on T1 images and appear less dark compared to hypodense “black holes” (which represent significant axon damage).

**TYPES OF MS**

An international panel was convened in the late 1990s to develop a common classification of subtypes of MS (45) (Fig. 25-1). This classification is useful as a way to group patients, though separations between subtypes may be more artificial than real.
Relapsing-Remitting MS

Relapsing-remitting patients have discrete attacks of new or worsened neurological symptoms that emerge over a few days and may resolve over a 4- to 8-week period with or without corticosteroid treatment. These patients often return to their preattack baseline but may have residual disability. Relapsing-remitting multiple sclerosis (RRMS) is the most common form of MS (46). RRMS frequently begins with optic neuritis (ON)—transient unilateral visual impairment lasting days to weeks, which may be associated with retrobulbar pain (47). Others with RRMS may initially experience tingling or weakness of a limb; paresthesias are the most common initial symptom of MS (48). In general, patients with sensory symptoms, and patients whose symptoms fully remit after early exacerbations, demonstrate better long-term prognosis (48). FDA and internationally approved medications are available for patients with RRMS (49–53). These medications in general reduce new exacerbation frequency, reduce measures of MRI activity, and may slow progression of disability. Note that patients who have exacerbations and are left with residual disabilities and those with complete resolution are both still considered RRMS (Fig. 25-1). Nevertheless, from a functional perspective these two patient groups are dramatically different.

Secondary Progressive MS

These are patients who initially have RRMS but after some years of disease have gradual worsening in between discrete attacks. They may have continued attacks or stop having attacks. They also tend to have fewer attacks and less MRI gadolinium enhancement than during the relapsing-remitting phase. A significant proportion of RRMS patients convert to secondary progressive multiple sclerosis (SPMS) at some time in their disease course. IFN-β1b and mitoxantrone slow the progression of SPMS if given in the early part of this stage, particularly when there are continued relapses (53,54). However, in older patients who do not have continued relapse activity, the beneficial effect of these medications is less pronounced (55). Patient with accrued deficits such as gait disorder and spasticity may benefit from rehabilitative approaches.

Primary Progressive MS

Primary progressive multiple sclerosis (PPMS) patients are often older at onset and have a progressive course without attacks. They tend to have fewer MRI lesions than RRMS or SPMS patients. PPMS occurs in about 10% of patients with MS. PPMS tends to affect both sexes equally, unlike RRMS (56). It usually becomes clinically apparent when they are in their 50s (57). PPMS more often starts with motor symptoms, often an asymmetric paraparesis, and tends to progress more rapidly, validating observations that patients starting with motor symptoms fare less well (vs. sensory symptoms), regardless of disease type (57). There are no FDA-approved immunomodulating agents for PPMS. Although some have been tried, none has shown sufficient benefit for regulatory approval. Some randomized,

blinded trial efficacy has been reported for methotrexate, but the effect was limited to upper limb function as measured by timed Nine-Hole Peg Testing (58). Pulse methylprednisolone has also been reported in some studies to be useful in this cohort of patients (59).

Progressive-Relapsing MS

Progressive-relapsing MS is a form of MS where the patient experiences progressive disease at the beginning but begins to have attacks later in their course. Probably no more than 5% of MS patients have this form. Effective medical treatment is unclear for this subgroup.

Fulminant MS

These are patients with rapid severe progression with relapses over a period of months. This type of MS affects a small percentage of patients.

OTHER ASSOCIATED DEMYELINATING DISORDERS

Optic Neuritis

Many patients with MS will present with ON, a sudden unilateral loss of vision, which can vary from a slight central scotoma to complete loss of light perception (47). In one long-term study of ON, 57% eventually developed MS (60). Both ON and ON/MS patients respond favorably to intravenous (IV) methylprednisolone (47). A recent Cochrane analysis of corticosteroid use for ON concluded that there is no evidence of a long-term benefit of high- or low-dose steroids; the improvement seen with high-dose corticosteroids in the ON trials appears to be limited to short-term outcomes (61). Data on patients with clinical isolated syndromes (CIS) such as isolated ON have shown that patients with multiple brain lesions on MRI are at higher risk of progressing to clinical MS within 1 to 2 years after presentation (62). Obtaining a brain MRI in patients with ON or other CIS is therefore an effective risk stratification methodology and potentially guides later therapy.

Transverse Myelitis

Transverse myelitis (TM) is an inflammatory disorder of the spinal cord (63). It may present after an infection in about half of cases. Most patients have rapid progression to a relatively complete spinal cord syndrome with paralysis, sensory level, and bowel and bladder involvement. Cerebrospinal fluid (CSF) may show a pleocytosis but usually does not show oligoclonal banding (64). Recovery is variable and often incomplete. There are no randomized trials of therapy for classic TM; therapies tried include IV solumedrol, IV immunoglobulins (IVIG), and plasmapheresis (65). A subset of patients, usually with incomplete TM, will develop MS (66). In addition, a subset of TM patients have a positive neuromyelitis optica-IgG (NMO-IgG), which is a subgroup of Devic’s disease (see below). TM patients who eventually develop MS typically present with asymmetric
and incomplete motor or sensory symptoms, whereas those with pure TM not converting to MS generally have symmetric findings and more profound disability (67). Key rehabilitation issues in TM include mobility, spasticity, bowel and bladder management, and avoiding decubitus ulceration and deep venous thrombosis during the acute illness.

**Devic’s Disease (Neuromyelitis Optica)**

The traditional concept of Devic’s disease consists of the subacute combination of relatively severe ON and TM with sparing of the brain (68). With new research, the classification of neuromyelitis optica has undergone significant revision and likely will continue to evolve (69). In general, patients have a combination of ON and TM, both of which can be severe, with initial relative sparing of brain involvement. The spinal lesions tend to be longer (three or more cord segments) than typical MS lesions in the cord (usually one segment only) (69,70). CSF may show a striking pleocytosis or a neutrophilic predominance which would be unusual for MS (71). A recently developed antibody, the NMO-IgG antibody, is positive in 70% of clinically defined neuromyelitis optica and is rarely positive in MS.

**DIAGNOSIS OF MS**

The most common presenting course for MS is relapsing symptoms affecting different neurological structures occurring over time (“lesions distributed in time and space”). Common symptoms include fatigue, ataxia, weakness, numbness and tingling, bladder dysfunction, spasticity, cognitive problems, depression, ON, and pain. Notwithstanding multiple revisions over time, the diagnosis of MS still requires the presence of signs and symptoms separated in space and time, or the presence of progressive symptoms with appropriate paraclinical evidence of demyelination (72–75).

**Diagnostic Criteria**

Over the years, several sets of diagnostic criteria have been developed. In the 1960s, Schumacher et al. based the diagnosis of MS entirely on clinical findings (72). In the early 1980s, these were modified by Poser et al. to allow laboratory tests to substitute for even more clinical findings (74,75). Specific MRI criteria to support the diagnosis of MS have been elucidated (76).

In general, the diagnosis of MS continues to be a clinical one with supporting evidence from MRI, CSF, and evoked potentials as needed (75). MRI assists in the diagnosis when it shows characteristic white matter lesions. Lesions in the periventricular, juxtacortical, subcortical, infratentorial, and spinal cord regions are all helpful. The MRI can be used to demonstrate dissemination in space, replacing examination findings in two areas. Repeated MRIs over time showing new lesion formation can also be substituted for further attacks in making a diagnosis. CSF and MRI findings are helpful in PPMS, particularly in ruling out other causes of a progressive myelopathy such as spinal cord tumor (75). The McDonald criteria allow for consistent diagnostic criteria to be applied and are helpful in epidemiological studies as well as patient recruitment into clinical trials (75).

**Magnetic Resonance Imaging**

MRI has become the most important paraclinical tool in the diagnosis and monitoring of MS and related diseases (77). A variety of MRI measures have been used to analyze MS in terms of diagnosis, progression, the acute lesion, subtypes of MS, and monitoring treatment trials (78).

MRI has also been crucial in modifying our concepts of disease pathogenesis and course in MS. In the 1980s, sequential MRI trials showed that new lesion formation occurs five to ten times as often as there were new clinical events, altering the understanding of disease activity during clinically quiescent times (79–81). In addition, MRI has supported the notion of early axonal injury with MS lesions which pathological studies have introduced. Longitudinal MRI measures of atrophy have shown slow but definite atrophy with time in patients with otherwise relapsing MS (82,83). Finally, MRI imaging particularly with FLAIR (Fluid Attenuated Inversion Recovery) has shown that cortical lesions are common in MS, a fact that went unrecognized prior to these imaging techniques (84,85).

MRI has also shown us that there is a limited correlation between the extent of measurable brain disease and measurable disability. Some patients with limited MRI change have significant deficits, and some with extensive white matter and gray matter changes have limited clinical deficits. The explanation for this is unclear but may involve disease in areas not measurably affecting function; compensatory changes that mask functional deficits; spinal cord involvement not imaged on brain MRI causing deficits; and changes that do not cause noticeable neural transmission deficits.

Conventional MRI imaging of MS lesions identifies MS lesions as often paraventricular, oval-shaped lesions, sometimes oriented tangentially to the ventricular surface. Lesions of the corpus callosum are common and extend in a fingerlike pattern, known as “Dawson’s fingers” (78). With disease progression, lesions may coalesce. With more chronic and severe disease, brain atrophy becomes apparent. T1 lesions (black holes) are correlated with axonal loss in the affected areas (86).

For several weeks after the occurrence of the breakdown in the BBB, affected brain will be permeable to gadolinium. Consequently, MRIs taken shortly after the IV injection of gadolinium will show “enhancement” or opaque regions on T1-weighted images. A characteristic enhancement pattern for MS is the “open ring sign,” where gadolinium enhances in an incomplete peripheral pattern, often pointing away from the lateral ventricles (87).

Other techniques that at present have a primarily research application include brain parenchymal fraction (83), magnetization transfer ratio (88), diffusion tensor imaging (89), and MR spectroscopy (90). These techniques may be most helpful in documenting axonal tract fibers loss and in predicting disability.
Evoked Potentials
The original work in evoked potentials came from Dawson's lab in the 1940s (91). Dawson and his colleagues showed that one could record neurological activity from the CNS with peripheral stimulation of a sensory pathway using summation of multiple stimulations by averaging. They and others demonstrated slowing of central conduction in MS patients. This was consistent with the concept of a demyelinating disorder as opposed to one with primarily axonal injury. Slowing in the visual evoked potential is the most classic example, but similar slowing can be seen with somatosensory evoked potentials and brainstem auditory evoked potentials (92). American Academy of Neurology guidelines suggested that visual evoked potentials are probably useful to identify patients at increased risk for developing clinically definite MS; somatosensory evoked potentials are possibly useful to identify patients at increased risk for developing clinically definite MS; and there is insufficient evidence to recommend brainstem evoked potentials as a useful test to identify patient at increased risk for developing clinical definite MS (93). There is no data to support repeated evoked potential testing for monitoring of ongoing disease.

Ocular Coherence Tomography
Ocular coherence tomography is a new technique which allows imaging of the retinal nerve fiber layer (RNFL) thickness (94). It is a sensitive, safe, and reproducible measure that allows analysis of features such as nerve fiber loss, ganglion cell loss, and macular edema. It is being used as an independent measure in clinical trials of medications in MS, and may be a robust marker of fiber loss indicating disease progression (95,96). Its role in clinical decision making is still being evaluated (95).

Lumbar Puncture
Lumbar puncture and CSF analysis are less often performed today, as improvements in imaging have generally replaced this invasive procedure. However, when done, identification of oligoclonal bands and increased IgG synthesis are associated with a positive diagnosis of MS (97). These findings indicate activity of the immune system in and around the CNS compartment, and as such are not specific for MS. Other diseases such as lupus, Lyme disease, and neurosarcoidosis also may show oligoclonal banding. Studies of the diagnostic accuracy of MRI, evoked potentials, and CSF have shown that if two of three of these tests are positive, the third does not add significantly to the diagnosis (98). Neurologists are reducing the frequency of CSF tests in patients with otherwise typical MS and in whom MRI and evoked potentials are characteristic. CSF studies are most useful in atypical presentations, where the MRI is non-specific, and where other diagnoses are being considered.

Electrodiagnostic Testing
In general there is no role for electromyography and nerve conduction testing in MS. The peripheral nervous system is usually spared in MS. There are some studies indicating involvement of the proximal aspect of the peripheral nervous system in patients with severe MS, sometimes affecting phrenic nerve function (99). There are rare patients with central and peripheral demyelination in which electromyography and nerve conduction studies would be useful (100).

MEDICAL MANAGEMENT

Corticosteroids
Corticosteroids and adrenocorticotropic hormone (ACTH) have been used in MS therapy for years, well before randomized controlled trials were undertaken in MS therapy. An early randomized trial of ACTH versus placebo for exacerbations of MS showed improvements in disability scores at 4 weeks in patients treated with ACTH compared with patients treated with placebo (101). Over the past 20 years, various studies have shown a similar short-term effect of steroid therapy during exacerbations. Three major randomized equivalence studies showed that IV methylprednisolone was similar to ACTH in terms of efficacy with a more convenient dosing schedule (102–104). Questions about the efficacy of IV methylprednisolone lead to three randomized trials versus placebo that reconfirmed the significant short-term benefit of various IV methylprednisolone regimens versus placebo on neurological function (105,106). A systematic review found a sufficient body of evidence to support the use of IV methylprednisolone in acute exacerbations of MS (107). The usual place for IV steroids is for acute exacerbations that cause functional deficits. The use of steroids has to be balanced against the considerable long-term risks which include occasional anaphylaxis against IV methylprednisolone, osteoporosis, cataracts, aseptic necrosis of the hips or shoulders, diabetes mellitus, etc.

Pulsed IV methylprednisolone given several times a year was shown to slow the development of T1 black holes, delay brain atrophy, and slow progression (108). In a trial of pulsed methylprednisolone for SPMS, patients on treatment were significantly less likely to be a treatment failure (59).

Interferons
Three IFN-β agents are FDA approved for MS. They have similar efficacy in reducing relapse frequency and new MRI activity and have similar side effect profiles.

IFN-β1b (Betaseron) was approved by the FDA for use in RRMS and became available in limited quantities in the United States in late 1993 (49). IFN-β1b was tested in patients with up to a moderate degree of disease severity. A three-armed phase III trial clearly demonstrated a dose-effect curve. (Comparison groups placebo, 1.6 million international units, 8 million international units s.c., alternate day dosing.) In the high-dose group, a 34% reduction in exacerbations and an 83% reduction in active MRI lesions were noted. Side effects of IFN-β1b include injection site reactions, elevation of liver function tests, and rare incidences of autoimmune hepatitis, glomerulonephritis, and immune thyroiditis. There may be an increase in depressive symptoms in patients on IFN therapy. Common side effects include flulike symptoms the night after the injection, which can usually be reduced by anti-inflammatory medications or
acetaminophen. Up to 35% of patients may develop neutralizing antibodies to IFN-β1b which seem to inhibit the efficacy of this medication. In patients with more severe functional symptoms spasticity may increase in IFN therapy.

In 1996, IFN-β1a (Avonex) became FDA-approved for RRMS, based upon a phase III trial demonstrating 18% reduction in exacerbations and a 33% reduction in MRI activity (50). In this study, there was a 37% lower risk of sustained progression of disability as measured by the Expanded Disability Status Scale (EDSS). Side effects include flu-like symptoms after the injection, and a lower rate of side effects similar to IFN-β1b. Five to ten percent of patients may develop neutralizing antibodies to IFN-β1a (50).

A third IFN-β is approved for the US market is IFN-β1a, s.c., three times a week (Rebif) (51). This agent demonstrated efficacy similar to IFN-β1b.

Glatiramer Acetate
Glatiramer acetate (GA) (Copaxone) is a mixture of synthetic polypeptides consisting of four amino acids in a specific molar combination. The proposed mechanism is binding to class II major histocompatibility complex with consequent inhibition of myelin-reactive T cells. GA was tested in RRMS patients with an EDSS of up to 5.0 and demonstrated a 29% reduction in exacerbations (52). A subsequent study has shown GA to reduce active lesions on MRI by 35% (109). Side effects include skin injection site reactions, rare arthritic symptoms, and an idiosyncratic reaction in 1/3,000 injections with chest tightness and shortness of breath lasting a few minutes.

Natalizumab (Tysabri)
Natalizumab contains humanized neutralizing IgG4-κ monoclonal antibodies against leukocyte α4 integrins. By blocking α4 integrins, natalizumab reduces the movement of mononuclear leukocytes into the CNS and small intestine. This medication is given as a monthly IV injection. One pivotal trial of natalizumab showed a reduction in sustained disability from 29% to 17% at 2 years, as well as a reduction in relapse frequency by approximately 70% and reduction in MRI new gadolinium enhancing lesions at year 2 by 92% (110). Two MS patients who were also on IFN therapy developed progressive multifocal leukoencephalopathy (PML) (111). These cases led to the initial withdrawal of natalizumab from market in 2004. Subsequently, it was reapproved with a specific monitoring program and consenting procedure for relapsing forms of MS, in particular for patients who have failed other standard agents (112). The risk of PML and other opportunistic infections in patients on natalizumab monotherapy is still unclear.

Mitoxantrone (Novantrone)
Mitoxantrone is an anthracenedione, a chemotherapy agent with potent effects on cellular and humoral immune mechanisms. It is the only FDA-approved chemotherapeutic agent for MS. In a clinical trial for secondary progressive patients with continued attacks, it significantly reduced disability progression, reduced attack rates, and reduced MRI measures (53). A dose-dependent cardiotoxicity limits use of the medication to about 2 years of total dose. In addition, there have been case reports of leukemia due to mitoxantrone in patients with MS (113). Other adverse side effects include bone marrow suppression, amenorrhea, and infertility. This drug is usually prescribed by MS experts or by oncologists experienced in handling this drug.

Other Agents in Trials
Presently, there are about 300 trials listed for MS with clinicaltrials.gov, a National Institute of Health Registry of clinical trials. These include combinations of presently available agents with other medications, medications thought to have neuroprotective effects, medications for symptom management in MS, chemotherapeutic agents, and a variety of novel agents that in general affect immune regulation. Agents with some proven efficacy in MS that are not FDA approved for MS include methotrexate, azathioprine, cyclophosphamide, rituximab, and IVIG. These agents have a moderate toxicity profile, and the data supporting their use are less robust due to smaller trials, less effective blinding, and other confounding factors. A small trial in active severe demyelinating disorders indicated that plasmapheresis was effective in patients who had failed IV steroid therapy and were less than 3 months from the onset of their symptom complex (114). Some studies are beginning to look at specific genetic markers as a potential measure of therapeutic responsiveness. MS is a very active area of research and therapeutics are expected to continue to change dramatically in the next few years.

REHABILITATION
It is generally agreed that rehabilitation is an important component of the management of patients with MS. A recent expert opinion paper published by the National MS Society summarized general recommendations for the use of rehabilitation in MS (115). In practice, rehabilitation professionals are faced with serious challenges when attempting to apply rehabilitative interventions to MS patients.

MS is overall a progressive neurologic disease, which does not follow the traditional model of rehabilitation (acute medical event or injury, functional recovery, and re-adaptation into the environment), with the notable exception of severe disease exacerbations. The best timing for rehabilitation interventions is not well defined. Too often, patients are referred late, when impairments are fixed and opportunities for functional improvement are limited. Payors usually allow a small number of therapy sessions every year when many MS patients have ongoing needs, and demonstration of progress is required to continue therapy, when preventing functional loss is essential in reducing the individual and societal costs of the disease.

Variability and unpredictability make it difficult to apply standard rehabilitation protocols in MS. Clinical presentations are highly heterogeneous, and intricate impairments are an obstacle to functional compensation. Symptoms and functional performance vary over time, even within the course of a day, but not always with a consistent pattern, giving a feeling of “trying to hit a moving target.” Furthermore, transient
worsening of MS symptoms is frequently encountered with exertion, and compromises patients’ adherence to therapy sessions and home exercise programs.

**Evidence Supporting the Use of Rehabilitation in MS**

In the expert opinion paper cited above, the National MS Society acknowledged a need for stronger evidence to support treatment recommendations. There is, however, a growing body of data regarding MS and rehabilitation.

**Neuroplasticity**

CNS plasticity has been more widely assessed after stroke and brain injury, and specific interventions, such as constraint-induced movement therapy or partial body weight supported treadmill training, have yielded interesting results. A recent review of the literature on fMRI in MS highlights evidence supporting the existence of cortical reorganization in MS patients (particularly in the visual, cognitive, and motor systems) (116). Morgen et al. reported a correlation between functional improvements after physical therapy, and changes in functional sensorimotor signaling patterns in the brain (117).

**Aerobic Exercise**

Deconditioning has been identified as a significant contributing factor to fatigue in MS (118). The benefits of exercise on fitness, activity level, subjective fatigue, and perceived health status, have been demonstrated in randomized controlled studies (119,120). In practice, it is often difficult for patients to initiate a sustainable exercise routine without the guidance of a rehabilitation professional, to find the appropriate type, intensity, and duration of exercises, and to encourage the patient through a usually difficult initial period for the first few weeks. Traditional outpatient therapy is appropriate. Short intensive programs, such as the Jimmy Heuga Center CAN DO wellness program, are also an option.

**Recovery after Exacerbations**

Even when treated with high-dose corticosteroids, MS exacerbations often leave patients with residual impairments (121). In fact, exacerbations have been identified as the main cause of accrual of disability in relapsing forms of the disease (122). Inpatient or outpatient multidisciplinary rehabilitation may also enhance functional recovery (123,124). It remains to be determined if neuroplasticity plays a role, or if most of the benefits of rehabilitation in this context stem from optimized adaptation and compensation strategies. Less intensive outpatient rehabilitation may not be as effective (125).

**Rehabilitation of Chronic Disability**

Several controlled trials showed objective and/or subjective benefits from multidisciplinary inpatient (126,127) or outpatient (128) rehabilitation, mostly in patients with a progressive disease course. If we assume that rehabilitation does not affect the disease process, these results suggest that some of the disability progression is not directly related to disease activity.

**Disease-Specific Outcome Measures for MS Rehabilitation**

Generic outcome measures used in rehabilitation, such as the Barthel Index and the Functional Independence Measure (FIM), can be applied to MS patients (126,129). However, these measures are often not sufficient, because they do not cover specific aspects of the disease (e.g., fatigue, visual disturbance), lack sensitivity to small but significant changes over time, or exhibit a ceiling effect (particularly with regard to cognitive disability). Many disease-specific instruments, although often not developed and validated in a rehabilitation setting, can be useful in clinical practice and in research studies (Table 25-1).

<table>
<thead>
<tr>
<th>Scale Name</th>
<th>Acronym</th>
<th>Dimension Assessed</th>
<th>Mode of Administration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expanded Disability Status Scale (212, 213)</td>
<td>EDSS</td>
<td>Impairments/ambulation</td>
<td>Standardized neurological examination/observation of walking performance</td>
</tr>
<tr>
<td>Incapacity Status Scale (212, 213)</td>
<td>ISS</td>
<td>Disability</td>
<td>Observation of performance/interview of patient and/or caregiver</td>
</tr>
<tr>
<td>Environmental Status Scale (212, 213)</td>
<td>ESS</td>
<td>Handicap</td>
<td>Interview of patient and/or caregiver</td>
</tr>
<tr>
<td>Multiple Sclerosis Functional Composite (214)</td>
<td>MSFC</td>
<td>Impairment/disability</td>
<td>Three simple tests (Timed 25-Foot Walk, Nine-Hole Peg Test, Paced Auditory Serial Addition Test—3 min)</td>
</tr>
<tr>
<td>Neuroperformance Testing (215)</td>
<td>NPT</td>
<td>Impairment/disability</td>
<td>Quantitative neurologic tests</td>
</tr>
<tr>
<td>Scripps Neurological Rating Scale (216)</td>
<td>SNRS</td>
<td>Impairment/disability</td>
<td>Standardized neurologic examination</td>
</tr>
<tr>
<td>Ambulation Index (217)</td>
<td>AI</td>
<td>Walking performance</td>
<td>Observation of gait performance on 25 ft</td>
</tr>
<tr>
<td>Multiple Sclerosis Impact Scale-29 items (218)</td>
<td>MSIS:29</td>
<td>Patient-reported walking performance</td>
<td>Patient survey</td>
</tr>
<tr>
<td>Multiple Sclerosis Walking Scale-12 items (219)</td>
<td>MSWS</td>
<td>Patient-reported walking performance</td>
<td>Patient survey</td>
</tr>
<tr>
<td>Multiple Sclerosis Quality of Life Inventory (221)</td>
<td>MSQOL</td>
<td>Patient-reported health status and symptom severity</td>
<td>Patient survey</td>
</tr>
<tr>
<td>Multiple Sclerosis Quality of Life-54 (220)</td>
<td>MSQOL-54</td>
<td>Patient-reported health status and symptom severity</td>
<td>Patient survey</td>
</tr>
</tbody>
</table>
Symptom Management

MS is a condition that leads to multiple symptoms affecting function (fatigue, bladder dysfunction, spasticity) and well-being (pain, paresthesias). Close attention to symptom management may be useful in allowing patients to function at their optimal level, while reducing the amount of discomfort.

Heat Intolerance

Worsening or recurrence of preexisting neurologic symptoms with heat (from a hot environment or from elevated body temperature) is a very common phenomenon, and was even used to support the diagnosis of MS in the past (130). Reduced safety factor for neurological transmission due to demyelinated or partially demyelinated axons is a likely mechanism (see Pathophysiology). Heat intolerance can limit a patient’s potential to participate in rehabilitation and exercise. Adjusting air or water temperature, using fans, cooling vests, or heat extraction units, and the medical recommendation of air-conditioning units in the home, are all practical and valid management tools.

Transient Neurological Events

A common but underrecognized problem in MS is transient neurological events (TNEs). These were reported in one series in as many as 20% of patients (131). These consist of stereotyped, brief (seconds), frequent (up to 100s of times a day) events with neurological symptomatology. Dystonic posturing of a limb, weakness, visual disturbances, dysarthria, abnormal sensations, and muscle spasm can all be seen, often in combination. These events are often not reported by patients without direct questioning, as they do not conform to the general pattern for an attack of MS. They are likely due to epaptic transmission between demyelinated fibers and are not a cortically based seizure phenomenon. They usually occur over a period of weeks and resolve, and may represent a new demyelinating event. Often an explanation is helpful to the patient to understand what is transpiring. If needed, treatment with low-dose antiepileptic medication can be effective (carbamazepine, topiramate, etc.).

Spasticity

Spasticity (defined as a velocity-dependent increase in stretch reflexes (132)) and decreased motor performance (weakness, loss of dexterity), both of which are components of the upper motor neuron syndrome, frequently coexist in MS patients, and represent a major source of disability. Treatment planning, particularly when functional improvement is sought, must take both into account, since interventions may have a different effect on each of them.

In a large survey of over 20,000 participants in the North American Consortium of MS (NARCOMS) Patient Registry, 84% of the subjects noted spasticity-related symptoms, and approximately one-third reported that spasticity was severe enough to affect their daily activities (133). Even though outcome measures and interventions used in other CNS conditions can all be applied to MS patients, evaluating and treating spasticity in MS can be challenging. Due to the presence of associated impairments, and due to the heterogeneity of clinical presentations, it is difficult to evaluate the impact of spasticity (and its treatments) on functional performance and subjective well-being. Furthermore, side effects from symptomatic therapies (particularly sedation and weakness) can make it difficult to reach effective doses. Clinical practice guidelines for the management of spasticity in MS have been developed based on a review of the literature and expert consensus (134).

Spasticity in MS is usually considered of spinal origin, although features of cerebral origin spasticity are also encountered. Patient complaints can be confusing; for example, a patient reporting increasing difficulty bending the legs and sensation of tightness with walking may be experiencing fatigability and abnormal sensations rather than spasticity. Examination will seek static (resistance to passive movement, decreased range of motion) and dynamic (synergy patterns, co-contraction of agonist and antagonist muscles, posture-dependent hypertonia, clonus, spasms) features suggestive of spasticity. It is essential to evaluate the patients at rest and during a functional activity such as walking. Even if the elements above allow for “diagnosis” of spasticity and its severity, they are not sufficient for treatment planning. The impact of spasticity-related phenomena on the patient and caregivers, the patient’s personal characteristics and preferences, and disease-related constraints will all influence treatment decisions. Therefore, outcome measures should include not only symptom severity (stiffness, pain, spasms) and impairment ratings (Ashworth or Tardieu scale, range of motion, muscle strength) but also measures of passive and active function and satisfaction with treatment.

Overall, there is limited published evidence to support the efficacy of most antispasticity agents in MS, and there are methodological concerns with most studies (135). Some of the treatments are used off label for this indication. Controlled clinical trials have demonstrated the efficacy of oral baclofen (136), tizanidine (137), gabapentin (138), benzodiazepines (139), and dantrolene sodium (140) in patients with MS. The benefits usually consisted of symptom relief (e.g., stiffness, spasms, spasticity-related pain), decreased resistance to passive movement, and improved range of motion. Objective functional benefits (e.g., on ambulation), however, were either not assessed or not found. Increased weakness, and sometimes functional deterioration, were reported with baclofen and benzodiazepines (139), but these occurred less frequently with tizanidine (137). Sedation is another major side effect of antispasticity medications, and may not allow titration to effective doses. In practice, the “start low, increase slow” rule must be applied, but high doses can be used if tolerated, and medications can be combined when spasticity is severe, or when dosing is limited by side effects. Reports of liver toxicity with some of these agents (particularly dantrolene, tizanidine, baclofen) must be kept in mind, particularly when patients are on potentially hepatotoxic disease-modifying therapies such as
IFN-β, azathioprine, cyclophosphamide, and mitoxantrone. More recently, levetiracetam was reported to be helpful with phasic phenomena associated with spasticity (141).

Chemodenervation and chemical neurolysis can be used to treat focal spasticity, or diffuse spasticity with focal problems. Onabotulinum toxin A was recently approved by the FDA for the treatment of upper extremity spasticity in several CNS conditions, including MS. There are only a few published clinical trials of botulinum toxin (BT) in MS (142). The advantages of local injections include a lower occurrence and severity of side effects compared to other treatment, and the ability to specifically target muscles. However, criteria for patient selection and dosing algorithms need to be more carefully studied in the context of MS. A high rate of therapy discontinuation associated with disease worsening has been reported (143).

Among the treatments for spasticity requiring a surgical intervention, intrathecal baclofen (ITB) therapy is by far the most frequently used in MS, and there is a fairly large body of evidence regarding its efficacy, side effects, and complications (144–146). ITB therapy is approved by the FDA for the treatment of severe spasticity of spinal or cerebral origin refractory to oral antispasticity medications, or when oral medications are not tolerated. The intrathecal route of administration allows for effective CSF concentrations with markedly smaller doses and plasma concentrations of baclofen. Therefore, a larger effect is achieved with reduced incidence of CNS sedation compared to the oral route. However, the risk of inducing weakness is higher, and for this reason ITB therapy has been used mostly in nonambulatory patients with severe lower extremity spasms and contractures. The benefits of ITB in this patient group should not be underestimated, and include relief of discomfort and pain related to spasticity, enhanced ease of care, improved posture, and improved ability to transfer (146). Some patients are able to improve from a mostly bedbound state to spending most of the day in a wheelchair. In the NARCOMS survey cited earlier, patients receiving ITB reported lower levels of spasticity and higher satisfaction with treatment, compared to patients treated with oral baclofen (133). Recent reports have shown that ITB can be utilized in ambulatory MS patients without loss of function (147,148). Benefits in this patient group include not only relief of spasticity-related symptoms and objective reduction of hypertonia but also subjective improvement of function and quality of life (149) and objective improvement of specific gait parameters (150). Opiates and clonidine have also been used intrathecally to treat intractable pain and spasticity in MS, alone or more frequently in combination with baclofen (151).

**Weakness**

Decreased voluntary motor output in the extremities and trunk can be observed with or without spasticity. Multiple factors account for the observed weakness, including decreased motor control from the CNS; immobility and deconditioning; spasticity; chronic elongation of muscles due to positioning; heat; and fatigue. Increased weakness with effort is thought to be related to nerve conduction blocks in the CNS, but could also be related to muscle fatigue with impaired muscle excitability and metabolism (152). In addition, other symptoms such as pain may inhibit motor performance.

Fluctuations in motor performance are the hallmark of MS, which complicate testing and treatment planning. The interview is essential in differentiating weakness from subjective fatigue or other causes of altered movement and defining the characteristics of weakness (localization, duration, timing, aggravating factors) and its impact. The Medical Research Council manual muscle testing, or even dynamometry (153), can be used to evaluate the severity of muscle weakness. Functional testing is essential, and in a busy clinical practice, may consist of simple tests such as the Timed 25-Foot Walk or the Nine-Hole Peg Test. In a rehabilitation setting, the patient can be challenged with a variety of tests and exercises and/or prolonged effort. In the ambulatory patient, motor fatigue can be evaluated with tests of walking endurance such as the 6-minute (154) or 2-minute (155) walk tests. Most tests of walking performance are performed on level ground, which does not reflect the diversity of terrains encountered in daily life, even within a patient’s home. The Mellen Center Gait Test (MCGT) was recently developed to reproduce the variety of terrains encountered in real-life situations (156).

There are few interventions aimed directly at muscle weakness. Spasticity management, particularly stretching, and functional training are indirectly helpful. Improvement of muscle strength has been reported after aerobic exercise (119) and progressive resistance exercises (157). These exercise programs must be highly customized to the specific limitations of MS patients, particularly in terms of intensity and duration of exercises. Since the increase in body temperature due to physical activity is thought to be at least in part responsible for motor fatigue, pre-exercise cooling may improve performance and increase the benefits of exercise (158). Symptomatic medications for fatigue, such as amantadine and modafinil, may improve motor performance and decrease motor fatigue. Treatments for spasticity could improve muscle strength by reducing the co-contraction of antagonist muscle groups, particularly when administered as local injections. In the context of an exacerbation of MS, high-dose steroids may improve muscle strength as they hasten neurological recovery.

4-Aminopyridine (4-AP) is a potassium channel blocker which facilitates conduction in demyelinated axons. An extended release form of 4-AP was recently approved by the FDA (Ampyra™, dalfampridine) to improve walking in patients with MS, based on significant improvement in walking speed in 2 phase III clinical trials. Beneficial effects on lower extremity muscle strength were also noted (159). The rate of patients responding to the drug in the trials was 35% and 43%. Dalfampridine is administered at the dose of one 10 mg tablet twice daily (approximately 12 hours apart), with or without food. A history of seizure and moderate or severe renal impairment are contraindications to this medication. Side effects occurring at a rate greater than 2% and more frequent with the drug compared to placebo were urinary tract infection, insomnia, dizziness, headache, nausea, asthenia, back pain, balance disorder, multiple sclerosis relapse,
paresthesia, nasopharyngitis, constipation, dyspepsia, and pharyngolaryngeal pain. The risk of seizure (in the absence of history of seizure) is low at the recommended dose. In phase II trials, higher doses of dalfampridine resulted in increased risk of seizure without substantial added clinical benefit.

Assistive devices and orthoses, in addition to their benefits in terms of safety, help compensate for weakness by reducing the energetic demands of activities and by replacing the action of weak muscles. The best illustration is the ankle-foot orthosis (AFO), which improves gait efficiency and safety by compensating for ankle dorsiflexor weakness and often ankle plantar-flexor spasticity. However, the AFO is a passive device and does not compensate for hip flexor weakness, which is a frequent cause of foot clearance deficit in MS. An active orthosis was recently developed under the name of hip flexion assist orthosis (HFAO) in order to address these problems. The HFAO consists of a proximal waist attachment, a medial and a lateral dynamic tension band, and a distal connector that attaches to the shoelaces. All the components are adjustable (Fig. 25-2). An uncontrolled pilot study of the HFAO on 21 ambulatory MS patients showed significant improvement in pain, walking speed, walking endurance, and performance on the MCGT, as well as muscle strength in the “affected” leg (160).

Functional electrical stimulation (FES) is another way to compensate for a lack of strength output in an active manner, by causing the contraction of select muscles to enhance function. In spinal cord injury and stroke, upper and lower extremity FES systems have been more extensively tested and are more commonly used. Several FES devices for foot drop (e.g., Odstock Dropped Foot Stimulator [ODFS], NDI Medical, Cleveland OH [USA]; NESS L300, Bioness Inc., Valencia, CA [USA]; WalkAide, Innovative Neurotronics, Austin TX [USA]) are currently used in MS patients. These devices produce peroneal nerve stimulation during the swing phase to help with foot clearance, and all but one use a heel switch to time the stimulation accurately. There is preliminary evidence suggesting that functional peroneal nerve stimulation increases walking speed and reduces the effort of walking in MS patients (161). Further studies are needed to refine selection criteria, recommendations for use, and patient/clinician expectations.

**Fatigue**

Fatigue is one of the most common symptoms accompanying MS (162). This appears to be most pronounced in the afternoon, and is associated with limited activity in 78% of patients in one survey (131). Fatigue is a subjective symptom defined as “lack of physical and/or mental energy which is perceived by the individual or caregiver as interfering with usual and desired activities” (118). In general, the magnitude of subjective fatigue often does not correlate with an objective observed performance. Fatigue may be a direct effect of immunological activity in the brain, similar to the fatigue associated with viral illness (163). In addition, functional MRI studies have shown that MS patients use more brain areas less efficiently than able-bodied normal controls to achieve the same task (164). This may be the physiological correlate of the complaint of “being unable to multitask,” as well as another cause for MS-related fatigue. Fatigue in MS does not correlate with brain measures of disease burden or atrophy (165). Subjective fatigue usually does not correlate with objective changes in motor or cognitive performance, but does correlate with depression. Fatigue measurement scales (e.g., Fatigue Severity Scale, Modified Fatigue Impact scale) are useful to track outcomes (166).
Clinical practice guidelines for the evaluation and management of fatigue in MS are available (118). A thorough baseline assessment is essential, both to characterize the symptom and to detect treatable causes or contributing factors such as medications, poor sleep, depression, infection, and thyroid dysfunction. Energy conservation strategies constitute the first line of treatment: timing of activities, rest periods and naps, and use of assistive devices to decrease the energy demands of activities. The benefits of exercise were described above; an optimal exercise routine will incorporate stretching, resistance exercise, and cardiovascular training. A recent trial of yoga in MS showed beneficial effects on fatigue over 6 months (167). Cooling techniques address heat sensitivity (168). All of the medications for MS fatigue are used off label. Amantidine, an antiviral agent, showed efficacy in some patients with MS-related fatigue in a small randomized clinical trials (169). Modafinil has been shown to reduce MS-related fatigue in standard doses (100 to 400 mg/day) in a single-blinded trial (170), but these results were not reproduced (171). A variety of stimulant medications have been used empirically but there is little evidence in the literature to support their use. A recent double blind, placebo-controlled trial showed that 650 mg of aspirin twice daily significantly reduced fatigue associated with MS (172).

**Bladder Dysfunction**

Bladder dysfunction affects up to 90% of MS patients (173) and correlates with disease severity and disability, but not with disease duration (174). Urinary symptoms affect daily activities, employment, social life, and quality of life in general. Furthermore, urinary tract infections, usually secondary to urinary retention, can cause worsening of MS symptoms and increased disease activity.

Although there are wide variations in the reported prevalence and incidence of urinary symptoms in MS, urgency and frequency are consistently noted as the most frequent, followed by incontinence, hesitancy, and retention (175). While the presence of urinary symptoms is highly correlated with abnormal test results, bladder dysfunction can often be found in the absence of complaints. This is particularly true with urinary retention. Workup for neurogenic bladder dysfunction usually includes urinalysis and urine culture, post-voiding residual (PVR) volume measurement, and urodynamic testing. A renal ultrasound and other upper urinary tract imaging studies may be ordered, particularly in patients with detrusor-sphincter dyssynergia (DSD) or indwelling catheter, but upper urinary tract complications are relatively rare in MS compared to other neurologic conditions (174). Education, with teaching of “bladder hygiene” (e.g., adequate fluid intake, avoiding bladder stimulants), is often helpful in reducing urinary symptoms, particularly those related to detrusor hyperreflexia. Other management options include pelvic exercises, medications, catheters (preferably intermittent catheterization [IC], sometimes indwelling catheter), and less frequently surgical interventions.

Detrusor hyperreflexia, the most common bladder function disorder in MS, is often treated with anticholinergic medications, such as hyoscyamine, oxybutynin (Ditropan), or tolterodine (Detrol). Side effects constitute the main limitation to the use of these medications, and may be less severe with extended-release formulations (176), or with the transdermal formulation of oxybutynin. More recently released anticholinergic medications include trospium chloride (Sanctura), solifenacin succinate (Vesicare), and darifenacin (Enablex). Desmopressin acetate (DDAVP) nasal spray or tablet may be useful in reducing nocturia and enuresis in MS patients (177). BT injections in the detrusor muscle can decrease detrusor hyperactivity for up to 9 months, but may cause temporary retention requiring IC (178). BT injections in the detrusor may be a useful complement to IC when patients experience incontinence between catheterizations. Surgical treatment options are considered when conservative management failed, and include augmentation cystoplasty (in most cases with an abdominal catheterizable stoma) and sacral denervation. Encouraging results have been reported with sacral root stimulation in carefully selected MS patients (Interstim device) (179).

DSD consists of sphincter contraction coinciding with detrusor contraction, leading to incomplete bladder emptying. DSD often occurs in combination with detrusor hyperreflexia; therefore it is strongly recommended to measure the PVR before initiating treatment with anticholinergics in patients who experience urgency and incontinence but do not already perform IC. Alpha-1 selective antagonists (e.g., terazosin) may improve bladder emptying. There are anecdotal reports of the efficacy of BT injections in the sphincter in MS patients with DSD. A randomized placebo-controlled trial failed to show a significant change in PVR after a single injection of BT-A in the sphincter, although voiding volume and detrusor pressures were significantly improved (180). Sphincterotomy or the placement of a urethral stent can be performed, for example when IC is not feasible, but requires adequate detrusor contractility to be fully effective. Another cause of urinary retention is detrusor hypocontractility (sometimes associated with detrusor hyperreflexia), which is more frequent in MS than true detrusor areflexia. IC is the most common treatment for detrusor hypocontractility. Bethanechol (Urecholine) is usually not effective in MS. Surgical interventions for urinary retention include the placement of a suprapubic catheter, which may be easier to manage than an indwelling Foley catheter, and incontinent urinary diversion.

**Bowel Dysfunction**

Constipation is the most commonly reported symptom of bowel dysfunction in MS, followed by fecal urgency and incontinence (181). The pathophysiology of neurogenic bowel dysfunction is not well elucidated. There are many factors contributing to bowel dysfunction, including immobility, inadequate diet and fluid intake, and side effects of medications. Patient and caregiver education is key to a successful bowel management program. In addition to remediating the
Sexual Dysfunction

Sexual dysfunction is frequent in MS patients, who often wish to remain sexually active (182). Sexual dysfunction was found to be correlated with lesion burden in the brain (183). Other consequences of MS, including concomitant neurologic symptoms (fatigue, decreased sensation, spasticity), depression, and marital relationship problems, often contribute to sexual dysfunction. Men usually report erectile dysfunction (ED), decreased libido, delayed ejaculation, and impaired genital sensation. The most common complaints in women include fatigue, anorgasmia, and decreased vaginal sensation. A thorough interview and examination, including the evaluation of perineal sensation, and in women pelvic examination, will guide treatment planning. Complementary testing is not often performed. Education and counseling are useful in both sexes, as a good relationship is key to intimacy. Phosphodiesterase-5 inhibitors (sildenafil [Viagra], vardenafil [Levitra], and tadalafil [Cialis]) are very popular in the treatment of male ED, although these medications do not necessarily resolve all the problems involved in sexual dysfunction. Other, less often used approaches include vibratory stimulation, vacuum pumps, papaverine or prostaglandin E1 injections, and implanted penile prostheses. Pharmacological treatment options for female sexual dysfunction are more limited. The results of a small double-blind trial of sildenafil in women with MS yielded disappointing results, with only a significant improvement of vaginal lubrication (184). The management of other contributing symptoms, vibratory stimulation, and adequate sexual positioning, can be helpful.

Tremor

Tremor in MS is often associated with ataxia, and may include rest, postural, and action components (“rubral tremor”) (185). The pathophysiology of this very debilitating symptom is not completely elucidated, although the role of the thalamus has been more extensively studied. Tremor severity and the outcome of interventions have often been measured with Fahn’s Tremor Rating Scale. Rehabilitation interventions include the use of assistive devices and technologies, and wrist weights (which can be difficult to use if weakness and fatigability are present) (186). Several medications were reported to be effective on MS tremor, including isoniazid, glutethimide, primidone, levetiracetam, carbamazepine, oral tetrahydrocannabinol, clonazepam, and propranolol, but in practice the results of pharmacological treatments are disappointing in the majority of cases. Thalamotomy, (187) and more recently deep brain stimulation, (188) have been used in MS patients with severe tremor. Positive results were reported in studies of small cohorts of MS patients, but patients were not often followed after 12 months. A rapid decrease in effect leading to the need for more frequent stimulator adjustments has been reported, as well as upper extremity weakness concomitant to the reduction of tremor.

Pain

The old saying “MS doesn’t cause pain” is contradicted by clinical experience and numerous publications. Up to 81% of MS patients experience pain at sometime in the course of their disease (189). Pain is often multifactorial and multifaceted in MS. It is sometimes thought to be a direct consequence of the disease process, such as neuropathic pain and pain related to inflammation and upper motor neuron damage. Musculoskeletal pain and headaches are also common.

Neuropathic Pain

Neuropathic pain is frequently encountered in MS (190). It may present as dyesthesia, allodynia, and/or neuralgia, and is often associated with other sensory disturbances. Partial TM can be associated with a painful band sensation around the trunk, sometimes called the “MS hug.” Chronic dysesthetic pain is often described as burning, but patients are sometimes at a loss when trying to characterize their pain. The pain is thought to originate from lesions in the CNS. Anticonvulsants and antidepressants (tricyclic and more recently selective serotonin reuptake inhibitors (SSRIs)) are commonly used to treat neuropathic pain in MS, despite a lack of solid published evidence to demonstrate their efficacy in this population. The potential side effects of these medications, particularly sedation, should not be overlooked (191). The use of opioids remains controversial (192). A recent randomized clinical trial and open-label extension study have shown that cannabinoids are effective in relieving neuropathic pain in MS (193,194), but their use is limited by side effects and legal considerations. Trigeminal neuralgia is experienced by up to 10% of MS patients in the course of their disease, usually involving the second and third divisions (195). This syndrome is believed to arise from ephaptic transmission of nerve conduction, which occurs with demyelination in the area around the trigeminal nerve entry zone. Carbamazepine and other anticonvulsants are usually effective. When the pain is refractory to oral medications, treatment options include glycerol injections, balloon ablation of the trigeminal nerve, and radiofrequency thermal or surgical rhizotomy.

Pain Related to Upper Motor Neuron Damage

Spasticity can be associated with pain. Nocturnal painful spasms have been described in MS (196). These are defined as transient painful extensor or flexor spasms of the legs, lasting seconds to minutes, often occurring at night. These may significantly interfere with sleep and increase daytime fatigue. An open-label trial of gabapentin at night showed that 20 of 22 patients experienced reduction in their nocturnal painful spasms with an acceptable side effect profile (196).

Pain Related to Inflammation

Retro- and periocular pain, aggravated or triggered by eye movement, often occurs in the context of ON and usually improves
with corticosteroid treatment. Retroorbital pain without visual disturbance is of unclear significance and may respond to steroidal or nonsteroidal anti-inflammatory agents.

Musculoskeletal Pain
Back, neck, upper and lower extremity pain of musculoskeletal origin are commonly encountered in the general population, and as a consequence may occur concomitant to MS, particularly as a result of aging. This issue should not be neglected, because it represents a source of added disability, discomfort, and medical complications. Indeed, MS and its treatments increase the risk of musculoskeletal problems. For example, abnormal posture and body mechanics as a result of weakness, spasticity, and/or loss of coordination, may cause excessive joint stress. Falls related to neurologic impairments can cause fractures and other injuries. Corticosteroid treatments increase the risk of aseptic necrosis of the femoral head, and in combination with immobility are a risk factor for osteoporosis and subsequent complications. Even if the diagnosis of musculoskeletal problems is straightforward, confusion may arise when it is combined with other types of pain, or when several etiologies are possible. For example, radicular pain can arise from demyelination as well as nerve root compression. When in doubt, additional testing and referral to a musculoskeletal specialist are helpful.

Headache
Although headache is a common complaint among MS patients, its relationship to MS is unclear. In one survey, headaches were no more common in MS patients than in age-matched controls (131). In another, there was a higher prevalence of headaches among MS patients over controls (197). Focal facial pain syndromes (e.g., trigeminal neuralgia) and retroorbital pain, particularly with eye movements, are likely to be MS-specific. Since MS patients often take multiple medications, headache might be related to drug side effects. In general, headache should be treated on its own merits, even if it is in some fashion related pathophysiologically to MS activity.

Visual Impairment
Visual disturbance occurs in a little over one fourth of all MS patients, and often fluctuates, similar to other MS symptoms, making it more difficult for the patient to adapt. Unfortunately, there is no treatment for the residual visual loss after ON. Even when visual acuity is satisfactory, contrast sensitivity and color perception may be impaired. Night driving may be compromised. Patients should undergo yearly eye examination surveillance for conditions amenable to treatment or correction. The use of steroids may increase the risk of developing cataracts, and uveitis is encountered in 1% of patients with definite MS. Large-print reading materials and computer-based magnification systems may be helpful in patients with severe visual compromise. Patients with photophobia will benefit from the use of sunglasses. Corrective surgery is rarely recommended in MS patients with diplopia. Alternative interventions include the use of an eye patch and prisms. Oscillopsia sometimes responds to anticonvulsants or baclofen (198).

Depression
Depression is a common problem in MS. A large population survey of 739 MS patients in King County, Washington, found the point prevalence of significant depressive symptoms to be 41.8% (199). Suicide is an unfortunate but not uncommon phenomenon in the MS population (200). Depression may be increased with IFN therapy for MS, but treatment appears to be effective despite this (201). Depression often contributes to fatigue and perceived cognitive difficulties. Monitoring for the presence of depression as well as a proactive approach to its treatment is crucial. Despite a lack of large placebo-controlled clinical trials in MS, antidepressant medications are widely used. Some antidepressants are considered “energizing” (e.g., paroxetine [Paxil], bupropion [Wellbutrin]) and may be of interest in the presence of severe fatigue. A recent placebo-controlled randomized clinical trial of paroxetine for major depression in MS patients showed no significant effect (202). Behavioral interventions such as counseling are often recommended.

Cognitive Impairment
Cognitive impairment represents a significant problem in persons with MS. Comprehensive neuropsychological (NP) testing suggests that over 40% of MS patients show cognitive impairment to at least some degree (203). Common complaints include forgetfulness, decreased attention/concentration, difficulty with multitasking, and “cognitive fatigue.” While some patients will exhibit obvious cognitive deficits and may meet the diagnostic criteria for dementia, in many cases the standard interview and testing during a routine visit will not detect significant abnormalities. Comprehensive NP testing typically demonstrates selective deficits in attention, information-processing speed, working memory, verbal and visuospatial memory, and executive functions (204). Disease-specific screening tools and short batteries have been validated (205) and single tests such as the Paced Auditory Serial Addition Test and the Symbol Digit Modalities Test have been suggested as outcome measures in clinical trials (204). Advanced imaging techniques (e.g., functional MRI, diffusion tensor imaging) may provide a better insight into the physiological and anatomical correlates to cognitive impairment than standard MRI.

The management of cognitive impairment often starts with educating (and often reassuring) the patient about their specific deficits and addressing contributing factors, such as depression, fatigue, heat intolerance, or medications. Sleep disorders, particularly obstructive sleep apnea, should not be overlooked. Rehabilitation interventions typically rely on the use of compensatory strategies (e.g., pacing, memory aides) and environmental adaptations, but there is an emerging literature on cognitive training in MS. Acetycholinesterase inhibitors used to treat Alzheimer’s disease (e.g., donepezil [Aricept], memantine [Namenda]), although not FDA-approved in MS, are used and may be helpful (206).
Speech and Swallowing Problems
More than one in five MS patients report problems with speech or communication (207). Dysarthria, decreased voice intensity, cognitive impairments, fatigue, and depression, may all contribute to decreased communication. Dysphagia is also relatively frequent, and may range from intermittent mild problems to chronic aspiration requiring percutaneous endoscopic gastrostomy (PEG) tube placement. Addressing these issues usually requires referral to a speech-language pathologist, and psychologist if indicated (208).

SOCIAL, VOCATIONAL, AND FINANCIAL MANAGEMENT
Similar to many chronic conditions, MS affects the interaction between the patient and the outside world (209). Caregiver burden is increased by the presence of psychiatric and cognitive impairments (210). Even if rehabilitation professionals are used to help patients and their families deal with the social, vocational, and economic consequences of disabling conditions, some characteristics of MS complicate management: the disease is typically diagnosed in early adulthood, when the patients are still in the process of building their personal and professional life; the evolution of the disease is unpredictable and symptoms fluctuate, making it difficult for the patient to adjust and plan ahead; disease-modifying therapies are expensive, and even with health insurance, an increasing portion of the cost is left to the patient; and functional limitations often result from a combination of impairments and subjective symptoms which may not be fully ascertained by a standard neurologic examination. Even if maintaining employment is desirable for social, psychological, and financial reasons, the worsening of MS symptoms with stress at work, the difficulty of getting reasonable accommodations in the workplace, and the negative impact of fatigue from work on family and leisure activities, may force the patient to stop working and apply for disability benefits (211).

WELLNESS
Patients with MS will develop the same pathologies that afflict the general population. In addition, MS (and sometimes treatments for MS) may increase the risk for secondary medical problems, which in turn may have an impact on MS symptoms. For example, urinary tract infections due to neurogenic bladder often cause a worsening of MS symptoms; osteoporosis decreases mobility; and steroid use increases the risk of fractures. However, individuals with MS have a tendency to neglect systematic screening for common health problems and to attribute any new symptom to MS. Therefore, it is essential that MS patients see a primary care physician on a regular basis and be referred to the appropriate specialty when needed. Also, basic wellness recommendations (e.g., diet, exercise) can help improve general health as well as MS symptoms.

CONCLUSION
MS is a complex, long-term, unpredictable, functionally and emotionally taxing disease. It is arguably one of the most difficult neurological diseases to manage. Needs vary widely between individuals and over time, and there are many circumstances in which a comprehensive approach involving rehabilitation may be beneficial. Simple rehabilitative strategies such as recommending a rollator or an AFO may make the difference between walking and not walking. However, rehabilitation is still underutilized in MS, and MS rehabilitation is not as well-defined as spinal cord injury or stroke rehabilitation. The principles commonly used for disease management can and should be applied to symptom management and rehabilitation: diagnose (or identify the problem), start appropriate treatment early, and monitor and adjust the treatment strategy over time.

REFERENCES


[Text continues with additional references]


Rehabilitation of Persons with Parkinson’s Disease and Other Movement Disorders

INTRODUCTION

Parkinson’s disease affects roughly 1% of the population older than age 50 in the United States, incidence increases with age, and the peak onset occurs in the sixth to eighth decade, making it one of the most common chronic diseases of adulthood (1). Life expectancy is near normal; however, there is increased risk for morbidity and mortality due to associated complications, such as falls, orthostatic hypotension, or aspiration pneumonia.

Rehabilitation has a vital and unique role in movement disorders. Parkinson’s disease is a chronic, progressive, neurological disease, which at various stages and manifestations affects all aspects of a patient’s functional state. The goal of this chapter is to review the clinical presentation and neuropathology of Parkinson’s and other movement disorders, with specific emphasis on the collaborative process of rehabilitation. The disablement process in Parkinson’s is based, just as in other medical and neurological disease states, on the continuum of impairments and how they affect functional limitations and subsequent disability (2,3). The impact of multifaceted treatment is covered in this chapter as well, placing high importance on the role of exercise with and without pharmacological management.

The shaking palsy or paralysis agitans was the description James Parkinson used in 1817 to represent what would later be referred to as “Parkinson’s Disease,” the progressive, central nervous system impairment that is characterized by the presence of two or more cardinal signs: bradykinesia, resting tremor, rigidity, and postural instability. The diagnosis of Parkinson’s disease can be challenging and is first and foremost a clinical diagnosis. The typical individual afflicted with Parkinson’s disease presents with symptoms on one side of the body, referred to as asymmetric onset. Typically, individuals with Parkinson’s disease can be subdivided into two groups: those with a predominance of rigidity and akinesia, with minimal or no rest tremor; and those with rest tremor predominant disease. However, other signs and symptoms may be present as well (Table 26-1). Parkinson’s disease can be incorrectly diagnosed as much as 25% of the time if based solely on clinical features and not on diagnostic trials, specifically responsiveness to dopamine treatment (4,5). The rate of progression of Parkinson’s varies significantly from one person to another.

Although the exact cause of Parkinson’s is not known, environmental and genetic factors have been implicated. One theory regarding the pathogenesis of Parkinson’s disease is that unidentified environmental triggers cause accelerated destruction of the dopamine-producing neurons in the substantia nigra in a genetically predisposed individual. Secondary parkinsonism refers to those disorders that have the same combination of signs and symptoms that are seen in primary Parkinson’s disease but result from a clearly identifiable insult to the central nervous system. Examples of such insults include exposure to 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP) (6–8), which causes the death of dopamine neurons and ischemic infarcts involving the vascular supply to the midbrain that can also result in the death of dopamine neurons. Dopamine-antagonist medications, (specifically antagonists of D2 receptors), such as the typical neuroleptics haloperidol and thioridazine, and other D2 antagonists, such as metoclopramide, have been known to exacerbate or even induce parkinsonism. Infectious conditions, such as St. Louis encephalitis, and von Economo’s and other influenza-related encephalitides, have also been associated with secondary parkinsonism.

ANATOMY AND PATHOPHYSIOLOGY

Motor impairment in Parkinson’s and other movement disorders stems from a common principle: There is a disruption in the modulation of the cortical and subcortical structures of the brain. The anatomic structures involved in the control of movement are referred to as pyramidal and extrapyramidal structures, or motor cortices and the basal ganglia, respectively (9–12). A basic understanding of these structures and their physiology forms the basis of symptom-driven therapy in the management of patients with movement disorders.

The several interconnected nuclei of the basal ganglia make up a large portion of the extrapyramidal system. The
FIGURE 26-1. Circuitry of direct and indirect pathways and pharmacology of the basal ganglia. The direct pathway projects gamma-aminobutyric acid (GABA)-mediated inhibitory input directly to the output nuclei of substantia nigra reticulata (SNr) and globus pallidus interna (Gpi). The indirect pathway sends an inhibitory GABA projection to the globus pallidus externa (Gpe), which subsequently projects GABA-mediated inhibitory input on the subthalamic nucleus (STN). The STN output to the SNr/GPi is an excitatory glutamate projection. Solid arrows reflect excitatory pathways and dashed arrows, inhibitory pathways. DA, dopamine; D1, D1 dopamine receptor; D2, D2 dopamine receptor; DYN, dynorphin; ENK, enkephalin; PPN, pedunculopontine nucleus; SNc, substantia nigra zona compacta; SP, substance P. Modified from Alexander GE, Crutcher MD. Functional architecture of the basal ganglia circuits: neural substrates of parallel processing (10).

TABLE 26.1 Clinical Presentations of Parkinson’s Disease

<table>
<thead>
<tr>
<th>Clinical Signs and Symptoms of Parkinsonism</th>
<th>Clinical Features</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Positive phenomena</strong></td>
<td></td>
</tr>
<tr>
<td>Tremor</td>
<td>Most common symptom. Distal involvement frequency: 3.5 Hz. Suppressed by sleep and activity, increased by fatigue or stress. EMG: rhythmic alternating bursts in agonist and antagonist muscles.</td>
</tr>
<tr>
<td>Rigidity</td>
<td>Increase in muscle tone during passive limb movement equal through entire range of motion; increases if contralateral limb is engaged in volition.</td>
</tr>
<tr>
<td>Flexed posture</td>
<td>Dominance of progravity flexor muscles (bowed head, chin toward chest, kyphotic thorax, protracted shoulders, internally rotated arms, flexed elbows, knees, and hips).</td>
</tr>
<tr>
<td><strong>Negative phenomena</strong></td>
<td></td>
</tr>
<tr>
<td>Bradykinesia</td>
<td>Slowness of movement, masked facies, decreased eye blinking, inability to move. Fatigue. EMG: delayed motor unit recruitment, pauses once recruited, inability to increase firing rate.</td>
</tr>
<tr>
<td>Loss of postural reflexes</td>
<td>Tendency to fall to the side (lateral pulsion) or backward (retropulsion); sitting en bloc (collapses in the chair when attempting to sit down).</td>
</tr>
<tr>
<td>Postural instability</td>
<td>Transient inability to move.</td>
</tr>
</tbody>
</table>

EMG, electromyography; Hz, hertz.

REHABILITATION AND COLLABORATION

Neurologists and, to some extent, primary care physicians largely oversee the care of persons with Parkinson’s and medically manage the patient. What has become increasingly clear in neurological and neuromuscular disease management is the necessity of the multidisciplinary approach, in order to offer holistic and comprehensive medical and functional care of the patient. A rehabilitation team should ideally incorporate neurology, physiatry, physical and occupational therapy, speech and language pathology, social work, psychology and/or psychiatry, and nutrition. It, of course, promotes a patient-centered approach with collaboration of all disciplines with the expertise to diagnose and treat various impairments (14).

Physiatrists often see patients with movement disorders, mainly those with Parkinson’s, in consultation. This practice may be in the form of a multidisciplinary clinic with neurologists, therapists, social workers, and psychiatrists or psychologists, or in a stand-alone physiatric practice. Physiatrists can be helpful for specific therapy prescription and analysis and treatment of gait abnormalities. Although a relatively low number of studies address the optimal therapeutic management plan, a collaborative effort by physiatrists and neurologists should logically provide the best therapeutic structure for the movement disorders population. Physiatrists who treat patients with various neurologic diagnoses should be able to assume a more active role in managing movement disorders.

The role of physiatry may be to assist with nonoperative orthopedic problems or basic musculoskeletal and pain impairments typical in this population. Physiatrists in the outpatient and inpatient settings collaborate and communicate with physical therapists, occupational therapists, and speech and language pathologists in order to optimize the function of the patient with a movement disorder. Therapists who subspecialize in Parkinson’s treatment and other movement disorders are invaluable links between patient and physician. Therapists offer not only highly skilled assessments and treatments of functional impairments but also longitudinal education of the patient and family that is essential for carryover and functional gains.

Therapists and physiatrists may encounter a person with Parkinson’s admitted to the rehabilitation unit for another condition, such as stroke, or after an orthopedic procedure or after the implantation of deep brain stimulators, which are used to treat some of the complications of Parkinson’s disease (discussed later in this chapter) (15). These patients pose a unique challenge to the rehabilitation team, who are cognizant of baseline parkinsonian impairments and the additional neurological, orthopedic, or other impairment (16). One of the vital roles of the rehabilitation team is to help the patient and his or her primary medical or surgical team understand how the new illness or procedure will likely influence his or her preexisting parkinsonian impairments.

The team designs a rehabilitation program in order to set appropriate therapeutic goals. The deconditioning syndrome in and of itself certainly affects the hospitalized patient with a movement disorder while in the acute-care setting and can impact upon the patient’s short- and long-term functional goals.

Nursing staff should be aware of the key importance of timely administration of anti-Parkinson drugs in order to minimize fluctuations in blood levels. Optimal pharmacological management impacts proper daytime participation with therapists and exercise (17). Those with Parkinson’s are at risk for nutrition impairment, constipation, and low-energy states, and therefore a nutritionist or dietitian is required to focus on preparing the appropriate caloric and supplement needs for the patient.

IMPAIRMENTS AND FUNCTIONAL LIMITATION IN PARKINSON’S

The concept of disability or disablement has been demonstrated in various ways, with the two most common being the International Classification of Impairments, Disabilities and Handicaps (ICIDH), and the Nage (Saad Z. Nage) framework in the 1960s. Nage introduced the principle of “disablement,” which is a process initiated with pathology leading to impairment that causes function limitation and in turn disability. What is intriguing about this framework is that disability can be positively affected at all four conceptual levels of the continuum. In Parkinson’s, pathology resides in the substantia nigra at the cellular level and results in the inability to create sufficient dopamine. One’s impairment may be muscle rigidity or bradykinesia, for instance, with functional limitation manifesting in gait dysfunction. Disability is the inability to optimally participate in community activities or carry a job, which occurs due to the functional limitation (2,3).

Physiatrists and the rehabilitation team assist in the care of persons with Parkinson’s by managing the various impairments that result from the condition (Tables 26-2 and 26-3).

<table>
<thead>
<tr>
<th>TABLE 26.2</th>
<th>Impairments Resulting From Parkinson’s Disease</th>
</tr>
</thead>
<tbody>
<tr>
<td>Motor</td>
<td>Gait</td>
</tr>
<tr>
<td></td>
<td>Movement initiation and execution</td>
</tr>
<tr>
<td></td>
<td>Bradykinesia</td>
</tr>
<tr>
<td>Tremor</td>
<td>Rigidity</td>
</tr>
<tr>
<td>Sensory/pain</td>
<td>Autonomic dysfunction</td>
</tr>
<tr>
<td></td>
<td>Orthostatic hypotension</td>
</tr>
<tr>
<td>Constipation</td>
<td>Cognitive</td>
</tr>
<tr>
<td></td>
<td>Difficulty in switching from one task to another</td>
</tr>
<tr>
<td>Depression</td>
<td>Anxiety</td>
</tr>
<tr>
<td>Psychotic features, such as hallucination</td>
<td></td>
</tr>
<tr>
<td>Gastrointestinal</td>
<td>Dysphagia</td>
</tr>
<tr>
<td>Constipation</td>
<td>Bladder dysfunction</td>
</tr>
<tr>
<td>Sexual dysfunction</td>
<td></td>
</tr>
</tbody>
</table>
### TABLE 26.3 Therapeutic Plan for Patients With Parkinson’s Disease

#### Medical and Nursing
- Firm bed to decrease contractures and improve bed mobility
- Gradual changing of positions, elastic stockings, abdominal binder, sodium tablets, and possibly pseudoephedrine, midodrine, and/or fludrocortisone for orthostasis
- Regular meals with proper diet (low protein); nutritional consultation
- Measure vital capacity and enforce incentive spirometry to prevent atelectasis and pneumonia
- Bowel program for gastrointestinal hypomobility (stool softeners, bulk-forming agents, cisapride, and suppositories may be required)
- Bladder evaluation and urodynamics; anticholinergics (e.g., oxybutynin chloride [Ditropan]) for hyperreflexic bladder
- Artificial tears for lack of blinking
- Sexual dysfunction evaluation
- Anticholinergic medications before mealtime to help facilitate oral and pharyngeal movements

#### Physical Therapy
- Relaxation techniques to decrease rigidity
- Slow rhythmic rotational movements
- Gentle range-of-motion and stretching exercises to prevent contractures, quads and hip extensor isometric exercises
- Neck and trunk rotation exercises
- Back extension exercises and pelvic tilt
- Proper sitting and postural control (static and dynamic); emphasize whole body movements
- Breathing exercises stressing both the inspiratory and expiratory phase
- Functional mobility training, including bed mobility, transfer training, and learning to rise out of a chair by rocking; may require a chair lift
- Stationary bicycle to help train reciprocal movements
- Training in rhythmic pattern to music or with auditory cues such as clapping may help in alternating movements. Standing or balancing in parallel bars (static and dynamic) with weight shifting, ball throwing
- Slowly progressive ambulation training (large steps using blocks to have patients lift legs, teaching proper heel-to-toe gait patterns, feet 12–15 in. apart, arm swing; use inverted walking stick, colored squares, or stripes as visual aids)
- Use of assistive devices (may need a weighted walker)
- Aerobic conditioning (swimming, walking, cycling)
- Frequent rest periods
- Family training and home exercise program

#### Occupational Therapy
- Range-of-motion activities of upper extremity with stretching
- Fine motor coordination and training, hand dexterity training using colored pegs or beads
- Hand cycling to help train reciprocal movements
- Rocking chair to help with mobilization
- Transfer training
- Safety skills
- Adaptive equipment evaluation, including Velcro closures, raised toilet, grab bars, eating utensils with built-up handles, and key holders
- Family training and home exercise program

#### Speech
- Deep breathing and diaphragmatic breathing exercises
- Articulatory speech training for dysarthria
- Facial, oral, and lingual muscle exercises
- Swallowing evaluation, including a modified barium swallow as needed
- Teaching compensatory strategies for safer swallowing

#### Psychology
- Psychological support for patient, family, and caregivers
- Cognitive assessment

---

This section highlights ways that rehabilitation professionals can help minimize the impact that these impairments, and therefore function limitations, have upon well-being and disability of those with Parkinson’s.

**Gait Dysfunction**

Gait dysfunction is the primary presenting symptom in as many as 18% of patients with Parkinson’s disease (18,19). Persons with Parkinson’s report that one of the leading causes for diminished quality of life is difficulty walking (20). Shortened step length is the focal disturbance of gait in a person with Parkinson’s, and the classic gait pattern is characterized by shuffling (21,22). Reduced hip, knee, and ankle flexion angles, decreased arm swing, limited trunk rotation, and increased forward flexed trunk are typical kinematic changes in gait among persons with Parkinson’s disease. Parkinsonian gait becomes progressively inefficient and unstable as a result of multiple impairments (23,24). There are two stereotypical variations of parkinsonian gait patterns. Many persons demonstrate changes in the fluidity of gait with episodes of hesitation to initiate gait or instances of “freezing” when the person stops and cannot initiate movement while walking (25). These difficulties in performing efficient gait patterns are often observed when the person is turning around, walking through a doorway threshold, or stepping around or over objects (23,26,27). The second parkinsonian gait pattern is festination, characterized by small steps increasing in speed and frequency with a forward trunk posture (26). These gait abnormalities place persons with Parkinson’s at varying risk for loss of balance or falls (24,28).

Although people with Parkinson’s may be able to perform straight-line walking, their overall gait is hampered by hypokinesia, rigidity, postural imbalance, and the fear of falling, especially in advanced stages of the disease (24,29–31). Deficits have also been implicated, thus aggravating the problem and may be the reason why people with Parkinson’s disease rely strongly on visual cues (32,33). The loss of proprioceptive feedback from the knee extensor load receptors is suspected to cause a reduction in the activation of leg extensor muscles (34). Other problems that contribute to the gait abnormality include impairment in programmed muscle activation in the lower limbs and diminished postural reflexes (33). These impairments result in progressive difficulty adapting postural responses and stability during gait, especially on uneven surfaces (24,33,35–37).

During early stages of Parkinson’s, primary impairments of hypokinesia, rigidity, proprioceptive deficits, decreased postural reflexes, and movement execution may have subtle effects on gait. During progression of the disease, the primary impairments are typically magnified and result in new secondary impairments (Table 26-4) (38). Muscle weakness due to the disuse and contractures from rigidity are examples of secondary impairments that further contribute to gait dysfunction. These abnormalities affect balance, alter gait, and increase the patient’s risk for falls (24,28). Another factor contributing to gait instability is orthostatic hypotension, which will be discussed further in this section.

Gait dysfunction in patients with Parkinson’s may be due to either the disease itself or as a side effect of pharmaceutical therapy (see Table 26-4). Levodopa therapy affects parkinsonian gait and movement, in that although the medication may facilitate gait during the “on” phase, deterioration may be seen as part of motor fluctuations during the “off” state. Its effect is largely due to the control of force and amplitude of limb movement, rather than improved automaticity or rhythmicity (39). Levodopa may also cause dystonia, which can negatively affect ambulation.

Direct and secondary impairments of Parkinson’s can lead to further decline if patients respond to these impairments and limitations by decreasing activity levels. This results in further muscle weakness, joint contractures, and worsened endurance. Discussion with the rehabilitation team and prompt referral to therapy can facilitate early treatment approaches to “break the cycle” and prevent resulting functional limitation in gait and other activities. Advances in technology are exploring ways to help monitor changes in gait and functional mobility in the home environment among persons with Parkinson’s using wearable sensors (40,41). In the future, such devices should offer clinicians the ability to titrate medication and make referrals to appropriate therapies based on these objective data.

### Table 26.4 Gait Disturbance in Parkinson’s Disease

<table>
<thead>
<tr>
<th>Gait Disturbance</th>
<th>Parkinsonian Direct Impairment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shuffling gait pattern: Decreased step and stride length; decreased cadence and velocity</td>
<td>Bradykinesia and hypokinesia; movement initiation and execution impairment</td>
</tr>
<tr>
<td>Festination</td>
<td>Impaired automatic motor task performance; movement execution impairment; impaired postural responses</td>
</tr>
<tr>
<td>Stoop (flexed) posture</td>
<td>Rigidity</td>
</tr>
<tr>
<td>Freezing, “start hesitation”</td>
<td>Impaired automatic motor task performance; hypokinesia; movement initiation and execution impairment</td>
</tr>
<tr>
<td>“Cautious” gait (fear of falling)</td>
<td>Impaired postural responses; rigidity</td>
</tr>
<tr>
<td>Impaired balance and unsteadiness due to lightheadedness</td>
<td>Impaired postural responses; orthostasis; ANS impairment</td>
</tr>
<tr>
<td>Dystonia, dyskinesia</td>
<td>Medication effect; Parkinson’s disease—direct impairment related movement disorder</td>
</tr>
</tbody>
</table>
Movement Initiation
The person with Parkinson's typically possesses slow movement initiation that may result from delayed activation of the motor cortex, impeding one's ability to initiate and execute normal movement (42). Although the corticospinal system is intact, the abnormal motor commands result in bradykinesia in the implementation and execution of a motor task. These impairments are exacerbated with high-complexity tasks. Highly refined motor activities such as handwriting require switching between motor components, and the limited ability of people with Parkinson's to adapt to changing task conditions makes this difficult. Indeed, it appears that slow initiation of movement may be a strategy adopted by those with the disease to facilitate motor tasks (43,44).

Dyskinesia and Dystonia
Dyskinesias, the excessive movement of muscles in the trunk or limbs that cannot be controlled voluntarily, are a common complication that typically develops after 7 or 8 years of levodopa therapy. It is thought to result from a complex interplay between the imperfect supplementation of dopamine, in the form of oral carbidopa/levodopa tablets, and the underlying, progressive loss of dopamine neurons that is the hallmark of Parkinson's.

Most commonly, dyskinesias are seen when the plasma concentration of levodopa is at its highest. This is referred to as “peak-dose” dyskinesia. It is believed to result from abnormal neuronal firing patterns in response to pulsatile stimulation of the dopamine receptors (45). In the early stages, it can be managed by reducing the dose of levodopa or taking levodopa less frequently. However, this typically results in a worsening of the symptoms of Parkinson's, such as rest tremor, bradykinesia, and rigidity. It is important to remember that if dyskinesias are relatively mild and not bothersome or dangerous to the patient, then most people prefer to move too much rather than too little. If dyskinesias become large in amplitude, to the extent that they risk injury to the patient, then they must be treated. Treatment options include decreasing the dose of levodopa, decreasing the frequency with which levodopa is taken, and adding amantadine to the medication regimen.

Dystonia, the involuntary contraction of a single muscle or multiple muscles that cause an abnormal posture, is also commonly seen in those with Parkinson's. The most common site for dystonia in those with Parkinson's is the foot. Typically, at the end of a dose of levodopa, when plasma concentrations are at their lowest, the foot begins to contract with flexion of the toes and inversion of the entire foot. With the ingestion of another dose of levodopa, the dystonia is relieved. Dystonia may also occur as a peak-dose effect, when plasma concentrations of levodopa are at their highest. Unfortunately, peak-dose dystonia is not easily treated. If dystonia presents as a focal problem, such as involving the neck muscles or the facial muscles, then local therapy with botulinum toxin injection is an option.

Orthostatic Hypotension
Up to 20% of the population with Parkinson's may experience orthostatic hypotension, which may be the result of central or peripheral mechanisms. This usually results from sympathetic outflow dysfunction that leads to impaired peripheral vasoconstriction. Intravascular volume depletion due to poor fluid intake or other mechanisms magnifies the problem. Because Parkinson's usually occurs in the elderly, other age-related medical conditions, such as cardiovascular diseases and use of drugs that cause hypotension, should be taken into consideration (46–48).

Lifestyle modification and education are important aspects of treatment. For example, patients should be counseled that warm or hot baths increase peripheral vasodilatation and may contribute to orthostasis. A heavy meal may also result in splanchnic vasodilatation and “steal” blood volume from elsewhere. Excessive straining while defecating and other tasks inducing the Valsalva maneuver should be avoided, and high-fiber diets and stool softeners should be incorporated as well. Nonpharmacologic methods include the use of compression leg stockings and abdominal binders. Management also includes eliminating unnecessary antihypertensive drugs or other medications that may cause hypotension and (47) potentially providing blood pressure support with fludrocortisone or midodrine.

Gastrointestinal Problems
Swallowing Dysfunction
Dysphagia typically results from loss of lingual control and inability to propel the bolus due to delay in the contraction of pharyngeal muscles. Up to 75% of people with Parkinson's experience dysphagia, and esophageal dysmotility may also occur (49–54). Abnormalities in striated muscles under dopaminergic control and smooth muscles under autonomic influence contribute to this complicated impairment. Videofluoroscopic swallowing evaluations help to determine which specific phases of swallowing are impaired. Modified barium swallows under fluoroscopy have shown that the most common abnormalities are motility problems, hypopharyngeal stasis, aspiration, and deficient positioning of the esophagus (55,56).

A speech pathologist may help by teaching oral-motor exercises and providing education on compensatory strategies to prevent penetration and aspiration. A recent randomized controlled trial evaluating the efficacy of the chin-tuck maneuver versus thickened liquids only for Parkinson's and dementia patients demonstrated that a chin-tuck cohort and thickened-liquids-only group equally prevented the development of pneumonia at 3 months follow-up; however, aspiration during videofluorographic assessment was prevented more readily with thickened liquids only (not using chin-tuck) (57). Importantly, when an individual cannot meet his or her caloric and fluid requirements, a gastrostomy feeding tube should be considered.
Nutrition

Parkinson's patients are at risk for weight loss and should be closely evaluated for nutritional losses (52). Protein intake should be monitored and limited, in that amino acids compete with levodopa for absorption. Clinicians may advise patients to take levodopa 1 hour before or after a meal in order to facilitate proper absorption. This is of particular concern to those in the late stages of the disease who have severe dyskinesias. Vitamin supplements should be considered for those who are unable to have well-balanced and adequate nutrition. The precise amount of vitamin supplementation needs to be monitored carefully because vitamin B6 supplementation may result in reduced absorption of levodopa from the intestines (58).

Delayed Gastric Emptying

Parkinson's patients experience early satiety or may experience nausea and vomiting due to delayed gastric emptying. Reduced peristalsis and gastroesophageal reflux present as complaints of "heartburn" or indigestion. This also puts the patient at risk for poor nutritional intake and therefore can exacerbate absorption of levodopa and other medications. Anti-Parkinson drugs themselves may also contribute to gastric emptying. On the other hand, promotility agents, such as metoclopramide (Reglan), may worsen dyskinesias (55).

Constipation

Constipation is a frequent obstacle in people with Parkinson's. Its cause is multifactorial, including altered sympathetic innervation of the gastrointestinal tract, various concomitant medications, overall limited mobility, and impaired hydration. Nonpharmacological treatments include adequate hydration, increased physical activity, and high-fiber diets. If these tactics do not resolve the problem, the use of a daily osmotic agent such as miralax or lactulose can be helpful.

Bladder Problems

One of the most common and earliest bladder abnormalities in patients with Parkinson's is nocturia (45). This, along with urgency and frequency, is thought to result from detrusor hyper-reflexia. In some patients, detrusor hyperreflexia and urinary sphincter problems occur. Urodynamic studies may be needed to diagnose the nature of the problem before appropriate treatment can be instituted. Management includes timed voiding, intermittent catheterization, and pharmacologic agents, such as peripherally acting anticholinergics. Redistribution of fluid intake, such that the bulk of fluid ingestion occurs earlier in the day, may be necessary to avoid nocturia.

Cognition

The onset of the cognitive symptoms in those with Parkinson's is usually slow. The cognitive domains that are most often affected include attention, memory and learning, executive functions, and visual-spatial functions. Verbal function and the ability to reason seem to be spared, although information processing may be slower. In a study of people who were nearly diagnosed with Parkinson's disease, researchers found that 36% showed some form of cognitive impairment (59). Specifically, these patients had difficulty with visual processing and executive functioning.

There is no specific medical treatment for the cognitive dysfunction seen in those with Parkinson's. However, minimizing the use of medications that may cause confusion is an important precaution, as those with baseline cognitive dysfunction are more likely to become transiently worse if treated with medications that are known to cause some confusion as a side effect.

Psychiatric Manifestations

Many people with Parkinson's experience depression. About half of the Parkinson's patients have dysthymia. Estimates of prevalence of depression in those with Parkinson's disease have ranged from 40% to 50% (60). There are unique characteristics to the depression found in those with Parkinson's. Depressed patients with Parkinson's experience higher rates of anxiety, are more often sad without feeling guilt or self-blame, and have lower rates of suicide despite higher rates of suicidal thoughts. The most commonly prescribed antidepressants for people with Parkinson's disease are the selective serotonin reuptake inhibitors (SSRIs). Dopamine replacement, the cornerstone of treatment for Parkinson's disease, can also result in psychiatric complications (61). As levodopa replacement is increased to ease worsening motor symptoms, some people may experience hallucinations, delusions, agitation, mania, or confusion.

As many as 40% of people with Parkinson's experience anxiety and panic attacks. They may result from central neurotransmitter dysfunction or as a reaction to Parkinson's and its complications. Thus, management includes optimization of an anti-Parkinsonian drug regimen, discontinuation of offending drugs, and institution of anxiolytic therapy.

Depression occurring in Parkinson's has been shown to be independent of disease severity and duration (62–65). The prevalence of depression in the Parkinson population has been shown to be roughly 30% as indicated by a 1996 review of 45 studies. Twenty percent of persons demonstrated moderate to severe depression on the Beck Depression Index in a community-based study (66). Many features of depression, such as psychomotor retardation, stooped posture, lack of initiation, decreased affect, poor sleep and appetite, are in fact similar to symptoms of Parkinson's (67,68). Thus, depression in people with Parkinson's may be underrecognized due to some of the similar somatic complaints that exist in both conditions. Those with depression and Parkinson's typically have more anxiety and suicidal ideation and less guilt. It is unclear as to the extent to which "reactive depression," or one's depressive symptoms as related to the chronic disease, exists as compared to neurochemically based or intrinsic depression; however, it is likely that both play a role. Those with Parkinson's indeed exhibit more depressive symptoms than do other chronic disease counterparts, showing that intrinsic processes likely predominate in Parkinson's (69,70). Regarding
treatment for depression in Parkinson’s, there is limited data pointing to specific medications. One report showed efficacy, albeit with a less desirable side effect profile, in a tricyclic antidepressant versus fluoxetine trial (71). Ideally, patients should be followed in a multidisciplinary clinic where Parkinson’s-specific counseling can be offered, as this can be a crucial adjunctive treatment option.

Pain and Parkinson’s
Roughly two thirds of people with Parkinson’s experience pain during the course of the disease (72). Pain in patients with Parkinson’s is either primary (central processes) or secondary to other conditions. Primary pain typically presents as an aching pain in the affected limb. It is more likely to occur in younger patients with predominant dystonia and akinesia and involves the side more affected by rigidity.

Limb rigidity is the most common cause of pain in patients with Parkinson’s and can be incorrectly diagnosed as cervical or lumbar radiculopathy (73). Many people with Parkinson’s have “pseudo-rheumatic disorders” that result in limb and joint deformities, and mechanically derived pain. Fractures should be ruled out whenever a person with Parkinson’s experiences a fall. Patients may also suffer from restless leg syndrome, a condition characterized by an unpleasant sensation in the legs accompanied by an overwhelming need to stretch or walk (72). Other manifestations of primary pain in people with Parkinson’s are headaches, characterized by deep, throbbing, occipital or neck pain. Oral and genital pain has also been reported (72).

Typical analgesics may not be effective. Treatment of the parkinsonian symptoms pharmacologically and nonpharmacologically, with the goal being to improve mobility and flexibility of the affected limbs, can help alleviate primary pain. Secondary pain may present as abdominal discomfort resulting from constipation, or shoulder and limb pain due to complex regional pain syndrome (formally known as reflex sympathetic dystrophy, RSD), an autonomic nervous system dysfunction. Painful limb dystonia is also an important cause of secondary pain and may be helped by botulinum toxin injections.

TREATMENT
Pharmacological Management
The introduction of levodopa several decades ago was an important event in the history of management of Parkinson’s disease. The availability of newer medications contributes to the evolution of pharmacologic therapy for the disease. To date, all available treatments help to control the symptoms of Parkinson’s disease but do not delay or halt its progression.

Experts agree that symptomatic therapy for Parkinson’s disease should be initiated when the symptoms of Parkinson’s disease are severe enough to result in functional impairments that affect the individual’s ability to move around safely, perform activities of daily living (ADL), or maintain employability (19).

Levodopa is commercially available as Sinemet, which is a combination of levodopa and carbidopa. The latter is a decarboxylase inhibitor that prevents the peripheral conversion of levodopa to dopamine. In the absence of adequate carbidopa, levodopa may result in peripheral side effects, such as nausea, vomiting, and hypotension. A sustained-release preparation, which decreases fluctuations in blood level of the drug but is not always reliably absorbed in the intestines, is also available.

Dopamine agonists are another group of drugs that have been used to treat individuals with Parkinson’s disease. Unlike levodopa, dopamine agonists do not need to be converted to an active product. They directly stimulate dopamine receptors. In addition, dopamine agonists appear to have longer duration of effect than does the immediate-release form of levodopa, thus providing more sustained stimulation of dopamine receptors. The older dopamine agonists, such as bromocriptine and pergolide, have been used as an adjunct to levodopa. The most serious side effect associated with their use is cardiac valve abnormalities. Because of this serious potential side effect, and the development of newer, more specific dopamine agonists, these drugs are not used as frequently as they once were. The newer dopamine agonists include pramipexole and ropinirole (74). Pramipexole and ropinirole are selective agonists of the D2 and D3 receptors that have been shown to be effective in early and advanced Parkinson’s disease (75–80).

Amantadine, an antiviral agent, has also been used to reduce dyskinesias. Although its exact mechanism is yet to be elucidated, it is known to have dopaminergic effects by promoting synthesis and preventing reuptake of dopamine, increasing dopamine release, and stimulating dopamine receptors. Up to 10% of patients taking amantadine report dizziness or insomnia, and up to 5% may experience constipation, diarrhea, or nausea. Headache, irritability, vivid dreams, and anxiety have also been described with regular use (81–85).

Anticholinergic agents, such as trihexyphenidyl and benztropine, have also been used to treat the symptoms of Parkinson’s disease. The rationale for their use in patients with Parkinson’s is to restore the imbalance of dopamine and cholinergic transmission in the brain (85). However, with the advent of the newer dopamine agonists, anticholinergic agents are no longer used as commonly as they once were.

Another class of drugs, the catechol-O-methyl transferase (COMT) inhibitors, such as entacapone and tolcapone, are used as adjuncts to carbidopa/levodopa treatment. They prevent the metabolic conversion of levodopa to 3-O-methyldopa peripherally by the enzyme COMT. Thus, there is increased availability of levodopa, resulting in enhanced benefits of levodopa in individuals with Parkinson’s disease (86). Table 26-5 summarizes these medications’ mechanisms of action, dosing, uses, and side effects.

Deep Brain Stimulation
Surgical procedures (Table 26-6) had once been the first-line therapy for patients with Parkinson’s disease, but with the introduction of levodopa in the late 1960s, surgical treatments have become less attractive because of the side effects and complications. However, patients with advanced Parkinson’s disease often develop significant motor complications (freezing,
<table>
<thead>
<tr>
<th>Drug</th>
<th>Mechanisms of Action</th>
<th>Usual Dosing</th>
<th>Side Effects</th>
<th>Other Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levodopa</td>
<td>Levodopa activates D1 and D2 dopamine receptors in the brain. Carbidopa is a peripheral dopa-decarboxylase inhibitor. It increases therapeutic potency and decreases gastrointestinal side effects of levodopa</td>
<td>Sinemet (levodopa/carbidopa) 25/100 titrated upward to achieve desired clinical effects (dosage strengths available: Sinemet, 10/100, 25/100, 25/250; Sinemet-CR, 25/100, 50/200)</td>
<td>Nausea, vomiting, hypotension (usually due to inadequate doses of carbidopa); dyskinesias; motor fluctuations; neuropsychiatric problems (e.g., confusion)</td>
<td>Cornerstone of treatment for Parkinson’s disease</td>
</tr>
<tr>
<td>Dopamine agonists:</td>
<td>Stimulate dopamine receptors</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bromocriptine</td>
<td></td>
<td>2.5–40 mg/d</td>
<td>Similar to levodopa; bromocriptine may cause red, inflamed skin (St. Anthony’s fire), which is reversible on drug discontinuation</td>
<td>Reduced incidence or levodopa-related side effects; selective stimulation of dopamine receptor subtypes; potential neuroprotection; limited antiparkinsonian effect</td>
</tr>
<tr>
<td>Pergolide</td>
<td></td>
<td>0.1–5 mg/d</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cabergoline</td>
<td></td>
<td>0.5–1 mg/d</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pramipexole</td>
<td></td>
<td>1.5–4.5 mg/d</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ropinirole</td>
<td></td>
<td>0.75–24 mg/d</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amantadine</td>
<td>Promotes synthesis and prevents reuptake of dopamine; increases dopamine release; stimulates dopamine receptors</td>
<td>100 mg every other day for 1 wk, with subsequent dose increase to up to 100 mg tid</td>
<td>Insomnia, confusion, hallucination, ankle edema, livido reticularis</td>
<td>Limited clinical efficacy</td>
</tr>
<tr>
<td>Anticholinergics:</td>
<td>Restore imbalance between dopaminergic and cholinergic neurotransmitters</td>
<td>0.5–1 mg bid and gradually increased to 2 mg tid</td>
<td>Significant cognitive (e.g., confusion, hallucination, memory impairment, etc.) and peripheral antimuscarinic side effects (e.g., dry mouth, blurred vision, constipation, etc.)</td>
<td>Limited clinical efficacy</td>
</tr>
<tr>
<td>Trihexyphenidyl</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Benztropine</td>
<td></td>
<td>0.5–2 mg bid</td>
<td>Dyskinesia</td>
<td>Increases levodopa availability; smoother levodopa plasma levels, thus, reducing levodopa-related side effects</td>
</tr>
<tr>
<td>COMT inhibitors:</td>
<td></td>
<td>200 mg/dose (up to 1,600 mg/d)</td>
<td>Diarrhea</td>
<td></td>
</tr>
<tr>
<td>Entacapone</td>
<td></td>
<td>100–200 mg tid</td>
<td>Liver toxicity/tolcapone</td>
<td></td>
</tr>
<tr>
<td>Tolcapone</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Selegiline hydrochloride (Deprenyl or Eldopryl)</td>
<td>MAO-B inhibitor</td>
<td>5 mg at breakfast and 5 mg at noon</td>
<td>Fulminant liver failure Hypertensive reactions may occur if taken with theophylline, ephedrine, carbidopa/levodopa and foods containing tyramine</td>
<td>May potentiate levodopa; some clinicians use it as monotherapy in early IPD. May delay the need for levodopa for about 9 mo. Used with levodopa to potentiate its effects and reduce the dose of levodopa</td>
</tr>
<tr>
<td>Rasagiline (Azilect)</td>
<td>MAO-B inhibitor</td>
<td>5 mg daily</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

COMT, Catechol-D-methyl-transferase; IPD, idiopathic Parkinson’s disease.
dyskinesias that cannot be managed with medication alone) (87). The development of deep brain stimulation (DBS), which allows a neurosurgeon to insert programmable leads into a discrete nucleus of the basal ganglia, has led to a resurgence of surgical treatment for advanced Parkinson's disease (88,89).

DBS was introduced by Benibid et al. in the late 1980s as an alternative surgical treatment for patients with Parkinson’s disease (90). There are many published reports about the clinical effect of DBS, and its use has emerged from being “experimental” to being part of the standard armament in the treatment of Parkinson’s disease. The procedure involves passage of a needle through brain tissue to place electrodes at specific target sites that deliver stimulation to a specific target in the basal ganglia, which is responsible for its clinical effects (91,92).

DBS simulates the surgical ablative procedures with high-frequency stimulation of certain target tissue without making a destructive lesion in the brain. It offers several advantages over ablative procedures: it accomplishes bilateral stimulation with minimal side effects; parameters can be adjusted postoperatively to improve symptoms or minimize side effects; and it is reversible so that the patient can still benefit from future treatment options. The disadvantages include risk for hemorrhage with needle passage, complications related to the device such as mechanical failure (lead placement or battery changes), and infection. The extent of stimulation is adjusted gradually, in the months following the surgery, in an outpatient setting.

The most commonly targeted site for DBS is the STN. Stimulation of the STN allows patients to spend a greater part of their day with minimal rigidity and rest tremor, while on lower doses of anti-Parkinson medication. Currently, research is being done to determine if stimulation at other sites, such as the pedunculopontine nucleus, would be beneficial for treating the postural instability seen in those with Parkinson’s disease (93–96). DBS is indicated in patients who have on/off fluctuations with severe immobility when “off” but continue to respond to levodopa and/or have levodopa-induced dyskinesias.

**Fetal Cell Transplantation**

Fetal cell transplantation offers the potential to replenish the supply of dopaminergic neurons in the substantia nigra. In a recent study where human embryonic dopamine neurons were transplanted into the brains of patients with severe Parkinson’s disease, there was no significant, long-term improvement in symptoms (97). A critical question, which remains unanswered, is how to ensure that the implanted cells form the proper connection with other cells in the brain. There are many obstacles to successful implantation, including identifying optimal transplant variables and target sites, and ethical concerns of using fetal cells (98).

### TABLE 26.6  Ablative and Stimulation Procedures for Parkinson’s Disease

<table>
<thead>
<tr>
<th>Ablative procedures</th>
<th>DBS procedures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thalamotomy</td>
<td>Thalamus (Vim nucleus)</td>
</tr>
<tr>
<td>Pallidotomy</td>
<td>Globus pallidus pars interna (Gpi)</td>
</tr>
<tr>
<td>Subthalamotomy</td>
<td>Subthalamic nucleus (STN)</td>
</tr>
<tr>
<td>DBS procedures</td>
<td>Restorative procedures</td>
</tr>
<tr>
<td>Thalamus (Vim nucleus)</td>
<td>Fetal cell transplantation</td>
</tr>
<tr>
<td>Globus pallidus pars interna (Gpi)</td>
<td>Stem cell transplantation</td>
</tr>
</tbody>
</table>

**THE ROLE OF EXERCISE AND PHYSICAL THERAPY**

The rehabilitation team must always keep in mind that treatment with therapeutic exercise is directed to the *person*, not solely the Parkinson’s. Whether our patients with movement disorders are free of comorbidities, or whether they have congestive heart failure or diabetes, the general health of the patient certainly affects the “functional health” of the patient. Exercise has played a major role in attenuating hypertension and reducing the rates of coronary artery disease, stroke, and type II diabetes mellitus (99,100). Aerobic exercise lowers blood viscosity, which in turn increases oxygen transport to vital organs, especially the brain. Exercise has repeatedly demonstrated positive effects on depression, anxiety, sleep, and general well-being (101). Of key importance especially in the elderly and postmenopausal women populations is the known benefit of weight-bearing and resistive exercise with respect to decreasing the onset of osteopenia (102,103).

In recent years, studies have emphasized the importance of exercise in neuroplasticity. There is evidence that exercise has protective benefits from the *onset* of Parkinson’s. A part of this process is likely that the brain is able to receive more oxygen as a result of exercise. Animal work has revealed that when subjects are exposed to MPTP, in order to artificially induce parkinsonism, exercise alone and exercise along with learning tasks and socialization protect against development of the symptoms (104). This effect may be the result of increased global oxygenation of the brain, in addition to increased amounts of neurotrophic factors, which are thought to limit cell death while fostering new cell growth, in the setting of aerobic exercise (105).

How does exercise promote neuroplasticity in Parkinson’s? Some basic concepts have been proposed: The neurons in the striatum that produce dopamine are highly “sensitive” to bursts of either activity or inactivity; therefore, there is a “use it or lose it” principle involved. Also, motor learning may occur more substantially if in the setting of “rewarding activities,” in that dopamine production can be reward based. In addition, synaptic activity and neuron structural changes are optimized with complicated and intense activities. Lastly, and of key intrigue to all patients and clinicians involved in Parkinson’s, when exercise is introduced at the earliest stages in Parkinson’s, its progression may be slowed (106).

Investigators have also noted that exercise can increase levels of dopamine and that increased frequency, intensity, and duration of exercise influence these systemic levels (107). Physical exercise has also been shown to boost dopamine levels.
in the striatum and is thought to increase due to a calmodulin-dependent pathway, in that experiments have revealed increased serum calcium levels in the brain as a result of exercise (108).

A large prospective study of more than 125,000 men and women investigated the relationship of physical activity and risk of Parkinson’s disease (109). The study reported results from data collected from mailed questionnaires between 1986 and January 2000. While no clear association was found among the women studied, the investigators reported that men who performed strenuous exercise ≥10 months a year during their early adult life had 60% lower risk of having Parkinson’s disease than men who regularly exercised ≤2 months a year (109). Ongoing studies are being performed to determine if moderate-intensity exercise will reduce the progression of Parkinson’s disease (110).

During the early stages of Parkinson’s disease, persons often are able to continue their current aerobic exercises or start programs with guidance from local fitness centers. Physiatrists can assess whether the patient is safe to exercise at such a facility with or without a physical therapist’s evaluation and make the appropriate referrals. Aerobic exercises such as swimming, stationary bike riding, walking (outdoor or treadmill), and dance classes promote overall health benefits and, in particular, endurance (111–114). As with any adult, in particular those more than 40 years of age, prescribing any aerobic program should include cardiovascular screening. Abnormal cardiovascular reflexes are common among persons with Parkinson’s disease. Orthostatic hypotension, cardiac arrhythmias, and exercise-induced hypotension can coexist with Parkinson’s disease (115). Routinely monitoring patients allows proactive referrals to rehabilitation therapies to maximize capabilities and ideally slow the rate of loss of function ability through early interventions of strengthening and flexibility exercises as well as any adaptive equipment or assistive device needs. Similarly, routine monitoring of Parkinson’s disease symptoms also promotes early recognition for speech and swallowing exercise and therapy needs.

Emphasis on maintaining safe, independent mobility is paramount. As rigidity and postural instability increase, modifications to aerobic exercise become necessary to prevent falls. Rigidity in the trunk and neck limits normal upper trunk and lower trunk rotation during gait, which contributes to less efficient gait patterns and decreased gait speed. The reduction in neck range of motion may also contribute to reduction in balance control (23). Forward trunk flexion and stooped head may contribute to festinating gait patterns.

Several studies using various visual markers, stationary and mobile, have been performed showing improvement in step length and velocity (116,117). Studies using visual cues have been valuable in introducing more methods for rehabilitation specialists to enhance treatment effectiveness. Unfortunately, implementation of set, stationary visual cues for improving gait among persons with Parkinson’s has limited carryover and is difficult to reproduce when ambulating in the community.

Auditory and other somatosensory cues may be helpful as well because they bypass the diseased basal ganglia (31,118,119). Continued research in the field of rhythmic auditory stimulus (RAS) has provided new evidence of the benefits of improving gait and limiting freezing episodes when listening to music or metronome-based rhythm. The use of RAS and visual and tactile rhythm cues has been investigated for portable home gait-training use (30,120).

Although traditionally results of the trials in the efficacy of exercise on Parkinson’s have been mixed, a recent review of the literature demonstrated that upon reviewing 14 randomized controlled trials with “moderate” quality of study, exercise was considered efficacious in physical functioning, health-related quality of life, strength, balance, and gait performance. The evidence in this analysis neither supported nor refuted the propensity to reduced falls and depression (104). The benefit of resistance exercise has long been established in other populations, with its role in increasing muscle mass, function, and strength and promoting bone health, and not until recently have researchers assessed the benefit on Parkinson’s (3). Five resistance interventions were evaluated between 2000 and 2007, revealing an overall increase in strength, balance, and muscle mass (57). Table 26-7 lists some of the fundamental

### TABLE 26.7  
Studies on Effects of Therapy and Exercise on Parkinson’s Disease

<table>
<thead>
<tr>
<th>Participants</th>
<th>Method</th>
<th>Intervention</th>
<th>Outcome Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Palmer et al. (1986) (37)</td>
<td>14 (12 males; mean age, 64.9 y)</td>
<td>Randomized parallel group design</td>
<td>Karate exercise (n = 7) vs. stretch exercises (n = 7)</td>
</tr>
<tr>
<td>Comella et al. (1994) (44)</td>
<td>18 participants</td>
<td>Randomized, single blind</td>
<td>Physical and occupational therapy program</td>
</tr>
<tr>
<td>Thaut et al. (1996) (190)</td>
<td>26 (18 males; mean age, 71.5 y)</td>
<td>Randomized parallel group design</td>
<td>Rhythmic auditory stimulation vs. self-paced training</td>
</tr>
<tr>
<td>Nieuboer et al. (2007) (191)</td>
<td>153 participants</td>
<td>Single-blind randomized crossover</td>
<td>Home-based gait cueing training program</td>
</tr>
<tr>
<td>Marchese et al. (2000) (192)</td>
<td>20 (13 males; mean age, 65.9 y)</td>
<td>Randomized parallel group design</td>
<td>Standard physiotherapy (n = 10) vs. novel cued group</td>
</tr>
</tbody>
</table>

ADLs, activities of daily living; EMG, electromyography; UPDRS, Unified Parkinson’s Disease Rating Scale.

**Specific Exercises for Parkinson's**

Specific exercises should include strengthening of back and neck extensor, hip musculature, and abdominal muscles in order to promote a more upright posture. Similarly, emphasis on stretching trunk flexor muscles, hamstrings, and heel cords will promote better posture and muscle length balance. These muscle groups are important for promoting good posture for swallowing and speaking as well as for safe mobility.

Designing and updating exercise programs to meet the changing needs of each person with Parkinson's disease allow the patient and the rehabilitation team the best opportunity to maximize functional ability throughout the progression of the disease.

Physical therapy, emphasizing gait retraining and motor control relearning, may help overcome postural instability. The individual can be trained to focus on foot positioning to provide a more stable base for posture and gait. Other helpful techniques include teaching the individual to turn around in a wide arc, rather that pivoting, in order to avoid the risk for loss of balance and postural stability, thus decreasing the risk for falls. Instructing the individual consciously to take longer steps with the help of verbal and visual (object) cues may help overcome the effects of freezing and festination. With the help of a neurologic music therapist, the physical therapist can provide exercises in a rhythmic fashion, that is, movements coupled with musical or drum-beating rhythm. Other specific exercises should include aerobic conditioning to improve endurance, strengthening of back extensor and abdominal muscles in order to promote a more upright posture, and stretching of trunk flexor muscles.

Assessment for an assistive device for gait will also be performed by the physical therapist. Depending on the predominant gait abnormality, a cane or rolling walker may offer stability while encouraging natural, fluid gait movement. The U-Step walker (InStep Mobility, Inc. http://www.ustep.com/walker.htm) has had increasing popularity in advanced Parkinson's. In the middle of a rolling walker, there is a laser beam that projects forward of the patient for guidance and cuing to help prevent freezing. Instructing the individual consciously to take longer steps with the help of verbal and visual (object) cues may help overcome the effects of freezing and festination.

**OTHER MOVEMENT DISORDERS**

The most common difficulty rehabilitation professionals encounter when faced with movement disorders is in correctly classifying them (Table 26-8) (121) (e.g., Is this dystonia or simply spasticity? Does this patient have a tic or focal myoclonus?). In contrast to the importance of correct classification, identifying the underlying etiology of a given movement disorder (e.g., ischemic, anoxic, traumatic) is often less important in its management. Regardless of the specific cause, the individual movement disorders are often managed identically (with the exception of medication-induced movement disorders).

**The Parkinson's Plus Syndromes**

The Parkinson's plus syndromes include multiple system atrophy (MSA), progressive supranuclear palsy (PSP), and diffuse Lewy Body disease (DLBD). In the first 1 to 2 years of illness, these disorders may mimic Parkinson's. It is only with careful clinical follow-up and monitoring of an individual's response, or lack of a response, to dopamine replacement, that a reasonable clinical diagnosis can be made.

MSA is characterized by urinary incontinence, reduced sweating, and orthostatic hypotension (122). In someone with MSA, these symptoms occur within the first 3 to 5 years of disease onset. In someone with Parkinson's disease, these symptoms may occur, but only 10 to 15 years after disease onset. Clinically, those with MSA do not improve with dopamine replacement.

PSP is characterized by the inability to exercise voluntary movement of the eyes (123). Another feature of PSP is a marked change in personality; affected individuals become apathetic and may appear to be depressed. In addition, PSP causes marked gait difficulty and those affected tend to have frequent falls within 1 or 2 years after developing symptoms. Thus, gait abnormalities become evident more rapidly in those with PSP than in those with Parkinson's disease.

DLBD is characterized by a marked decline in intellectual function, visual hallucinations, and signs of bradykinesia, rigidity, and possibly restless tremor (124). In DLBD, the

---

**TABLE 26.8 Movement Disorders**

<table>
<thead>
<tr>
<th>Movement Disorder</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chorea</td>
<td>Brief, rapid, forceful, dysrhythmic, discrete, purposeless, flinging of limb</td>
</tr>
<tr>
<td>Athetosis</td>
<td>Slow, writhing movements and inability to maintain position of limb or body part</td>
</tr>
<tr>
<td>Ballismus</td>
<td>Large amplitude, flinging movement of limb (usually proximal)</td>
</tr>
<tr>
<td>Dystonia</td>
<td>Sustained muscle contraction that leads to repetitive twisting movements of variable speed and abnormal posture</td>
</tr>
<tr>
<td>Tremor</td>
<td>Rhythmic, oscillatory movements of a body part</td>
</tr>
<tr>
<td>Tic</td>
<td>Intermittent, repetitive, stereotypical, abrupt, jerky, typically affecting the face and head</td>
</tr>
<tr>
<td>Stereotypy</td>
<td>Purposeless, uniformly repetitive, voluntary, movement of whole body areas</td>
</tr>
<tr>
<td>Akathisia</td>
<td>Subjective restlessness, compulsion to move about</td>
</tr>
<tr>
<td>Myoclonus</td>
<td>Sudden, brief, irregular, contraction of a group of muscles</td>
</tr>
</tbody>
</table>
dementia and visual hallucinations progress rapidly within the first 3 years of symptom onset. The bradykinesia and rigidity progress more slowly. Unfortunately, treatment is limited because dopamine replacement, which would alleviate the motor symptoms of bradykinesia and rigidity, worsen the visual hallucinations and can produce an acute worsening of the underlying dementia.

One feature common to all the Parkinson’s plus syndromes is that the symptoms do not improve with dopamine replacement. This is in stark contrast to Parkinson’s, in which the symptoms do improve significantly with dopamine replacement. It is this lack of response to dopamine, as well as the monitoring of the progression of symptoms, that help to distinguish these syndromes from Parkinson’s and from one another.

**Tremors**

Tremors are regular, oscillatory movements produced by alternating but synchronous contractions of antagonistic muscles (125–127). Their rhythmic quality distinguishes them from other involuntary movement disorders (128,129). Clinically, tremors are classified as resting, essential, or action. A resting tremor occurs when the affected limb is at rest and the muscles are relaxed. Resting tremors are typically seen in those with Parkinson’s or one of the Parkinson’s plus syndromes. An essential tremor is a visible and persistent tremor that occurs when the arms or head is in a specific position or when they move. Most cases of essential tremor are hereditary and are ameliorated with alcohol, benzodiazepines, propranolol, or primidone. The action tremor is one that exists with any voluntary muscle contraction. Specifically, an “intention tremor” is a subtype of action tremor, in that the type of muscle contraction is kinetic as opposed to isometric or postural (130).

**Dystonia**

Dystonia refers to a sustained muscle contraction that causes repetitive, twisting movements of variable speed, leading to abnormal posture. Dystonia may be focal, segmental, multifocal, or generalized (131–134). Focal dystonia involves single body parts (e.g., blepharospasm, “writer’s cramp”), whereas the segmental variety affects two or more contiguous regions (e.g., craniofacial). Multifocal dystonia consists of abnormalities in noncontiguous body parts. Generalized dystonia involves segmental cranial dystonia and at least one other body part. Unilateral dystonia is also called hemidystonia.

Regardless of anatomic distribution, dystonic contractions typically begin intermittently and become severe and persistent, leading to sustained abnormal postures. Secondary dystonia is associated with neurologic disorders, such as brain injury, cerebral tumor, and infections (135–137). (Secondary dystonias may also be caused by medications, such as phenothiazines and sertraline [138].)

Impairment of basal ganglia output is thought to play a role in the genesis of dystonia (139). In a series of 22 subjects with hemidystonia, 73% had prior hemiparesis or basal ganglia lesions on imaging studies (140). Lesions in the putamen have been linked to hemidystonia, whereas bilateral putaminal involvement may be responsible for generalized dystonia (131). Torticollis and hand dystonia are thought to result from involvement of the head of the caudate nucleus and thalamus, respectively (140). Disease of the thalamus and subthalamus and derangement of hypothalamic function have also been suspected (141,142).

The current lack of understanding of the exact pathophysiology of dystonia has made it difficult to define specific pharmacological therapy. Treatment has become a “trial-and-error” process, often leading to frustration for the patient and clinician (132,133,143). Oral medications used to treat the various dystonias include dopamine agonists, antagonists, and depletors; anticholinergics; benzodiazepines and baclofen (GABA_A and GABA_B agonists, respectively). Unfortunately, these systemic medications have numerous adverse effects. Local injection of botulinum toxin is a safe and efficacious alternative, particularly for focal dystonias. The efficacy of physical therapy techniques, such as soft-tissue mobilization, cervical muscle strengthening and stretching, and orthotic intervention have not been well studied to date. Similarly, limited success has been achieved through behavioral modifications, including hypnosis, biofeedback, and relaxation techniques.

**Cervical Dystonia**

Cervical dystonia, the most common focal dystonia, involves the sternocleidomastoid, trapezius, and posterior cervical muscles. It gives rise to patterned, repetitive, and spasmodic movement that causes the head to twist (rotational torticollis), extend (retrocollis), flex (anterocollis), or tilt toward the shoulder (laterocollis). One or more of these head movements may occur simultaneously. Walking or standing worsens the condition, but the patient may be able to return the head to midline by placing the hand on the jaw or chin. The neck movement may be associated with blepharospasm, lip or chewing movements, and tremor.

Because they play a role in maintaining normal head posture, the basal ganglia and the vestibulocochlear reflex pathway have been implicated in the development of cervical dystonia (144–147). Disturbances of neurotransmitter systems have also been described in dystonias (148). Abnormalities in blink reflex recovery have suggested involvement of the brainstem (149). Earlier cervical and upper-limb trauma have also been implicated in development of cervical dystonia (150–152).

With respect to treatment, baclofen, benzodiazepines (e.g., clonazepam), anticholinergics (e.g., trihexyphenidyl), carbamazepine, and dopamine agonists or antagonists have been trialed, but side effects may limit their use. Local injection of botulinum toxin into the offending muscles (i.e., sternocleidomastoid, trapezius, splenius capitis) has been successful and not associated with significant complications. In few cases, dysphagia may develop due to local spread of the toxin to neighboring pharyngeal and laryngeal muscles.
Chorea, Athetosis, and Ballismus

Chorea, derived from the Greek word for dance, is an involuntary movement that is brief, rapid, forceful, and dysrhythmic (153). In general, choreic movements are discrete and purposeless and involve distal body regions such as the hands and feet. Athetosis refers to writhing, snakelike involuntary movements that involve larger and more proximal muscle groups, such as the entire arm or leg. It is believed that the contralateral STN, caudate nucleus, and putamen may be responsible for chorea (154–157). Bilateral thalamic involvement has been described as well. Choreic movements have been evident in Sydenham’s chorea, hyperthyroidism, cerebral arteritis, polycythemia vera, systemic lupus erythematosus, Huntington’s disease, and phenothiazine intake (137,158).

The movements in athetosis are typically slower than choreiform movements. They are characterized by writhing movements and the inability to maintain the position of virtually any body part (e.g., fingers, wrists, toes). Although the limbs are most commonly affected, the axial musculature may be involved as well. Athetoid movements may be seen in Wilson’s disease, cerebral palsy, and basal ganglia disease. In some cases, athetosis is drug induced. When it appears with chorea, it is referred to as choreoathetosis.

In contrast to chorea, movements related to ballismus are of large amplitude and involve the proximal limbs. The movements often come on suddenly, with no clear pattern. Because ballismus is frequently unilateral, it is also commonly referred to as hemiballismus. Involvement of the contralateral STN has been suggested, but other subcortical structures may also be involved (155–157,159,160). It is thought that a lesion in the contralateral STN disrupts the inhibitory pathways to the globus pallidus, leading to dopamine hyperactivity in the striatum (161). Bilateral involvement (biballism) is seen in bilateral basal ganglia disease (162). Ballismus has also been associated with metabolic abnormalities, such as hyperglycemia, neoplasms, systemic lupus erythematosus, and encephalitis (156,163).

Antiepileptic drugs, such as phenobarbital and valproic acid, may be of benefit in chorea, athetosis, and hemiballismus. In hemiballismus, dopamine antagonists, such as haloperidol and phenothiazines, and dopamine-depleting agents, such as reserpine and tetrabenazine, may be helpful (160). GABA agonists, such as clonazepam, may also ameliorate chorea in that GABA appears to mediate the inhibitory action of the STN (160). Stereotactic thalamotomy may be considered in severe conditions refractory to medications (164).

Tics, Stereotypy, and Akathisia

Tics are intermittent, stereotypical, repetitive, jerky movements. Although the individual is aware of such movements, he or she finds it difficult to resist performing the action (165). Tics can be either motor or vocal. Many tics are associated with purposeful tasks such as eye blinking and throat clearing. They generally do not interfere with willed, voluntary movements and may be volitionally suppressed to some extent. Tics may also be classified as simple, such as grimacing, or complex, as in the many tics characteristic of Gilles de la Tourette syndrome (166). In general, tics disappear during sleep and worsen during stressful situations.

Stereotypy is purposeless, uniformly repetitive, voluntary movement of whole body areas. Examples include head nodding, head banging, body rocking, and arm jerking, seen in individuals with mental retardation and amphetamine addiction (141).

Akathisia is formally defined as inner restlessness and compulsion to move about. Although a subjective experience by definition, it may manifest overtly as the inability to stand or sit still, or as an urge to pace constantly. In some, the only finding is toe tapping or leg shaking. Akathisia is often seen in the early stages after traumatic brain injury, when it is often difficult to distinguish from other behavioral sequelae, such as agitation. Akathisia is thought to result from dopamine blockade in the frontal area (167). Thus, antidopaminergic medications, such as neuroleptics, may also induce this disorder. Dopamine agonists, such as ropinirole, are sometimes used. Clomipramine, clonidine, propranolol, piracetam, and clonazepam (168,169) may also be beneficial. Neuroleptic-induced akathisia may be treated with amantadine, but it is best managed by reducing the dose or discontinuing the offending drug (170).

Psychogenic Movement Disorders

Psychogenic movement disorders result from various psychiatric conditions, and up to 9% of conditions presenting with neurological symptoms are believed to have no “organic” basis (171,172). A recent evaluation of self-reporting of disability in patients with psychogenic movement disorders versus those with Parkinson’s revealed that the severity of perceived disability was equivalent in both populations (173). In a retrospective evaluation of 4,470 patients, 405 (about 9%) were found to have psychogenic disorders, with motor disorders as one of the more common symptoms (171). In a series of 842 patients with movement disorders, 3.3% were diagnosed as having clinically documented psychogenic movement disorders (172).

Common psychogenic movement disorders are tremor, dystonia, myoclonus, tics, chorea, hemiballismus, and parkinsonism (174). One feature suggestive of psychogenic movement disorders is an inconsistent or fluctuating clinical presentation of symptoms. Psychogenic movement disorders also usually have an acute onset and static course characterized by spontaneous remissions. They are worsened by attention and dampened by distraction. Typically, they are no more responsive to medications than placebo.

When psychogenic movement disorders are entertained, attempts should be made to make a psychiatric diagnosis even if the diagnoses of psychogenic movement disorders are based on neurological examination. Common psychiatric diagnoses include conversion, malingering, somatoform disorder, factitious disorder, depression, and anxiety. Associated psychiatric diagnoses (usually depression), precipitating events, and secondary gain are obvious in up to 60% of cases of psychogenic movement disorders. Behavioral management, encouragement
and support, in addition to rehabilitation, have shown some efficacy in managing psychogenic gait disorders; however, when symptoms are present ≤12 months, there is high likelihood for long-term disability (175).

**Drug-Induced Movement Disorders**

Medications are common causes of movement disorders and should be considered in the evaluation of movement disorders (176). Neuroleptics are among the most common of the medications that can directly trigger movement disorders. Extrapyramidal syndromes, including akathisia, parkinsonism, dystonia, and tardive dyskinesia, represent untoward motor side effects of antipsychotic drug therapy. The mechanism is thought to involve postsynaptic blockade of dopamine receptors (169,177,178).

Acute extrapyramidal syndromes present within a few days following administration of neuroleptics and may persist days after withdrawal of the offending agent. Although considered different manifestations of the same underlying etiology, drug-induced akathisia, dystonia, and parkinsonism have unique motor and mental symptoms that help distinguish one from another (178). A less common extrapyramidal reaction to neuroleptics is acute laryngeal dystonia (176).

Tardive dyskinesia, characterized by orofacial dyskinesia, dystonia, and choreoathetosis, is thought to result from hypersensitivity of dopamine receptors in the basal ganglia (due to prolonged postsynaptic receptor blockade by neuroleptics). Advanced age, female gender, history of alcohol or substance abuse, diabetes, and smoking are considered risk factors for tardive dyskinesia in older individuals (177). In contrast to tardive dyskinesia, the parkinsonian side effect is thought to arise from blockade of dopamine receptors in the striatum.

Other drugs have been implicated in the development of movement disorders. For instance, lithium, methyldopa, and metoclopramide may bring about parkinsonism (179). Choreoathetosis may result from tricyclic antidepressants, oral contraceptives, amphetamines, pemoline, and lithium (180). Diphenhydramine and flecainide have been reported to cause dystonia, asterixis has been associated with carbamazepine, and phenytoin has been implicated in the development of choreoathetosis (181,182,183,184). Table 26–9 lists drugs commonly encountered in physiatric and neurological practice that may induce movement disorders (185–189).

**OTHER TREATMENT CONSIDERATIONS**

**Botulinum Toxin Injection**

Local intramuscular injection of botulinum toxin is arguably the most important advance in the nonsurgical management of dystonia and other movement disorders. The clinically available neurotoxins, derived from *Clostridium botulinum* serotypes A and B, block neuromuscular transmission by inhibiting the release of acetylcholine (but not its synthesis or storage). Injected intramuscularly, the toxin results in partial denervation. Clinical effects are seen 24 to 72 hours after injection, and peak effects occur 4 to 6 weeks later. The average duration of effect is 3 to 4 months. Injections are guided by surface anatomy, electromyography, or electrical stimulation.

Early indications for treatment were blepharospasm and strabismus. Over the years, further disorders have been found to be successfully managed by the technique. Among these are the various dystonias: cervical dystonia (spasmodic torticollis), hemifacial spasm, orolingual dystonia, cranial dystonia, limb dystonia, and "occupational cramps" (190,192–194).

Physiatrists can employ their knowledge of anatomy, kinesiology, and electromyography when treating with botulinum toxin. Because it affects neuromuscular transmission, the toxin is contraindicated in disorders of neuromuscular junction, such as myasthenia gravis and myasthenic syndrome. However, successful use in the treatment of spasmodic torticollis in a patient with myasthenia gravis has been reported (195). Possible adverse effects include excessive weakness in the injected and adjacent muscles. Flu-like symptoms and allergic reactions have been described, but this has rarely occurred in our experience. The clinical effects wear off after 3 to 4 months, and reinjection may be necessary. Readministration

<table>
<thead>
<tr>
<th>TABLE 26.9 Some Medications That May Induce Movement Disorders</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dopamine antagonists</strong></td>
</tr>
<tr>
<td>Haloperidol</td>
</tr>
<tr>
<td>Metoclopramide</td>
</tr>
<tr>
<td><strong>Dopamine agonists</strong></td>
</tr>
<tr>
<td>Levodopa</td>
</tr>
<tr>
<td><strong>Antihypertensives</strong></td>
</tr>
<tr>
<td>Methylphenidate</td>
</tr>
<tr>
<td><strong>Monoamine oxidase inhibitors</strong></td>
</tr>
<tr>
<td>Phenylethylamine</td>
</tr>
<tr>
<td><strong>Antiepileptics</strong></td>
</tr>
<tr>
<td>Phenytoin</td>
</tr>
<tr>
<td>Carbamazepine</td>
</tr>
<tr>
<td>Valproic acid</td>
</tr>
<tr>
<td>Gabapentin</td>
</tr>
<tr>
<td>Felbamate</td>
</tr>
<tr>
<td><strong>Adrenergic agents</strong></td>
</tr>
<tr>
<td>Amphetamines</td>
</tr>
<tr>
<td>Methyldopa</td>
</tr>
<tr>
<td>Caffeine</td>
</tr>
<tr>
<td><strong>β-Adrenergic agonists</strong></td>
</tr>
<tr>
<td>Others</td>
</tr>
<tr>
<td>Antihistaminics</td>
</tr>
<tr>
<td>Tricyclic antidepressants</td>
</tr>
<tr>
<td>Buspirone</td>
</tr>
<tr>
<td>Lithium</td>
</tr>
<tr>
<td>Cimetidine</td>
</tr>
<tr>
<td>Oral contraceptives</td>
</tr>
<tr>
<td>Cocaine</td>
</tr>
<tr>
<td>Selective serotonin reuptake inhibitors</td>
</tr>
</tbody>
</table>

of the toxin within 3 months of the previous injection is discouraged to avoid potential development of antibodies.

**Intrathecal Baclofen**

Intrathecal baclofen is helpful in alleviating focal limb and axial dystonia (196,197). Recently, we successfully treated a man with posttraumatic hemiballismus, and dystonia was successfully treated with intrathecal baclofen. After intrathecal infusion of baclofen, there was a dramatic decrease in the frequency and amplitude of ballismus. Although early reports suggest that intrathecal baclofen may be helpful, its efficacy has not been supported by controlled trials. Furthermore, outcomes from intrathecal baclofen therapy have not been compared directly with those from oral medications, botulinum toxin, and other neurological techniques.

This procedure is an alternative when oral medications are ineffective and their side effects are intolerable, or when the severity of the condition requires more than the recommended dose of botulinum toxin. Although these injections are a neurological procedure, physiatrists may participate in screening patients with movement disorders who might benefit from intrathecal baclofen. Trained physiatrists also play a critical role in postimplantation rehabilitation care by monitoring response and complications and prescribing appropriate therapy interventions.

Patients who respond to a trial dose of intrathecal baclofen are considered for surgical implantation of the infusion system. The pump, with the drug reservoir, is placed in the abdominal wall and connected to a catheter that has been introduced to the intrathecal space, usually at the lower thoracic level. An external programmer adjusts the dose, rate, and mode of drug delivery by radiotelemetry. The drug reservoir can be refilled through transcutaneous insertion of a Huber-type needle into the reservoir port.

Simple dose titration and safe concurrent use with other therapies, such as oral drugs and botulinum toxin, are among the advantages of intrathecal baclofen. Common adverse effects include drowsiness, weakness, and dizziness, which subside with dose reduction. Pump-related problems include catheter kink, fracture, dislodgment, and disconnection, which are corrected surgically.

**REFERENCES**


CHAPTER 26 | REHABILITATION OF PERSONS WITH PARKINSON'S DISEASE AND OTHER MOVEMENT DISORDERS


Rehabilitation of Spinal Cord Injury

The earliest reference to spinal cord injury (SCI) is found in the Edwin Smith Surgical Papyrus, written between 2,500 and 3,000 BC, as “an ailment not to be treated” (1). Much has changed in the last 50 years in spinal cord care as it relates to increasing survival, life expectancy, community reintegration, and quality of life (QOL). Major advances include the specialized spinal cord centers of care; model SCI centers funded by the National Institute on Disability and Rehabilitation Research (NIDRR), Department of Education, United States; establishment and growth of organizations and journals dedicated to SCI; and the development of the subspecialty of SCI Medicine in 1998. The subspecialty of SCI medicine addresses the prevention, diagnosis, treatment, and management of traumatic and nontraumatic (NT) etiologies of spinal cord dysfunction (2). Physicians with training in a variety of related specialties are eligible for fellowship training and can become certified after passing a written exam. The advances of the last decade alone have been dramatic in terms of the understanding of the pathology of the initial and secondary aspects of the injury, and the barriers that must be overcome to enhance recovery. Newer techniques to improve function and intervene at the cellular level for possible cure are being developed that will further allow individuals who sustain a SCI to be more independent in the future.

Epidemiology of Traumatic SCI

Incidence and Prevalence

The National SCI Statistical Center (NSCISC) database has been in existence since 1973 and captures approximately 13% of all new traumatic SCI that occur in the United States each year, and has been used to develop an epidemiological profile (3–6). When compared to population-based studies, persons in this database are representative of all SCIs except that more severe injuries, nonwhites, and injuries due to acts of violence are slightly overrepresented. The incidence of traumatic SCI in the United States has remained relatively constant, at approximately 40 new cases per million population, or approximately 12,000 cases per year. The incidence of SCI in the rest of the world is consistently lower than in the United States (7). The prevalence in the United States is estimated at 250,000 persons.

Age, Gender, and Race

SCI primarily affects young adults, with most injuries occurring between the ages of 16 and 30, but over time the average age at injury has steadily increased. Approximately 5% of traumatic SCI occur below the age of 15. From 1973 to 1979, the average age at injury was 28.7 years, but since 2000 is 39.5 years, with 11.5% of persons injured older than 60. Reasons for the observed trend toward older age at injury may include changes in the referral patterns to model SCI systems and the location of the systems that contribute data to the NSCISC, survival rates of older persons at the scene of the accident, or age-specific incidence rates. In adults, men suffer traumatic SCI more commonly than women, at a 4:1 ratio. In the pediatric population, the difference is less dramatic (see Chapter 74). State registries and NSCISC data reveal higher incidence rates of SCI for African Americans than Whites, although this trend may have changed in 2000 (4). It is unknown whether changing locations of or referral patterns to model SCI systems, or race-specific incidence rates are responsible for this trend.

Etiology and Time of Injury

Motor vehicle crashes (MVCs) including automobile, motorcycle, and bicycle, rank first (since 2005, MVCs account for 42% of cases), followed by falls (27.1%), acts of violence (15.3%) (primarily, gunshot wounds [GSW]), and recreational sporting activities (7.4%) (6). MVCs account for a lower percentage of cases among men than women while men have a higher percentage of SCI that are due to GSW, diving mishaps, and motorcycle crashes (4). The likelihood of SCI from a MVC is higher in non-sedan cars (e.g., sport utility vehicles) involved in rollover crashes (8). MVC is the leading cause of traumatic SCI until age 45, when falls become the leading cause. Falls are most often from low heights, resulting in a cervical lesion (9). There is an increased frequency of cervical spinal stenosis, placing the elderly population at a greater risk of SCI with relatively minor trauma.

Acts of violence caused 13.3% of SCI prior to 1980, and peaked between 1990 and 1999 at 24.8%, before declining to 15.3% since 2005 (6). Violence causing SCI is more common in minority groups. Diving injuries account for the majority of SCI due to recreational sports, followed by snow skiing, surfing, wrestling, and football. Recreational sports and acts of violence decrease with advancing age, as a cause of injury.

In children, the incidence of SCI in the United States is 1.99 cases per 100,000. MVCs are the leading cause, and in children involved in a MVC, two thirds are reported as not wearing a seat belt. Alcohol and drugs are reportedly involved...
in 30% of all pediatric cases. Boys are twice as likely to have a SCI as girls, with an overall mean age of 14.6 years. Incidence of pediatric SCI varies by region, with the South and Midwest regions having almost twice as many as in the Northeast (10).

Traumatic SCI occurs with greater frequency on weekends, with the greatest incidence on Saturday. Seasonal variation exists, with peak incidence occurring in July followed closely by August and June. The seasonal pattern is more pronounced in the northern part of the United States.

**Associated Injuries**
Spinal cord injuries are often accompanied by other significant injuries. The most common include broken bones (i.e., ribs, long bones), loss of consciousness, and traumatic pneumothorax. The nature and frequency of these injuries is, typically, associated with the etiology of the SCI. For example, pneumothorax occurs more frequently with GSW as compared with other causes of SCI. Noncontiguous spinal injuries (a lesion separated by at least one normal intervening vertebra from a spine fracture or subluxation/dislocation) occur in 12% to 28% of cases as documented by magnetic resonance imaging (MRI) (11,12).

**Neurological Level and Extent of Lesion**
Traumatic SCI, most commonly, causes cervical lesions (approximately 50%) followed by thoracic and then lumbo-sacral lesions. The C5 segment is the most common lesion level, followed by C4, C6, T12, C7, and L1, as per documentation at the time of discharge from inpatient rehabilitation programs (7). Since 2000, the most frequent neurologic category at discharge from rehabilitation of persons reported to the NSCISC database is incomplete tetraplegia (34%), followed by complete paraplegia (23%), incomplete paraplegia (18.5%), and complete tetraplegia (18.3%) (6). Pediatric SCI more often results in paraplegia and neurologically complete injuries than adult SCI. Less than 1% of persons experience complete neurologic recovery at discharge. The percentage of persons with incomplete tetraplegia has recently increased, while that of those with complete paraplegia and tetraplegia has decreased slightly.

The etiology of injury is strongly associated with the level and severity of the injury. Most recreational sports-related injuries, falls, and approximately 50% of MVC result in tetraplegia, whereas acts of violence usually result in paraplegia (4,7). Neurologically complete injuries are more likely to occur as a result of acts of violence and among younger age groups. Thoracic injuries are most likely to be neurologically complete while most lower level lesions are incomplete injuries. Cervical injuries are most commonly classified as either ASIA Impairment Scale (AIS) A or D.

**Discharge Placement**
Approximately 88% of persons discharged from a model system are discharged to a private residence within the community (in most cases their homes before injury), with 5.6% being discharged to a nursing home (NH) (6). The remaining patients are discharged to hospitals, group living situations, or other destinations. Predictors of NH placement include being ventilator-dependent, older in age, tetraplegia with nonuseful motor recovery, unmarried, unemployed, and having either Medicaid or health maintenance organization insurance (17). There has been a significant trend toward an increasing percentage of persons being discharged to NH since 1995 with the advent of shorter inpatient rehabilitation lengths of stay (LOS) (17,18).

Average days hospitalized in the acute care unit for those who enter a model system immediately following injury has declined from 25 days in 1974 to 15 days in 2005. Similar downward trends are noted for days in the rehabilitation unit (from 115 to 39 days). Overall, mean days hospitalized (during acute care and rehabilitation) are greater for persons with neurologically complete injuries (6).

**Life Expectancy**
Mortality rates are significantly higher during the first year after injury than during subsequent years, particularly for severely injured persons. Life expectancy of persons with SCI has improved significantly from 50 years ago, but remains below normal. Over the last three decades, there has been a 40% decline in mortality during the first 2 years after injury. While overall life expectancy has improved, the decline in mortality over time in the post 2-year period is small (19). Predictors of mortality after injury include male gender, advanced age,
TABLE 27.1 Life Expectancy (Years) for Post-Injury by Severity of Injury and Age at Injury (for Persons Surviving At Least 1 Year Post-Injury)

<table>
<thead>
<tr>
<th>Age at Injury</th>
<th>No SCI</th>
<th>Motor Functional at Any Level</th>
<th>Paraplegia</th>
<th>Low Tetraplegia (C5-8)</th>
<th>High Tetraplegia (C1-4)</th>
<th>Ventilator-Dependent at Any Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>58.4</td>
<td>53.0</td>
<td>45.8</td>
<td>41.0</td>
<td>37.4</td>
<td>23.8</td>
</tr>
<tr>
<td>40</td>
<td>39.5</td>
<td>34.5</td>
<td>28.2</td>
<td>24.2</td>
<td>21.2</td>
<td>11.4</td>
</tr>
<tr>
<td>60</td>
<td>22.2</td>
<td>18.0</td>
<td>13.2</td>
<td>10.4</td>
<td>8.6</td>
<td>3.2</td>
</tr>
</tbody>
</table>


ventilator dependence, injured by an act of violence, high injury level (particularly C4 or above), a neurologically complete injury, poor self-rated adjustment to disability, poor community integration, poor economic status indicators, and having either Medicare or Medicaid third-party sponsorship of care (19–21).

Life expectancy estimates (Table 27-1) are typically based on neurological level of injury (NLI), degree of injury completeness, age at injury, and ventilator dependency (6). For persons with complete injuries, mortality rates are higher for those with high tetraplegia (C1-3) than for those with mid or low tetraplegia, and the latter have higher mortality than those with paraplegia. The distinction between injury grades is more important for those with the highest levels of injury, but not for those with lower injuries. For persons with paraplegia, there is no significant difference between AIS grades of A, B, or C injuries. Those with higher paraplegia (T1-6) have a higher mortality than lower injuries (18).

Causes of Death
Diseases of the respiratory system are the leading cause of death following SCI, with pneumonia being the most common. Heart disease ranks second, followed by sepsisemia (usually associated with pressure ulcers [PUx], urinary tract or respiratory infections), and cancer (6,7,20–22). The most common location of cancer is the lung, followed by bladder, prostate, and colon/rectum.

Pneumonia is by far the leading cause of death for persons with tetraplegia while heart disease, sepsisemia, and suicide are more common among persons with paraplegia. The suicide rate is highest in younger patients and in persons with paraplegia (20,22). Among persons with incomplete motor-functional (American Spinal Injury Association [ASIA] D) injuries at any neurologic level, heart disease ranks as the leading cause of death (24%), followed by pneumonia (11%). Heart disease is the primary cause of death in persons injured for more than 30 years and in patients over 60. While genitourinary (GU) disease (i.e., renal failure) was the leading cause of death 30 years ago, this has declined dramatically, most likely due to advances in urological management.

Lifetime Costs
Data are available from the model systems for the direct costs of the SCI, with the indirect costs (i.e., lost wages, fringe benefits) not included in these estimates (6). Estimates from a 1998 publication of the total annual costs of SCI were $9.73 billion (23). The average yearly health care and living expenses and the estimated lifetime costs that are directly attributable to SCI vary by the year postinjury (first year versus subsequent years), by the level, and severity of injury (6).

ACUTE MEDICAL AND SURGICAL MANAGEMENT
Clinical practice guidelines (CPG) have recently been published on the early management of adults with SCI (24). The treatment of a traumatic SCI begins at the scene. An injury to the spinal column should be suspected whenever trauma occurs. As such all trauma victims should have their spine immobilized, preferably with a rigid cervical collar with supportive blocks on a backboard, with straps to secure the entire spine in patients with a potential spinal injury, and should be transferred onto a firm padded surface while maintaining spinal alignment to prevent skin breakdown. Movement should be via logrolling until spinal injury has been ruled out. Traditional cardiopulmonary resuscitation (CPR) methods should be utilized, minimizing trauma to a potentially unstable cervical spine that is utilizing the jaw-thrust maneuver to access the airway. After injury, prompt resuscitation, stabilization of the spine, and avoidance of additional neurologic injury and medical complications are of greatest importance. During the first seconds after SCI, there is release of catecholamines with an initial hypertensive phase. This is rapidly followed by a state of spinal shock, defined as flaccid paralysis and extinction of muscle stretch reflexes below the injury level, although this may not occur in all patients. Ditunno et al. proposed four phases of spinal shock from initial loss of reflex activity to hypertreflexia (25). Neurogenic shock, as part of the spinal shock syndrome, is a direct result of a reduction in sympathetic activity below the level of injury, consists of hypotension, bradycardia, and hypothermia, and is common in the acute postinjury period. Parasympathetic (PS) activity predominates, especially in persons with injuries at or above the T6 level. Treatment of hypotension involves fluid resuscitation (usually 1 to 2 L) to produce adequate urine output of greater than 30 cc/h. In neurogenic shock, further fluid administration must proceed cautiously, as the patient
is at risk for neurogenic pulmonary edema, and vasopressors are utilized. Maintenance of mean arterial pressure at approximately 85 mm Hg during the first week postinjury has been associated with improved neurological outcomes (24, 26).

Bradycardia is common in the acute period in cervical spinal level injury and may be treated, if below 40 per minute or if symptomatic, with intravenous (IV) atropine (0.1 to 1 mg), or prevented with atropine given prior to any maneuver that may cause further vagal stimulation (i.e., nasotracheal suctioning). While significant bradycardia typically resolves within 6 weeks, episodes of persistent bradycardia beyond this time may occur in some severe injuries. Some patients may require implantation of a cardiac pacemaker to facilitate safe mobilization (27).

Respiratory assessment is critical for acute SCI patients, and should include arterial blood gases and measurement of forced vital capacity (VC) as an assessment of respiratory muscle strength (28). A VC of less than 1 L indicates ventilatory compromise and the patient usually requires assisted ventilation. Serial assessments should be obtained for those with borderline values. A nasogastric tube should be inserted during the initial assessment period to prevent emesis and potential aspiration. A Foley catheter should be inserted with an acutely for urinary drainage and facilitates accurate assessment of urine output and should be left in place until the patient is hemodynamically stable and strict attention to fluid status is no longer required (24).

Upon presentation to the emergency department, a baseline neurological examination should be performed, maintaining spinal precautions. Imaging studies including x-rays, computed tomography (CT) scan, or MRI should be employed to assess spinal fracture, instability, and/or spinal cord pathology. A standard trauma series includes cross-table laterals and AP views of the cervical and thoracolumbar spine. Because of the incidence of noncontiguous fractures (10% to 40%), once one fracture is identified careful inspection of the rest of the spine is imperative. CT scans often provide improved visualization of the C1 and C7 vertebrae while MRI provides optimal visualization of the neuronal structures. The spine should remain immobilized until an injury has been definitively excluded or the spine is stabilized either surgically or by application of an appropriate orthotic device. In patients with a stiff spine and midline tenderness, the clinician should suspect a fracture (even if plain x-ray is negative), especially in the presence of spondylodiscitis, ankylosing spondylitis, or diffuse interstitial skeletal hyperostosis (DISH) (24). In cases of cervical dislocations (if patient is cooperative), weights can be applied to Gardner-Wells tongs to achieve cervical distraction and spinal realignment. Application of a halo or surgery will follow. Forty-seven percent of patients with spine trauma and 64% of patients with SCI have concomitant injuries, including head, chest, rib, and long bone fractures (29). Therefore, a thorough assessment of the total patient is imperative.

Stab wounds and GSWs generally do not produce spinal instability and therefore usually do not require surgical stabilization or orthotic immobilization. Objects that are embedded around the spinal canal (i.e., knife) should be left in place with removal performed in the operating room under direct visualization of the spinal canal. Bullets that pass through the abdominal viscera are treated with broad spectrum antibiotics and tetanus prophylaxis (30, 31). Bullets do not have to be removed; however, they can be if accessible while performing another surgical procedure.

In many trauma centers in the United States, intravenous (IV) methylprednisolone (MP) is given to adults after an acute SCI. Mechanisms of action for MP include improving blood flow to the spinal cord, preventing lipid peroxidation, free radical scavenger, and having anti-inflammatory function. The National Acute SCI Study (NASCIS) 2 reported that IV MP given within 8 hours of injury (30 mg/kg bolus and 5.4 mg/kg/h for 23 hours) improves neurologic recovery at 6 weeks, 6 months, and 1 year, although functional recovery was not clearly studied (32). NASCIS 3 reported that if initiated within 3 hours of SCI, MP should be continued for 24 hours, whereas if initiated at 3 to 8 hours after SCI it should be continued for 48 hours (33). The administration of MP is not extended beyond 8 hours from SCI or in those with penetrating injuries, as they have shown no benefit and their use is associated with a higher incidence of infections (34, 35). The benefits and safety of utilizing the NASCIS protocol has been questioned, due to the fact that the findings have not been consistently replicated, concerns regarding methodology and analysis, as well as possibly increased morbidity and mortality in persons administered steroids (36–38). The neurosurgical guidelines consider the use of high dose MP to be a treatment option rather than a standard (39), and the CPG states that there is no clinical evidence to definitively recommend the use of any neuroprotective pharmacologic agent, including steroids, in the treatment of acute SCI to improve functional recovery (24).

Additional recommendations of the CPG include transferring the SCI patient to a specialized center as soon as possible to decrease complications and hospital LOS. Patients with acute SCI, especially high level tetraplegia, should be assessed for evidence of concomitant traumatic brain injury (TBI) (i.e., assessing for loss of consciousness or post-traumatic amnesia [PTA]). Early stabilizations should be considered for extra-skeletal fractures. In cases of high-energy injuries, aortic injury should be evaluated. For anesthesia, avoid the use of succinylcholine after the first 48 hours postinjury (potentially fatal hyperkalemic response). While priapism is frequently seen, it is usually self-limited and does not require treatment. Lastly, it is important to maintain normoglycemia in critically ill, mechanically ventilated patients (24).

Not all SCI is associated with a spinal fracture or dislocation (24) and may result from forced extreme range of spinal movement without mechanical abnormality. A high index of suspicion for SCIWORA (SCI without radiological abnormality) is important when evaluating adolescents with sports-related neck trauma or victims of child abuse (especially in children who may be suffering from physical abuse) (24). See Chapter 59 for pediatric-related SCIWORA.
Spinal Stability and Principles of Spinal Stabilization

White and Panjabi proposed the most widely accepted theory on spinal instability defining it as “the loss of the ability of the spine, under physiologic loads, to maintain its pattern of displacement so that there is no initial or additional neurological deficit, no major deformity, and no incapacitating pain” (40). This definition is applicable at all levels of the axial spine. Radiographic criteria have been established for the diagnosis of clinical instability of the spine. Denis described the widely accepted “three-column theory” for thoracolumbar fractures, where the spine is divided into three columns (Fig. 27-1). The anterior column comprises the anterior vertebral body, the anterior longitudinal ligament and the anterior half of the annulus fibrosus. The middle column consists of the posterior vertebral body, the posterior longitudinal ligament and the posterior column comprises the posterior half of the annulus fibrosus. The posterior column includes all the posterior elements (including the pedicles). In this three-column theory of Denis, spinal instability is present if any two of the three columns are violated (41).

Injuries that are primarily ligamentous, such as facet dislocations, are unstable and require internal stabilization procedures (42). The primary goal of surgical intervention in acute SCI is to decompress the neural elements, and either an anterior or a posterior approach may accomplish this. The approach chosen depends on the expertise of the operating surgeon and the specific pathophysiology of the injury. Since the most common etiology of SCI occurs from retropulsion of bone and/or disc material from a ventral location into the spinal canal, an anterior approach may be preferable. Anterior surgery, however, is associated with increased complications including recurrent laryngeal nerve lesions leading to speech and swallowing disorders. After adequate neural decompression is accomplished, the spine is stabilized and fused. Fusion is typically performed by using autologous bone which is most frequently harvested from the iliac crest. The fibula can also be used as a donor site for autograft bone; however, this is usually reserved for cases that require more than a single level of fusion. Surgical hardware is utilized to help fixate bones in order to allow a fusion to occur. The hardware, however, is only a temporary fixation device that facilitates the eventual long-term bony fusion. From a posterior approach, techniques include use of interspinous wiring with bone grafting or placement of lateral mass plating with bone grafting.

Postoperatively, or if surgery is not required, an orthosis is usually prescribed, and maintained for approximately 3 months. The type of spinal orthotic chosen depends on the level of spinal injury. Generally, for the occipito-C2 levels the Halo-vest may be used, although some surgeons will utilize a head-cervical orthosis (HCO) (i.e., Miami J Collar (Jerome Medical)). An HCO is utilized for the C3-7 levels; for the T1-3 levels a cervicothoracic orthosis is used (i.e., extended HCO or Yale brace). From T4 through L2, a thoracolumbar spinal orthotic (TLSO) is utilized, however at L3 and below a lumbosacral orthotic (LSO) with the incorporation of one hip/thigh (spica attachment to a LSO or TLSO) will ensure satisfactory immobilization of the low lumbar and sacral spine is required.

Specific Injuries to the Spine

Fractures of the atlas are commonly referred to as “Jefferson” burst fractures. These are usually stable injuries (i.e., may occur after a football spearing injury) that may be treated with a Halo-vest orthosis. Unstable Jefferson fractures usually require posterior surgical stabilization. Odontoid fractures are classified into three basic types. Type I odontoid fractures are very rare and involve a fracture of the tip of the odontoid process. Type II odontoid fractures are much more common, particularly in the elderly population, and involve a fracture through the base of the odontoid process, at its junction with the C2 vertebral body. Type III fractures extend from the base of the odontoid into the body of the C2 vertebra proper. Type I odontoid fractures typically require no specific surgical intervention. Type III odontoid fractures are typically treated with an external orthosis (either Halo-vest or HCO) for 3 months. Type II odontoid fractures may be treated with an external halo/vest orthosis; however, there is a high failure rate with this treatment and internal stabilization may be needed.

Fractures of the pedicles of C2 are usually bilateral and are commonly referred to as “hangman’’s” fractures. These can occur during an abrupt deceleration injury, that is, a MVC with the person’s head hitting the windshield, and are most often stable injuries treated with external orthoses. When disruption is more significant, treatment with a halo in slight extension...
is utilized. In the case of an unstable hangman’s fracture, open surgical fusion may be necessary.

Pure bony injuries in the subaxial spine (C3-7) without substantial neurologic compression may heal with an external orthosis alone. However, most cases of an acute SCI secondary to cervical fractures will have ligamentous injuries and will require open surgical intervention to decompress and/or fuse the cervical spine. The most common burst fracture in the cervical spine occurs at the C5 vertebral level.

The most common thoracic spinal injury involves fracture of the T12 vertebra. Unstable injuries are treated with stabilization and fusion procedures. In the lower thoracic spine, either an anterior or posterior approach may be utilized. Anterior surgery is, typically, performed via a thoracotomy, or a thoracoabdominal approach, when a corpectomy is performed. This is followed by bone grafting and stabilization with either a screw/plate or a screw/rod construct. Anterior surgery in the upper thoracic spine is difficult technically and is rarely performed. Posterior thoracic surgery will classically involve the use of a hook/rod stabilization construct. Stable fractures can be treated with an orthosis (i.e., custom-molded TLSO).

A Chance fracture involves a horizontal splitting of the vertebra extending from posterior to anterior through the spinous process, pedicles, and vertebral body. Despite the extent of vertebral damage, these fractures tend to be stable. They most commonly occur at the thoracolumbar spine (T12, L1, or L2 level). Chance fractures usually result from an acute hyperflexion of the back and were known as “seat-belt” fractures with the advent of lap seat belts in cars. A head-on collision would cause the passenger wearing a lap-belt to suddenly be flexed at the waist, creating stress on the posterior elements of the vertebra. With the combination of shoulder and lap belts, most Chance fractures seen today result from falls or crush-type injuries where the thorax is acutely hyperflexed.

In the lumbar spine, L1 burst fractures are most common. These frequently result from a fall from a height and will result in an injury partially affecting the conus medullaris and/or the cauda equina (CE). Anterior surgery in the lower lumbar spine is difficult due to the presence of the iliopsoas muscles and the great vessels, and as such, below the level of L3 is rarely performed.

The role and timing of spinal surgery including decompression has not been fully clarified in patients with SCI. Early spinal decompression within 24 hours, and perhaps within 8 hours, may improve neurological recovery, particularly in patients with incomplete injuries (24,43). Most previous studies have failed to demonstrate significant neurologic recovery with early surgery (<72 hours) among patients with complete or incomplete lesions (44,45). Preliminary results of the STASCIS (Surgical Treatment of Acute SCI Study) trial showed that 25% of patients who received decompressive surgery within 24 hours of their injury experienced a 2-grade or greater improvement on the ASIA scale, compared with 4% of those in the delayed-treatment (>24 hours) group (46). Further study is needed regarding the timing of surgery and neurological recovery. Data have shown that early surgery reduces LOS in the acute hospital, facilitates rehabilitation, decreases hospital costs, and reduces postoperative complications (47–49). The current indication for emergent surgical treatment is progressive neurologic deterioration.
fluid (CSF) by lumbar puncture, usually performed between the L3 and L4 lumbar vertebrae.

Thirty-one pairs of spinal nerves (8 cervical, 12 thoracic, 5 lumbar, 5 sacral, and 1 coccygeal pair) emerge from the spinal cord. At and below the thoracic level, the nerve roots exit just caudal to the corresponding vertebra while in the cervical region the nerve roots exit through the intervertebral foramina just rostral to the corresponding vertebra. This is due to the fact that there are eight cervical nerve roots and only seven cervical vertebrae; the C8 nerve root exits through the intervertebral foramen just rostral to the first thoracic vertebra.

The spinal cord has two enlargements: cervical and lumbar. The cervical enlargement includes the C5-T1 nerve roots to form the brachial plexus which innervates the upper extremities (UEs). The lumbar plexus (roots L1 to L4) and lumbosacral to form the brachial plexus which innervates the upper extremity. The cervical enlargement includes the C5-T1 nerve roots to form the brachial plexus which innervates the upper extremities (UEs). The lumbar plexus (roots L1 to L4) and lumbosacral to form the brachial plexus which innervates the upper extremities (UEs). The sacral spinal nerves emerge from the conus medullaris and contain PS and somatic motor fibers innervating the muscles of the bladder wall and external sphincter, respectively.

The spinal cord receives its blood supply from one anterior and two posterior spinal arteries (PSAs) as well as anterior and posterior radicular arteries. The anterior spinal artery (ASA) arises in the upper cervical region and is formed by the union of two branches of the vertebral arteries. The ASA supplies the anterior two thirds of the spinal cord including the gray matter and anterior and anterolateral white matter. The ASA varies in diameter according to its proximity to a major radicular artery. It usually is narrowest in the T4-8 region of the spinal cord. There are two PSAs that originate from the vertebral artery. The PSAs supply the posterior one third of the spinal cord consisting of postero lateral and posterior white matter of the spinal cord.

The blood supply from the anterior and posterior arteries is sufficient for the upper cervical segments. Segmental arteries that arise from the aorta supply the ASA and PSAs in the thoracic and lumbar regions. The radicular arteries arise from the vertebral, cervical, intercostal, lumbar, and sacral arteries and supply the remaining segments of the spinal cord. The major radicular artery that supplies the lumbosacral enlargement of the spinal cord is known as the artery of Adamkiewicz. It usually arises from the left intercostal or lumbar artery at the level of T6-L3 and provides the main blood supply to the lower two thirds of the spinal cord. There are less radicular arteries that supply the midthoracic region of the spinal cord and are smaller in diameter and therefore create a “watershed zone” of the spinal cord at this level. With clinical situations where there is low blood flow to the spinal cord (i.e., clamping of the aorta for surgery above the renal artery), this level of the cord is most affected (T4-6 level).

The internal structure of the spinal cord is such that a transverse section of the spinal cord reveals a butterfly-shaped central gray matter surrounded by white matter. The gray matter of the spinal cord contains cell bodies and primarily neurons, dendrites, and myelinated and unmyelinated axons. Autonomic neurons are located laterally and exit by the ventral root and innervate smooth muscle. Lower motor neurons (LMN) are located ventrally, exit by the ventral roots and innervate striated muscle. The white matter consists of ascending and descending bundles of myelinated and unmyelinated axons (tracts or fasciculi). The ascending pathways relay sensory information to the brain while the descending pathways relay motor information from the brain.

Sensory tracts or ascending pathways are composed of axons from peripheral sensory nerves whose cell bodies are located in the dorsal root ganglion (DRG) and ascend toward the brainstem. Receptors for pain and temperature enter the spinal cord and synapse in the dorsal horn of the gray matter. The fibers cross over within one to two vertebral segments, then travel in the lateral spinothalamic tract and ascend to the ventral posterolateral (VPL) nucleus of the thalamus. The fibers then ascend in the internal capsule to reach the postcentral gyrus, which is the primary somatic sensory area of the brain. Pressure and light touch (LT) fibers enter the cord in the same fashion, and pass into the ipsilateral dorsal white column and bifurcate. One branch immediately enters the dorsal horn gray matter, synapses, and crosses over within one to two segments, while the other branch remains ipsilateral, and ascends in the dorsal column for as many as ten spinal segments. The ipsilateral branch ultimately enters the dorsal horn, synapses, and crosses over to join the other branch in the ventral white column, forming the ventral spinothalamic tract. These axons travel in the same pathway as the lateral tract to reach the postcentral gyrus, which interprets these sensations.

The posterior columns transmit three different sensations: proprioception, fine touch, and vibration sense. Their nerve fibers reach the DRG and immediately pass into the ipsilateral dorsal white columns and ascend to the medulla. Axons that enter the cord at the sacral, lumbar, and lower thoracic levels are situated in the medial part of the dorsal column (i.e., the lower part of the body) called the fasciculus gracilis. Those axons that enter at the thoracic (above T6) and cervical levels are situated in the lateral part of the column (from the upper part of the body) and are termed the fasciculus cuneatus. Axons of each fasciculus synapse in the medulla and form a bundle termed the medial lemniscus, which ascends to the postcentral gyrus.

The cerebellum is the control center for the coordination of voluntary muscle activity, equilibrium, and muscle tone. The spinocerebellar tract is a set of axonal fibers originating in the spinal cord and terminating in the ipsilateral cerebellum that conveys information to the cerebellum about limb and joint positions (proprioception). The lateral corticospinal tract is the main tract for voluntary muscle activity. Its origin is the precentral gyrus of the frontal lobe of the brain. Their axons descend through the internal capsule to the medulla oblongata. Approximately 80% to 90% of the axons cross at the pyramidal decussation to the contralateral side of the medulla and descend in the lateral white columns of the spinal cord, in the lateral corticospinal tract. At each level of the spinal cord, the axons from the lateral tract peel off and enter the gray matter of the...
ventral horn and synapse with secondary neurons. The 10% to 20% of uncrossed axons that continue down on the same side of the cord travel in the ventral corticospinal tract. The axons of the ventral tract then cross over at the corresponding level of muscles that it innervates. Both tracts travel from the precentral gyrus to the ventral horn as a single uninterrupted neuron and are termed upper motor neurons (UMN), while the secondary neurons that they synapse on, are termed lower motor neurons.

**NEUROLOGICAL ASSESSMENT**

The most accurate way to document impairment in a person with a SCI is by performing a standardized neurological examination, as endorsed by the International Standards for Neurological Classification of SCI Patients (50). These standards provide basic definitions of the most common terms used by clinicians in the assessment of SCI and describe the neurological examination. Key terms are defined in Table 27-2. The examination is composed of sensory and motor components, and is performed with the patient in the supine position to be able to compare initial and follow-up exams. The information from this examination is recorded on a standardized flow sheet (Fig. 27-3) and helps determine the sensory, motor, and NLI, sensory and motor index scores, and to classify the impairment. An online course for examination and classification utilizing the Standards is available through the ASIA website (www.asia-spinalinjury.org).

**Sensory exam:** The sensory exam is performed separately for LT and pin prick (PP) modalities. Each of 28 dermatomes (Fig. 27-1) is tested and graded 0 for absent, 1 for impaired, 2 for normal (or intact), or NT for not testable. The face is used as the reference point for testing sensation in each dermatome. A grade of 2 indicates the sensation is equal to that of the face. For the pin exam, a grade of 1 indicates the ability to distinguish sharp from dull; however, the sensation is qualitatively different as compared to the face (i.e., either less sharp or hyperesthetic). If the patient cannot distinguish the sharp from the dull aspect of the safety pin used for testing, then the score is 0. In questionable cases, eight out of ten correct answers are suggested as a standard for accuracy. A score of 0 is also given if there is no sensation. For the LT exam, a cotton tip applicator is used. A score of 1 is recorded if the sensation is less than on the face and a 0 if there is no sensation at all. The lowest sacral segment, S4-5 (anal musculocutaneous junction), should be tested with the pin and cotton swab as well. It is important to document the modality of sensation spared, as preservation

<table>
<thead>
<tr>
<th>TABLE 27.2</th>
<th>Glossary of Key Terms</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Key muscle group:</strong> Ten muscle groups that are tested as part of the standardized spinal cord examination.</td>
<td></td>
</tr>
<tr>
<td>Root Level</td>
<td>Muscle Group</td>
</tr>
<tr>
<td>C5</td>
<td>Elbow flexors</td>
</tr>
<tr>
<td>C6</td>
<td>Wrist extensors</td>
</tr>
<tr>
<td>C7</td>
<td>Elbow extensors</td>
</tr>
<tr>
<td>C8</td>
<td>Long finger flexors</td>
</tr>
<tr>
<td>T1</td>
<td>Small finger abductors</td>
</tr>
</tbody>
</table>

**Motor level:** The most caudal key muscle group that is graded 3/5 or greater with the segments cephalad graded normal (5/5) strength.

**Motor index score:** Calculated by adding the muscle scores of each key muscle group; a total score of 100 is possible.

**Sensory level:** The most caudal dermatome to have normal sensation for both PP/dull and LT on both sides.

**Sensory index score:** Calculated by adding the scores for each dermatome; a total score of 112 is possible for each PP and LT.

**Neurologic level of injury:** The most caudal level at which both motor and sensory modalities are intact.

**Complete injury:** The absence of sensory and motor function in the lowest sacral segments.

**Incomplete injury:** Preservation of motor and/or sensory function below the neurologic level that includes the lowest sacral segments.

**Skeletal level:** The level at which, by radiological examination, the greatest vertebral damage is found.

**Zone of partial preservation:** Used only with complete injuries, refers to the dermatomes and myotomes caudal to the neurological level that remain partially innervated. The most caudal segment with some sensory and/or motor function defines the extent of the ZPP.

of pin sensation in the lower sacral segments yields a better prognosis for neurological recovery.

To test for deep anal sensation (DAS), a rectal digital exam is performed. The patient is asked to report any sensory awareness, touch, or pressure, with firm pressure of the examiner’s digit on the rectal wall. DAS is recorded as either present (yes) or absent (no).

The maximum sensory score is 112 (calculated by adding the scores from the 28 dermatomes—maximum score of 56 for each side of the body) for LT and pin sensation. The sensory level is defined as the most caudal level where sensation for LT and PP are both graded as 2 (normal) for both sides of the body. If the LT level is C6 and PP is C5, the overall sensory level is C5. In a case where sensory loss begins at, or just above the nipple line (T4 dermatome), often a patient is credited with the T3 dermatome being spared. If the sensation is absent in the T1 and T2 dermatomes despite the presence of some sensation at the T3 dermatome it is recommended that the T3 dermatome be scored as absent. It is felt that this sparing above the nipple line is from C4 innervation.

**Motor exam**: The motor exam is conducted using conventional manual muscle testing (MMT) technique (on a scale from 0 to 5) in ten key muscle groups, five in the upper limb (C5-T1 myotomes) and five in the lower limb (L2-S1), on each side of the body (Table 27-2). Key muscles were chosen based upon their myotomal innervations and ability to be tested in the supine position. Most muscles are innervated by two root levels (e.g., the elbow flexors are innervated from C5 to C6). When a key muscle tests initially as a grade 5, it is presumed to be fully innervated by the contributions from the two roots. If a muscle initially grades a three fifth, it is presumed to have full innervation of its more proximal segment (in the case of the elbow flexors, innervation from the C5 myotome). The maximum motor index score is 100 (calculated by adding the scores—maximum of 50 for each side of the body). Voluntary anal contraction is tested by sensing contraction of the external anal sphincter around the examiner’s finger and graded as either present (yes) or absent.

When examining a patient with an acute injury below T8, the hip should not be flexed passively or actively beyond

---

**FIGURE 27-3.** ASIA flow sheet.
90 degrees as this may place too great a kyphotic stress on the lumbar spine. Therefore, it may only be possible to test hip flexor muscle strength isometrically. While asking the patient to lift the leg straight off the bed, the patient’s movement is resisted and the examiner’s judgment is required to grade the muscle force as 2 through 5.

If a muscle’s range is limited by contracture that exceeds 50% of the normal ROM, the muscle is to be listed as NT (not testable); if less than 50% loss of range—the MMT scoring can be applied. NT is also used if the muscle cannot be tested; that is, a cast in place limiting MMT. If spasticity interferes with the exam significantly, then NT should be documented. If pain limits full patient effort and the examiner feels that the initial contraction given represents normal strength, the muscle should be graded as 5*, while indicating the reason for this scoring (e.g., pain).

The motor level is defined as the most caudal motor level with a score of ≥3, with the more cephalad key muscles grading a 5. For injuries with no corresponding motor level (i.e., above C4, T2-L1), the last normal sensory level is used. For example, a person with normal strength in all key muscles of the UEs, 0/5 strength in the key muscles of the LES, and a T4 sensory level, would be assigned a motor level of T4. Similarly, in a case where the elbow flexors (C5 level) grade 3/5 on both sides, with a sensory level on the left at C4 and on the right C3; the motor level on the left would be C5 and on the right C3. This is due to the C4 dermatome on the right being scored as impaired—it is presumed that the C4 myotome is also impaired. Therefore, the motor level is designated as C3, since the patient does not meet the criteria of having a key muscle ≥3/5, with the levels above (in this case C4) scoring as normal. On the left side, the C4 dermatome is normal so the C4 myotome is considered normal and as a result, the left motor level is C5.

The NLI is the most caudal level, at which both motor and sensory modalities are intact on both sides of the body. The motor and sensory levels are the same in less than 50% of complete injuries, and the motor level may be multiple levels below the sensory level at 1-year postinjury (51). In cases where there is no key muscle level available (i.e., cervical levels at and above C4; T2-L1; and sacral levels below S2), the NLI is that which corresponds to the sensory level.

If there are non–SCI-related causes of weakness, this should be documented and taken into account when classifying the injury. For example, in a patient with a T8 fracture and complete paraplegia who also has a left brachial plexus injury, notation should be made that the sensory and motor deficits in the left arm are due to brachial plexus injury, not SCI, and the patient may still be classified with a NLI of T8. The motor level and UE motor index score better reflect the degree of function as well as the severity of impairment and disability, relative to the NLI, after motor complete tetraplegia (51).

**AIS Classification**

The International Standards have evolved over the past 25 years and have become accepted as the most appropriate method to describe the neurological impairment of SCI for clinical and research use and have been incorporated into the International Core SCI Data Set. In 1982, the ASIA first published the *Standards for Neurological Classification of SCI*, adopting the Frankel Scale (52). The Standards were replaced in 1992 by the AIS (53) and revised a number of times since (1996 and 2000) with reprinting in 2002 and 2006 (50,53,54).

The patient’s injury is classified utilizing the AIS, separating the injury into a neurologically complete versus incomplete injury. A neurologically complete injury is defined as an injury with the individual having no “sacral sparing.” Sacral sparing refers to having one or more of the following residual findings: LT or PP in the S4-5 dermatome (can be on either side, impaired or intact); DAS or voluntary anal contraction preserved. If any of these components are present, the individual has sacral sparing and therefore has a neurologically incomplete injury. Patients who have an incomplete injury initially (i.e., sacral sparing) have a significantly better prognosis for motor recovery than those without preservation of the lower sacral segments.

Table 27-3 describes the steps to classify the SCI and Table 27-4 outlines the AIS. A reference manual and training video are available that provide a detailed explanation of the examination elements (ASIA, Atlanta, GA. www.asia-spinalinjury.org). The zone of partial preservation (ZPP) refers to the dermatomes and myotomes caudal to the NLI that remain partially innervated in persons with a neurologically complete injury (AIS A). The ZPP should be recorded as the most caudal segment with some sensory and/or motor sparing but only in persons with a neurologically complete injury.

A neurologically complete injury is classified as AIS A. Persons with sensory sacral sparing are classified as an AIS B. To be classified with a motor incomplete injury (AIS C or D), the subject must have either (a) voluntary anal sphincter contraction or (b) sensory sacral sparing with sparing of motor function more than three levels below the motor level (50,54). To differentiate an AIS C from D, the individual with a motor incomplete injury AIS D has at least half of key muscles below the NLI with a muscle grade of 3 or more (AIS C would be less than half). It is important to recognize the distinction using the motor level to determine if one with sensory sacral sparing has a motor incomplete injury (AIS B vs. C), yet uses the NLI when differentiating an AIS C from D.

**Incomplete SCI Syndromes**

Incomplete SCI syndromes include central cord (CCS), Brown-Sequard, anterior cord, conus medullaris, and CE syndromes (CES), and can result from traumatic as well as NT injuries (55). CCS is the most common, accounting for approximately 50% of incomplete injuries and 9% of all traumatic SCI. CCS is characterized by motor weakness in the UE greater than the LE, in association with sacral sparing. CCS most frequently occurs in older persons with cervical spondylosis who suffer a hyperextension injury from a fall, but it may also occur in persons of any age and is associated with other etiologies, predisposing factors, and injury mechanisms (56). The postulated mechanism of the injury involves compression of the
CHAPTER 27  REHABILITATION OF SPINAL CORD INJURY

TABLE 27.3 Summary of the Steps to Determine the Classification of Individual with SCI

1. Determine sensory levels for right and left sides.
   Sensory level = Most caudal segment of the spinal cord where both PP and LT are normal and all rostral segments are normal.
   Start from the top of the worksheet, and go down until you see a “1” or “0” for either LT or sharp/dull, and go up one level. That is the sensory level.

2. Determine motor levels for right and left sides.
   Motor level = The lowest key muscle that has a grade of at least 3, providing that the key muscles … above that level are normal.
   Note: In regions where there is no myotome to test, the motor level is presumed to be the same as the sensory level.

3. Determine the single neurological level.
   Note: This is the lowest segment where motor and sensory function is normal on both sides, and is the most cephalad of the 4 levels.
   Single neurological level = The most rostral of sensory and motor levels.
   For example, if
   Sensory level: Right C4 Left C5
   Motor level: Right C5 Left C6
   then, single neurological level = C4

4. Determine whether the injury is complete or incomplete (sacral sparing).
   Sacral sparing = Sensory or motor function in the sacral segments, S4-5.
   • Sensory function = S4-5 dermatome or DAS.
   • Motor function = voluntary anal sphincter.

5. Determine AIS grade
   (a) Is injury incomplete? No, AIS = A. Record ZPP.
      If the injury is complete, the worksheet will read “N-0-0-0-0-N”
   (b) If yes, is injury motor incomplete? No, AIS = B
      (Yes = voluntary anal sphincter or motor function more than three levels below the motor level on a given side.)
   (c) If yes, are at least half of the key muscles below the (single) neurological level graded 3 or better? No, AIS = C
   (d) If yes, AIS = D.
   (e) If sensation and motor function is normal in all segments, AIS = E.
   Note: AIS E is used in follow-up testing when an individual with a documented SCI has recovered normal function. If at initial testing no deficits are found, the individual is neurologically intact; the AIS does not apply.

TABLE 27.4 ASIA Impairment Scale

<table>
<thead>
<tr>
<th>Grade</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Complete: No motor or sensory function is preserved in the sacral segments S4-5.</td>
</tr>
<tr>
<td>B</td>
<td>Incomplete: Sensory but not motor function preserved below the neurological level and includes the sacral segments S4-5.</td>
</tr>
<tr>
<td>C</td>
<td>Motor function is preserved below the neurological level, and more than half of the key muscles below the neurological level have a grade of &lt;3.</td>
</tr>
<tr>
<td>D</td>
<td>Incomplete: Motor function is preserved below the neurological level, and at least half of the key muscles below the neurological level have a grade of more than 3.</td>
</tr>
<tr>
<td>E</td>
<td>Normal: Motor and sensory function are normal.</td>
</tr>
</tbody>
</table>

Note: For an individual to receive a grade of C or D, he/she must be incomplete, that is, have sensory or motor function in the sacral segments S4-5. In addition, the individual must have either (a) voluntary anal sphincter contraction or (b) sparing of motor function more than three levels below the motor level.


It has been postulated that the pathophysiology of the upper limbs being more affected than that of the lower limbs is that the motor tracts of the upper limbs are located more centrally, while those of the lower limbs are located more peripherally in the spinal cord (56). MRI evidence has revealed that this clinical pattern may not be based on locations of the arm and leg fibers within the corticospinal tract, but rather that this tract carries fibers that mainly innervate distal limb musculature, and therefore, the functional deficits are more pronounced in the hands when the tract is the primary site of damage (57).

CCS generally has a favorable prognosis for functional recovery. Recovery occurs earliest and to the greatest extent in the LEs, followed by bowel and bladder function, proximal UE, and then distal hand function. Prognosis for functional recovery of ambulation, ADL, and bowel and bladder function is dependent upon the patient’s age (< or >50 years of age), with a less optimistic prognosis in older patients (>50 years old) relative to younger patients (55,58–60). Specifically, younger patients are more successful in becoming independent in ambulation (87% to 97% vs. 31% to 41%), bladder function (83% vs. 29%), and dressing (77% vs. 12%) (58) than older patients. Older newly injured individuals, however, with an initial classification of AIS D tetraplegia, have a good prognosis for recovery of independent ambulation (60).
Brown-Sequard syndrome (BSS) is defined as a lesion similar to a hemisection of the spinal cord, and accounts for 2% to 4% of all traumatic SCI (55,61–63). In the classic presentation, there is (a) ipsilateral loss of all sensory modalities at the level of the lesion; (b) ipsilateral flaccid paralysis at the level of the lesion; (c) ipsilateral loss of position, sense, and vibration below the lesion; (d) contralateral loss of pain and temperature below the lesion; and (e) ipsilateral motor loss below the level of the lesion. This is due to the crossing of the spinothalamic tracts in the spinal cord, as opposed to the corticospinal and dorsal columns that cross the brain stem. The pure form of BSS is rare and the Brown-Sequard Plus Syndrome is more common (BSPS) (64), which refers to a relative ipsilateral hemiplegia with a relative contralateral hemianalgesia. Although BSS has traditionally been associated with knife injuries, a variety of etiologies, including closed spinal injuries with or without vertebral fractures may be the cause (64,65).

Recovery usually takes place in the ipsilateral proximal extensors and then in the distal flexors (66,67). Motor recovery of any extremity having a pain/temp sensory deficit occurs before the opposite extremity. Overall, patients with BSS have the greatest prognosis for functional outcome of the incomplete syndromes, as 75% to 90% of patients ambulate independently at discharge from rehabilitation and nearly 70% perform functional skills and ADL independently (62,64). In predicting outcome, when the upper limb is the predominant site of weakness, patients are more likely to ambulate at discharge. Recovery of bowel and bladder function is also favorable.

The anterior cord syndrome involves a lesion affecting the anterior two thirds of the spinal cord while preserving the posterior columns. This may occur with flexion injuries, from retropulsed disc or bone fragments compressing the cord, direct injury to the anterior spinal cord, or with lesions of the ASA which provides blood supply to the anterior spinal cord (55,62,68). There is a variable loss of motor as well as PP sensation with a relative preservation of LT, proprioception, and deep-pressure sensation. Patients usually have a 10% to 20% chance of muscle recovery (69).

Posterior cord syndrome (PCS) has been omitted from recent versions of the International Standards, and is the least common of the SCI clinical syndromes with an incidence of less than 1% (55,62). There is a loss of proprioception and vibration sense, but with preservation of muscle strength, temperature, and pain sensation due to a selective lesion of the posterior columns. PCS has been linked to neck hyperextension injuries, PSA occlusion, tumors, disk compression, and vitamin B12 deficiency. Prognosis for ambulation is poor, secondary to the proprioceptive deficits.

Conus medullaris and Cauda Equina (CE) injuries. The conus medullaris, which is the terminal segment of the adult spinal cord, lies at the inferior aspect of the L1 vertebrae. The segment above the conus medullaris is termed the epiconus, consisting of spinal cord segments L4-S1. Lesions of the epiconus will affect the lower lumbar roots supplying muscles of the lower part of the leg and foot, with sparing of reflex function of sacral segments. The bulbocavernous (BC) reflex and micturition reflexes are preserved, representing an UMN or suprasacral lesion. Spasticity will most likely develop in sacral innervated segments (e.g., toe flexors, ankle plantar flexors, and hamstring muscles). Recovery is similar to other UMN spinal cord injuries.

Conus medullaris lesions present with UMN and LMN aspects. Lesions affecting neural segments S2 and below will present with LMN deficits of the anal sphincter and bladder due to damage of the anterior horn cells of S2-4. Bladder and rectal reflexes are diminished or absent, depending on the exact level and extent of the lesion. Motor strength in the legs and feet may remain intact if the nerve roots (L3-S2) are not affected, that is, “root escape.” Trauma and tumors are among the most common etiologies responsible for this condition.

Injuries below the L1 vertebral level usually affect the CE or nerve roots supplying the lumbar and sacral segments producing motor weakness and atrophy of the LEs (L2-S2) with bowel and bladder involvement (S2-4), impotence, and areflexia of the ankle and plantar reflexes. CES is a LMN injury. In CES there is a loss of anal and BC reflexes. CE injuries have a better prognosis relative to UMN injuries for neurological recovery, most likely due to the fact that the nerve roots are more resilient to injury as they are histologically peripheral nerves and therefore regeneration can occur. CE injuries may represent a neuropraxia or axonotmesis, and demonstrate progressive recovery over a course of weeks and months. CES can occur as a result of trauma, tumors, spinal stenosis, disc compression, infection, or postsurgical epidural hematoma (55).

Separation of CE and conus lesions in clinical practice is difficult, because some of the clinical features of these lesions overlap. Pain is uncommon in conus lesions but is frequently a complaint in CES. Sensory abnormalities occur in a saddle distribution in conus lesions and, if there is sparing, there is usually dissociated loss with a greater loss of pain and temperature while sparing touch sensation. In CE lesions, sensory loss occurs more in a root distribution and is not dissociated.

**FUNCTIONAL EVALUATION**

The International Standards (50) are the most widely accepted instrument of impairment, and interrater reliability is very good overall (70). The Modified Benzel Scale is a 7-grade scale used in the Sygen studies (71) that expands the AIS D category into three separate grades. The International Classification for Surgery of the Hand in Tetraplegia (72) is most commonly used when dealing with upper limb reconstruction procedures. A committee has developed a classification for autonomic functions, including blood pressure (BP), heart rate, and temperature regulation; bladder function; bowel function; and sexual function (70,73).

Functional limitation in SCI can be measured by the Grasp and Release Test that evaluates the impact of implanted upper limb prosthesis (74) or the capabilities of UE instrument...
(75,76). To measure walking in a standardized environment as it relates to walking, the Timed Get Up and Go test, 6-minute walk test or 10 m walk have been utilized (70,77). The Walking Scale for SCI (WISCI) is a valid scale that ranks walking based on various combinations of braces, assistive devices, and level of personal assistance (78,79).

Activity performance, what an individual can do in his/her environment, can be measured by the functional independence measure (FIM), the Canadian Occupational Performance Measure, the Quadriplegic Index of Function (QIF), and the spinal cord independence measure (SCIM) (70,80–84). The FIM as a generic instrument has shortcomings when applied to SCI and while the FIM was added to the Standards in 1992 (53), it was removed in the 2000 revisions (54). The QIF and SCIM are more sensitive and a better indicator of motor recovery as compared to the FIM, since it can reflect small gains in function which parallel small strength gains (84,85). The QIF is limited in use to persons with tetraplegia and has been used sparingly. The SCIM addresses indoor and outdoor activities (86) and is gaining in popularity.

Individuals with the same SCI level and severity may have different levels of activity performance due to differences in adaptive equipment, personal assistance, and accessibility of their environment. Impairments are not highly correlated with community integration, that is, the inability to walk does not prevent one from working. The CHART is the most commonly used measure of community functioning (participation) for persons with SCI (87). QOL is not related to impairment, but rather is highly associated with social support, community integration, and resumption of life and family roles (88).

**PROGNOSTICATING NEUROLOGICAL RECOVERY**

Prognosticating neurological recovery early after a traumatic SCI is important to provide accurate information to patients and their families and to guide the rehabilitation course. The major factors in predicting recovery early after a traumatic SCI include the initial NLI, the initial motor strength, and most importantly, whether by examination the injury is classified as neurologically complete or incomplete (89). This information is best provided by performing the neurological physical examination using the International Standards (90). The examination at 72 hours postinjury is superior to that at 24 hours postinjury for predicting recovery (91,92).

Between 30% and 80% of patients with initially complete tetraplegia will recover one root level of function. Most UE recovery occurs during the first 6 to 9 months, with the greatest rate of change during the first 3 months. Motor recovery can continue during the second year, especially for patients with initially 0/5 strength. The initial strength of the muscle, immediately below the motor level, is the most significant predictor of achieving antigravity strength by 1-year postinjury (93–95). If the first level below the NLI has 0/5 strength at 72 hours to 1 week, approximately one third of patients will recover to 3/5 strength in that muscle. If there is some initial strength present, approximately 75% to 95% of patients will regain antigravity strength at 1 year. At 1 month, if there is no motor strength in the first level below the NLI, then only 27% will improve to 3/5 by 1 year (94). Presence of sensation in the dermatome corresponding with that motor level, improves the chances of recovery. The faster an initial 0/5 muscle starts to recover some strength, the better the prognosis for recovery. Only a small percentage of subjects have motor recovery below the first level caudal to the motor level. Recovery at the C4 level to the C5 level, both motor and sensory, may be less and slower than at the lower cervical segments, especially if there is initially no strength at the C5 myotome (96).

Persons with incomplete tetraplegia have a better prognosis for recovery (97,98). UE motor recovery is approximately twice as great in incomplete tetraplegia as in complete tetraplegia, with the potential for varying degrees of LE motor recovery and functional ambulation. For patients who are sensory incomplete initially, the prognosis for motor recovery is more favorable in those with sparing of pin sensation rather than LT sensation alone (92,97–99). The basis of a more favorable outcome for pin sparing may be explained by the close anatomical relationship of the motor tracts (lateral corticospinal tract) to the sensory tracts carrying pain and temperature fibers (lateral spinothalamic tract). Ambulation can be predicted by having PP sensation sparing in 50% or more of the dermatomes in the LE (L2-S1) (99). Up to 40% of patients initially classified as AIS B will improve to AIS C and up to 40% to AIS D (92). Functional and neurological recovery is even more favorable for patients with an initial motor incomplete injury. Approximately 60% to 80% of patients initially classified as AIS C will improve to AIS D. The majority of motor recovery occurs within the first 6 months after injury, and the early return of motor function suggests a better functional outcome. Motor recovery in the UEs and LEs occurs concurrently, rather than sequentially.

In complete paraplegia, the potential for LE motor recovery improves with lower initial NLI; 15% of patients with a NLI between T9 and T11, and 55% of those with an initial NLI below T12, demonstrate recovery of some strength in the lower limbs (100). Most movement gained is in the proximal LE musculature and may represent recovery of partially injured lumbar roots or “root escape.”

Individuals with incomplete paraplegia have the best prognosis for LE motor recovery and ambulation (101). Eighty percent of individuals with incomplete paraplegia regain antigravity hip flexors and knee extensors at 1 year. Individuals with no LE motor control at 1 month may still show significant return by 1 year.

**AIS A Conversion and Late Recovery**

The vast majority of patients initially designated as having a complete injury will remain an AIS A. There are older reports of up to 10% of initial AIS A converting to AIS B and 10% regaining some motor function at later follow-up (92,102). Burns et al. reported that at year 1 or later, less than 7% (2/30)
of AIS A subjects initially tested within 2 days without factors affecting exam reliability, converted to AIS B status, and none developed volitional motor function below the zone of injury (103). Aspects of an early examination that may render it unreliable include factors that may affect cognition (i.e., TBI, drugs) and communication barriers (i.e., on ventilator). More recent literature has reported greater amount of conversion from initial to follow up examinations. A comprehensive review of the existing literature by the International Campaign for Cures of Spinal Cord Injury Paralysis (ICCP) reported a conversion rate of ~20% at one year post-injury, with 10% converting to AIS B and an additional 10% regained some motor function (AIS C and D) (104). The conversion rate of AIS A was twice as high for persons with tetraplegia compared to paraplegia. A recent evaluation of the European Multi-center study on human spinal cord injury (EM-SCI) reported that 70% of persons with initial (within 15 days) AIS A remained AIS A at one year, with 17.3% improving to AIS B, 5.8% to AIS C, and 7.2% to AIS D (105). Late conversion (i.e., change in status from a neurologically complete to incomplete injury after 30 days) has been reported to occur in 4% to 10% of cases (92,94,106,107). Late conversion can occur even years after injury, although usually only to AIS B with very few (<2%) to motor incomplete status, or C (108).

Other Predictors of Neurological Recovery

The presence of spinal shock may play a role in prognosis; for the same degree of SCI, the presence of spinal shock implies a more rapid evolution of injury and a poorer prognosis. The order that reflexes return in the postinjury period may help prognosticate outcome (24,109). Absence of the BC reflex (S3-4 roots) or the ankle reflex (S2-4 roots) after the acute period (24 to 72 hours) suggests injury to the conus medullaris or CE (i.e., LMN injury). As such, prognosis regarding recovery and also the potential use of rehabilitation intervention (e.g., electrical stimulation [ES]) can be determined. The delayed plantar response, which may be the first of all reflexes to return, occurs within hours or days following SCI, and its persistence shows a high correlation with complete injuries and a poor prognosis for LE motor recovery and function (ambulation) (109,110). The presence of the crossed adductor reflex to patellar tendon taps in the acute stage is highly predictive of functional motor recovery (111). Older individuals may have a less favorable outcome in regard to neurological recovery, functional ambulation, and bowel and bladder independence than younger patients with similar severity of the injury, and they have more associated medical complications (112).

Radiological and electrodiagnostic results early after injury assist in confirming the prognosis obtained from the clinical evaluation. The type of fracture often correlates with the severity of the injury. Bilateral cervical facet dislocation, thoracolumbar flexion/rotation injuries, and transcanal bullet locations more commonly occur with neurologically complete injuries. An intramedullary hemorrhage correlates with a more severe initial neurologic deficit, and carries a poorer prognosis. Hemorrhage location corresponds anatomically to the level of the neurologic injury. The greater the extent of cord signal abnormality on MRI, the greater the chance of having a complete injury (113–116). Lesions in patients with incomplete SCI (AIS B through E) were found to have a mean length of ≤20 mm, whereas those in patients with neurologically complete injuries had a mean length of 40 mm (116). The presence of cord edema greater than 1 vertebral body segment is also a poor prognostic finding. Smaller degrees of cord edema and especially the absence of abnormal signal intensity on MRI in the spinal cord are considered positive predictors for neurologic recovery. In the chronic stage after SCI, persons with persistent cord signal changes on MRI demonstrate little improvement in AIS grades relative to the improvements of patients with resolution of signal abnormalities.

Electrophysiological techniques include nerve conduction studies, late responses (H-reflex and F-wave), somatosensory-evoked potentials (SSEP), motor-evoked potentials (MEP), and sympathetic skin responses (SSR) that can supplement clinical and neuroradiological findings (89,117–122). These tests, however, are most useful in differentiating lesions between the central and peripheral nervous system. They may also help in diagnosing a neurologically complete versus incomplete injury in uncooperative or unconscious patients, or to rule out a conversion disorder, since they do not require the cooperation of the patient. They are not a routine part of the acute investigation of a newly injured individual to offer prognosis for neurological or functional outcome.

Nontraumatic Spinal Cord Injury

NT SCI is a growing population for admission to inpatient rehabilitation and includes such etiologies as spinal stenosis with myelopathy, spinal cord compression from a neoplasm, multiple sclerosis (MS), transverse myelitis (TM), infection (viral, bacterial, fungal, parasitic), vascular ischemia, radiation myelopathy, motor neuron diseases, syringomyelia, vitamin B12 deficiency, and others. Spinal stenosis and spinal cord tumors are the most common causes of NT SCI presenting for inpatient rehabilitation in the United States (123,124).

Details regarding most of these disorders are covered elsewhere in the text. TM is an inflammatory disorder of the spinal cord, with a female to male ratio of 4:1 that peaks in the second and fourth decades. TM most commonly affects the thoracic spine (50%) and diagnosis is best established by MRI. The time course of progression of symptoms that include back pain, motor weakness (the UEs are less often involved), abnormal sensations, and neurogenic bowel and bladder, is from hours to weeks. Although some patients recover with minor or no residual problems, most remain with residual permanent impairments. Poor recovery is predicted by rapid progression, back pain, and spinal shock (125). Most patients will have only one episode of TM, but a small percentage may have a recurrence.

SCI from epidural abscess is most commonly seen in immunocompromised and diabetic patients. Staphylococcus aureus is the most common pathogen. Epidural hematoma can be associated with anticoagulation, vascular malformations, or myelodysplastic diseases. Arterial disease from a thrombus or embolism to the spinal arteries or from other vascular diseases, including vasculitis and diabetes, or from thoracoabdominal aneurysm repair (reported in 9% to 18% of cases), may cause spinal cord ischemia/infarction.
Late radiation myelopathy is a delayed complication of radiation to lesions of the spine or adjacent tissues that develops months or years after treatment. The incidence of this complication is correlated with the total radiation dose, the dose fraction, and the length of the spinal cord irradiated. Clinically, there is weakness, loss of sensation, and sometimes a Brown-Séquard-like syndrome develops. The injury predominantly affects the white matter in the lateral spinal cord. The diagnosis of late radiation myelopathy remains one of exclusion and prognosis for significant recovery is poor.

Spinal cord tumors can be primary or metastatic, intradural, or extradural (see Chapter 43 for details). The majority of spine and spinal cord tumors are metastatic in origin and 95% are extradural. Approximately 70% of spinal metastasis occurs in the thoracic spine, with clinical presentation of pain, which is typically worse at night and in the supine position. Neurological changes including weakness, sensory loss, and bowel/bladder involvement indicates cord involvement. The most common sources of secondary tumors are metastatic lung, breast, and prostate. The strongest predictors of functional outcome are the tumor type and the preoperative neurological status (126).

As compared to persons with traumatic-related SCI, individuals with NT SCI are more likely to be older, female, married, and retired. NT SCI is more common than trauma in persons over 40. Persons with NT SCI usually have a less severe neurological impairment as compared with traumatic SCI, as they more often present with motor incomplete (90%) lesions (127). There is a lower incidence of secondary medical complications including spasticity, orthostasis, deep vein thrombosis (DVT), PUs, autonomic dysreflexia (AD), and wound infections during rehabilitation in patients with NT SCI (128–130). However, because cervical stenosis and neoplastic SCC has a peak incidence between the ages of 50 and 70 years, these individuals may have other premorbid medical issues that may impact their rehabilitation. No difference is noted for depression, urinary tract infections (UTIs), heterotopic ossification (HO), pain, or gastrointestinal (GI) bleeds.

No statistical differences were found in acute care LOS and admission to rehabilitation FIM scores (123,127,131) in patients with NT SCI relative to traumatic SCI. Inpatient rehabilitation LOS are shorter for persons with NT injury secondary to tumors, but FIM efficiency and home discharge rates are overall comparable to traumatic SCI. There is a favorable discharge to home for patients who have an incomplete injury, are married, have an established bowel and bladder management program, intact skin, male gender, and who are cognitively intact.

### Rehabilitation of SCI

Rehabilitation begins in the intensive care setting and includes addressing the SCI-specific needs to help the individual meet their potential in terms of medical, physical, social, emotional, recreational, vocational, and functional recovery (132). If early medical complications can be prevented, the inpatient rehabilitation course is facilitated, and the total cost of care is lessened.

The medical aspects of the SCI specialist’s acute recommendations can be formulated into a problem list (Table 27.5). The most important aspects include bowel, bladder and pulmonary management, deep venous thrombosis, GI prophylaxis, and proper positioning in bed to prevent contractures and skin breakdown, with the proper turning frequency (at least every 2 hours). Once the spine is stabilized by surgery or orthoses, therapy and nursing will incorporate ROM that will help prevent contractures. The shoulder, elbow, hip flexors, and heel cords are most important to range because they are the most
frequently observed contractures on presentation to the acute rehabilitation unit and can potentially serve as a source of pain and functional limitation. Resting splints for paralyzed UEs may also be useful for feeding and other self-care skills in the early period. For persons with high level tetraplegia who have undergone tracheostomy, early introduction of communication aids is important. Once medically stable, the patient should be transferred to a specialized spinal cord rehabilitation unit.

A specialized SCI center is most beneficial to provide for comprehensive management as it offers access to experienced SCI physicians and therapists, including psychology, vocational and SCI educational services, an active peer support program, and an opportunity to undergo rehabilitation with other patients who have similar impairments (133). A comprehensive SCI education program is essential for educating the patients and their families about SCI-related issues. Additional opportunities available at larger SCI centers include availability of trial equipment for mobility and accessibility to high level assistive technology. The interdisciplinary approach of the rehabilitation team, including the patient and family, is important for the optimal care of the individual with SCI. A sample rehabilitation prescription is listed in Table 27-6. As the LOS shortens in acute rehabilitation, coordination and communication with the entire team is needed to allow for a timely and safe discharge. Frequent team conferences with an early home evaluation are recommended.

**Functional Goals**

Once the patient’s motor level of injury, AIS, and prognosis for neurological recovery are determined at the onset of rehabilitation, short- and long-term functional goals are formulated and a therapy prescription established. The CPG on outcomes following traumatic SCI delineates expected (at 1 year postinjury) functional outcomes, and equipment needs based on the level of injury for persons with SCI with a motor complete injury (134). Table 27-7 lists the expected functional outcome at 1 year postinjury and Table 27-8 lists equipment usually prescribed for persons for each level of injury.

The projected long-term goals are a starting point for the rehabilitation prescription. The rehabilitation program should be individualized to meet each person’s strengths, weaknesses, and individual circumstances. Monitoring progress at the team conference, with the patient as an active participant, helps to identify limiting factors along with the patients’ additional needs. Discharge planning should be discussed as early as the first team conference, assuring a timely, but most importantly, a safe discharge. The ideal outcome may not always be achieved for each patient, as there is a significant amount of variability in individual outcomes despite similar levels of injury based upon the age, gender, and medical comorbidities.

**C1-4 Level**

Persons with motor levels above C3 will usually require long-term ventilator assistance, whereas most individuals with lesions at C4 will be able to wean off the ventilator. Respiratory equipment including a ventilator, a method for secretion management (i.e., suction machine or mechanical insufflator/expireflator), back-up ventilator batteries, and a generator in case of power failure, should be obtained. Contact should be made with the local power company and emergency services to alert them of the patient’s needs prior to discharge.

Rehabilitation goals for persons with high cervical SCI primarily include prevention of secondary medical complications, education and training of the patient and family members, prescription of appropriate durable medical equipment (DME), and environmental modification. The patient should be independent in instructing others in providing care including weight

---

**TABLE 27.6 Sample Physical and Occupational Therapy Prescription**

| Diagnosis: C7 ASIA A Tetraplegia |
| Goals: (see outlined goals) |
| Precautions: Skin, respiratory, sensory, orthostasis, safety, risk for AD, and others as needed for the specific patient (i.e., bleeding if on Coumadin) |
| Physical therapy |
| PROM to bilateral LE, with stretching of hamstrings and hip extensors. |
| Mat activities. |
| Tilt table as tolerated. Start at 15 degrees, progress 10 degrees every 15 min within precautions, up to 80 degrees. |
| Sitting balancing (static and dynamic). |
| Transfer training from all surfaces including mat, bed, wheelchair, and floor. |
| Wheelchair propulsion training and management. |
| Teach and encourage weight shifting. |
| Standing table as tolerated. |
| Deep breathing exercises. |
| FES for appropriate candidates. |
| Family training. |
| Community skills. |
| Teach home exercise program. |
| Occupational therapy |
| Passive, active assisted, active ROM/exercises to bilateral UEs. |
| Allow for some finger tightness to enhance grasp. |
| Bilateral UE strengthening. |
| Motor coordination skills. |
| ADL program with adaptive equipment as needed (dressing, grooming, feeding). |
| Functional transfer training (bathroom, tub, car, etc.). |
| Splinting and adaptive equipment evaluation. |
| Desktop skills. |
| Shower program. |
| Kitchen and homemaking skills. |
| Assistive technology devices, as needed. |
| Wheelchair training (parts and management). |
| Home evaluation. |
| Family training. |
| Teach home exercise program. |
shifts, ROM, positioning, donning orthoses, transfers, and in setting up their environmental control unit (ECU). Additional goals include independence in power wheelchair mobility, using breath control, mouth stick, head array, tongue, or chin control mechanisms. The wheelchair should be equipped with a pressure relief cushion and recline and/or tilt features for independent pressure relief. If the patient can control a power chair, then both a power chair and a manual positional wheelchair with a high back and tilt or recline should be prescribed for use as a back-up wheelchair for assisted mobility in the home and in the community as needed (136). Once properly set up, persons at these levels of injury should be independent in using assistive technology. This includes lower level technology devices (i.e., adapted phones and page turners) and higher level devices (known as electronic aides of daily living [e-ADL]) that control one or more electronic appliances (i.e., television, radio, lights) via voice activation or switch access. These will allow for independence in communication and controlling their local environment. A type of lift to assist in transfers and a padded commode/shower chair should be prescribed. An attendant-operated van with a lift and tie-down, or accessible public transportation is necessary for community mobility.
The patient’s family and social support, financial resources, personal preferences, educational, vocational and avocational goals, and living arrangements after discharge must be fully considered during rehabilitation. Post-acute medical, psycho-social, and rehabilitation care should be prescribed in the home or outpatient setting. Although the evaluation process requires input from many specialist team members, the physiatrist is ultimately responsible for medical justification of all equipment decisions and should be directly involved in the DME evaluation and prescription process.

<table>
<thead>
<tr>
<th>Orthotics</th>
<th>C1-4</th>
<th>C5</th>
<th>C6</th>
<th>C7</th>
<th>C8-T1</th>
</tr>
</thead>
<tbody>
<tr>
<td>BFO (mobile arm support)</td>
<td>X</td>
<td>+/–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Resting hand splint</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Long opponens splint</td>
<td>X</td>
<td>X</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Spiral splint</td>
<td>–</td>
<td>X</td>
<td>X</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Powered tenodesis splint</td>
<td>–</td>
<td>X</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Wrist-driven tenodesis splint</td>
<td>–</td>
<td>–</td>
<td>X</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Rachet tenodesis splint</td>
<td>–</td>
<td>X</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Short opponens splint</td>
<td>–</td>
<td>–</td>
<td>X</td>
<td>X</td>
<td>–</td>
</tr>
<tr>
<td>Universal cuff</td>
<td>–</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>–</td>
</tr>
<tr>
<td>Lumbrical Bar</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>X</td>
</tr>
<tr>
<td>Mouthstick</td>
<td>X</td>
<td>+/–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

| Transfers/mobility | | | | | |
| Power/mechanical lift | X | X | X | – | – |
| Transfer board | X | X | X | X | X |
| Power w/c with tilt/recline | X | X | +/– | – | – |
| Power w/c | – | – | X | +/– | – |
| Power assist w/c | – | X | X | +/– | – |
| Manual w/c | X | X | X | X | X |

| Feeding | | | | | |
| Adapated equipment (plate, etc.) | X | X | – | – | – |
| Utens. with built-up handles | – | – | X | X | – |

| Grooming and dressing | | | | | |
| ADL splints (wash mitt, razor holders) | – | X | X | – | – |
| Dressing equipment (pant loops, sock aide, dressing stick, long shoe horn, etc.) | – | – | X | X | – |

| Communication | | | | | |
| Environmental control unit | X | X | +/– | – | – |
| Computer | X | X | X | X | – |
| Book holder | X | X | X | X | – |

| Bathing | | | | | |
| Grab bars | – | – | X | X | X |
| Recl shower/commode chair | X | – | – | – | – |
| Tub seat/shower chair (padded) | – | X | X | X | X |
| Hand-held spray attachment | X | X | X | X | X |

| Beds | | | | | |
| Full electric hospital bed | X | X | X | – | – |
| Full specialized mattress | X | X | +/– | +/– | – |
| Overlay mattress | – | – | X | X | +/– |

Restoring respiratory function for persons with high level of injuries (C2 and above) have been achieved by the use of various ES (see later under pulmonary section). These procedures may assist in weaning off of the ventilator, and enhance speech production and improve QOL (135,136). Persons with a NLI at C4 who have some elbow flexion and deltoid strength may be able to use a mobile arm support (MAS) or balanced forearm orthosis (BFO) to assist with feeding, grooming, and hygiene. Once the elbow flexors have antigravity strength with adequate endurance, these devices are
no longer needed. A long straw or a bottle that the person can easily access to drink fluids should be obtained as early as possible.

The benefit of specialized acute inpatient rehabilitation for persons with such high levels of injury is justifiable despite their inability to initially tolerate 3 hours a day of therapy and having what may seem as limited goals. The SCI medical and nursing care during the first few months after injury are crucial for monitoring, treating, and preventing medical complications that can lead to future morbidity and mortality. Patient and family education, emotional and social support, and exposure to advanced technology that may allow independence in the proper environment (i.e., power mobility, assistive technology) may be the difference between returning to their family/community and living in a long-term facility.

C5 Level
The C5 motor level adds the key muscle group of the elbow flexors (biceps), as well as the deltoids, rhomboids, and partial innervation of the brachialis, brachioradialis (BR), supraspinatus, infraspinatus, and serratus anterior. It is important during the acute period after SCI to prevent elbow flexion and forearm supination contractures caused by unopposed flexion activity by stretching, splinting, and if needed antispasticity injections. A long opponens splint, with a pocket for inserting different utensils (i.e., Universal cuff), is important to assist with many tasks including feeding, hygiene, grooming, and writing. Most functional activities will require the use of assistive devices; however, tendon transfers may be considered after neurological recovery is complete.

The addition of the elbow flexors should allow for use of a joystick for a power wheelchair and can allow manual wheelchair propulsion on level surfaces with either rim projections (lugs) or plastic-coated handrims with a protective glove. A power wheelchair, with a power recline and/or tilt mechanism, is usually still required in addition to the manual wheelchair. Push rim-activated power-assist wheelchairs may also be advantageous. These are manual wheelchairs that have a motor linked to each rear hub. With each manual propulsion by the user, there is supplementary power provided by the motor. Therefore, the force required by the user for propulsion over the same distance is decreased when compared to a regular manual wheelchair. This feature is particularly useful for those with tetraplegia or those with paraplegia, and overuse injury causing shoulder pain. The use of power-assist wheelchairs can improve the ADL in persons with tetraplegia, when compared to the use of regular manual wheelchairs (137).

Another recent significant wheelchair development is the Independence iBOT 4000 Mobility System. One of its unique functions is the ability to negotiate curbs and stairs. The iBOT is not necessarily appropriate for all power wheelchair users, as there are certain prerequisites that the user must meet. In order to climb the stairs, adequate UE function to grab the handrail and to stabilize the iBOT while initiating the climbing movement is necessary or this can be performed with the assistance of a caregiver.

Persons with this level of injury will require almost total assistance for their bowel program. A padded commode/shower chair is recommended as the gravity from the upright position assists with the program and the padding helps to prevent skin breakdown. Bladder management is a decision based upon discussion with the SCI specialist and urologist, urodynamic results, amount of assistance available, and lifestyle circumstances. Intermittent catheterization (IC) usually cannot be performed independently, and will need to be performed by another person. If using a leg drainage bag, electronic devices to help empty the bag are available. Driving a specially modified van is possible at this level, with a lift for access allowing the patient to be fully independent in this activity.

C6 Level
The C6 level adds the key muscle group that performs wrist extension (extensor carpi radialis), as well as partially innervating the supinator, pronator teres, and latissimus dorsi. Active wrist extension can allow for tenodesis, the opposition of the thumb and index finger with flexion as the tendons are stretched with wrist extension. One should avoid overly stretching the finger flexors initially after injury (“selective tightening”) in C5 and C6 motor level patients to avoid potentially losing the tenodesis action. Tenodesis may allow some individuals with this level of injury to perform activities without splints, that is, feeding. Tenodesis splints can be fabricated but are frequently discarded by patients.

Feeding, grooming, and UE hygiene are usually independent after assistance with setting up the utensils; however, clothing modifications such as Velcro closures on shoes, loops on zippers, and pullover garments are recommended. Assistance for meal preparation and for other homemaking tasks is still required. Transfers may be possible using a transfer board and with loops for LE management, but most often requires assistance. Although persons with a C6 motor level can propel a manual wheelchair with plastic-coated rims, a power wheelchair is often required for long distances, especially if the individual will be returning to the workplace. Power-assist wheels may be of benefit as well. IC may be possible for males after assistance for set-up with assistive devices (138) but this technique is more difficult for females.

C7, C8 Levels
The C7 motor level adds the elbow extensors (triceps) as the key muscle group; C8 the long finger flexors. The C7 level is considered the key level for becoming independent in most activities at the wheelchair level, including weight shifts, transfers between level surfaces, feeding, grooming, upper body dressing, and light meal preparation (139). Uneven surface transfers, lower body dressing, and house cleaning may require some assistance. The independent use of a car is possible if the individual can transfer and load/unload their wheelchair.

IC in males can be performed although it is more difficult for females, especially if LE spasticity is present. Surgical options for females including a continent urinary diversion with an umbilical stoma can allow for a cosmetic means to
catheterize easily (140). Bowel care on a padded commode seat, especially suppository insertion, may still require assistance or the use of adaptive devices (i.e., suppository inserter).

**T1-12 Thoracic Levels**

Individuals with all levels of paraplegia should be independent with basic ADL including LE dressing, and mobility skills at the wheelchair level on even and uneven surfaces. This includes advanced wheelchair techniques such as curbs, ramps, wheelies (balancing the wheelchair on the two rear wheels), and floor to wheelchair transfers. Bowel and bladder management should be independent.

For most individuals with higher levels of complete thoracic injury, community ambulation is not a functional long-term goal. The lower the level of injury, the greater the trunk control due to abdominal and paraspinal muscle innervation. For lower levels of thoracic injuries, there is improved trunk control to allow for ambulation training with bilateral LE orthoses, as an exercise and for short distance household ambulation, once they have mastered basic wheelchair mobility skills.

**L1-2 Levels**

Muscles gained at these levels include the hip flexors and part of the quadriceps. While the person may be able to ambulate for short distances, a wheelchair will still be required for functional mobility needs. Bladder care is usually by IC. Individuals with these levels of injury can drive a car with hand controls.

**L3-4 Levels**

The knee extensors are fully innervated with some strength of ankle dorsiflexion (L4 myotome). Ambulation usually requires ankle foot orthoses (AFOs) with canes and crutches. Bowel and bladder management should be independent. These injuries are, typically, LMN in nature, and bowel management is usually by contraction of abdominal muscles and manual disimpaction. Suppositories will not be effective because of the loss of sacral reflexes. Bladder management is usually performed via IC or valsalva maneuver if postvoid residuals are within normal limits and urological work-up reveals no contraindication to this method. Absorbent pads can be used.

**L5 and Below**

These individuals should be independent in all activities unless there are associated problems, that is, severe pain, cardiac conditions, etc.

**Specific Activities in Therapy**

**Range of Motion**

Shoulder ROM is important to prevent pain in persons with all levels of injury. In persons with C5 and C6 motor levels, ROM should be especially addressed to diminish the development of elbow flexion and supination contractures. As previously mentioned, for those individuals with active wrist extension and weak or no finger function, the finger flexors should not be fully stretched, but instead, allowed to tighten somewhat and naturally curl to improve grip strength and function utilizing a “tenodesis” action. Lying prone when medically able is helpful in preventing and stretching the hip flexors. Prevention of heel cord tightness and contracture is important for proper positioning of the feet on the wheelchair footplate. Stretching of the lumbar spine is initiated when tightness or spasticity interferes with function, but is often avoided to provide the patient with increased postural stability and balance in the short and long sitting positions.

**Modalities**

Modalities such as heat, cold, and ultrasound can be used above the level of the sensory loss, but in areas with diminished or no sensation, one must be cautious as the risk of causing a burn is high.

**Pressure Reliefs**

Pressure reliefs are essential to prevent the occurrence of PUs. When supine, turning is initially performed every 2 hours and is progressively extended with close monitoring for signs of erythema that does not quickly dissipate (i.e., early stage one pressure sore). The use of a mirror will greatly enhance the ability to monitor the sacral area and ischial tuberosities (ITs). Persons with high level injuries (at C5 and above) will usually require a wheelchair with a tilt and/or recline mechanism. Some individuals with a C5 motor level can perform an anterior weight shift (with loops attached to the back of the wheelchair to assist in returning to the upright position), or a lateral weight shift. These weight shifts are more effective in pressure relief than having someone tip the wheelchair backward to 35 or 65 degrees (141). Individuals with an injury at C7 and below are usually able to perform independent push-up pressure relief, although other types of weight shifts are still recommended. A pressure relief should be performed for up to 2 minutes every 20 to 30 minutes while in the wheelchair (142). Computerized pressure mapping can be used to locate areas of high pressure in the patient’s wheelchair cushion and serve as a teaching tool to demonstrate the pressure relief technique that provides the best results (143).

**Transfers**

Initially, transfer training can be started on the mat, with subsequent progression to functional surfaces. Transfer training involves both the patient and the care providers. A lift transfer (one or two persons) is for individuals who are dependent for transfers. These can be performed with a lift that may be manual or electric; free standing, rolling, or attached to a structure (i.e., ceiling). Stand-pivot transfers are for patients who can weight-bear through the LEs with sufficient extension ROM in the hips and knees. Transfer (often referred to as “sliding”) board transfers allow transfer between seated surfaces, with the use of a board that bridges the gap between them. One should, however, be careful not to slide across the board, preventing shearing forces leading to skin breakdown. The board is a sturdy, smooth, flat surface that is made of wood or plastic and comes in different sizes, styles, and handgrips for different transfer situations. The use of the board may decrease
the pressure on the shoulders relative to a sit-pivot transfer. Sit-pivot transfers are performed by patients with a strong upper body, good short sitting balance, and good control of the head and shoulder movements. An overhead swivel trapeze bar can be attached to the bed, with the person lifting their body up swinging across to another surface. This can be a source of shoulder pain and is not recommended to be used on a consistent basis. Floor to chair/chair to floor transfers are particularly useful in accidental situations such as a fall, when the individual can then return to the wheelchair independently. This is a physically demanding technique, requiring strong upper body strength in both the arms and the shoulders.

When transferring into the wheelchair, positioning is critical. The wheelchair should be positioned next to the surface being transferred to, at about a 30 to 45 degrees angle to it. The brakes should be locked and the footrest and armrests should be out of the way. The transfer technique and method used should take into account the presence of PUs at surfaces such as the sacrum and the IT.

Generally speaking, individuals with paraplegia should be independent of transfers, whereas those with tetraplegia may have varying levels of transfer independence. In most instances, those with a NLI at or below C7 can perform independent transfers, with or without the use of a transfer board, though some individuals with a C6 injury may be independent with the board. Individuals with a C5 injury are expected to seek the assistance of one person with or without a transfer board, whereas individuals with an injury at or above C4 will be dependent. The exact transfer technique used by individuals with SCI varies tremendously. Devices such as cushions to raise the seat height, frames, and mechanical devices that can also assist individuals with SCI to stand up may be very helpful in transfers.

Standing
Standing with the use of a tilt-table or standing-frame, after an acute SCI, decreases hypercalciuria and may retard or lessen bone loss (144); however, it has not been shown to reverse osteoporosis in chronic SCI. Standing should proceed with caution in patients with chronic SCI since bone mineral density is often at or below fracture threshold (145). Additional physiologic benefits of standing include a decrease in spasticity, enhancing bowel and bladder programs secondary to the effect of gravity, and preventing PUs by facilitating pressure relief on normally weight-bearing areas (146,147). Tilt tables can be used early in the rehabilitation process in the treatment of orthostatic hypotension (OH).

Ambulation
There are four general levels of ambulation; community, household, ambulation for exercise, and nonambulatory (148). Community ambulation requires independence in performing transfers, capable of going from the sit to stand position, and ambulating unassisted in and outside the home for reasonable distances (>150 ft) with or without braces and assistive devices. Household ambulation is the ability to ambulate only within the home with relative independence, but may require assistance for transfers. Ambulation for exercise is for a person who requires significant assistance for ambulation.

Physiological benefits of walking includes potentially decreasing the progression of osteoporosis, reducing urinary calcinosis and spasticity, aiding in digestion and improving the bowel program because of the effect of gravity, and preventing PUs as with standing (149,150). In addition, standing and walking enables reaching objects not obtainable from the wheelchair level and affords access to areas that are not wheelchair-accessible, such as through narrow doorways or into a bathroom that is not appropriately modified. While ambulation following SCI has physiological and psychological benefits, it also has significant drawbacks including increased energy consumption, with decreased speed of ambulation when compared to the relatively normal energy expenditure and velocity of wheelchair use; weight-bearing through the UEs that may predispose an individual to shoulder, elbow, and wrist problems; and poor long-term follow through (151–153).

Community ambulation requires bilateral hip flexor strength >3/5, and one knee extensor to be at least 3/5, with a maximum amount of bracing of one long leg and one short leg brace (148). Prognosis for ambulation can be done early after injury and is determined by the initial level of the injury and the AIS classification (94,98,101). Forty-six per cent of persons with incomplete tetraplegia advance to community ambulation at 1 year, with an additional 14% performing household ambulation. Patients with initially motor complete injuries have a greater chance of ambulation than those with initial sensory incomplete injuries. For those with AIS B, PP sparing offers a better prognosis for regaining the ability to ambulate. Approximately 5% of complete paraplegics (the lower the level of injury, the greater the chance of ambulating) and 76% of incomplete paraplegics regain community ambulation. The percentage of persons with incomplete tetraplegia able to achieve community ambulation is lower than for incomplete paraplegia with equivalent lower extremity motor strength (LEMS), because the UE strength may be compromised and insufficient to enable assistive device ambulation if required. AIS grades at inpatient rehabilitation admission are helpful in predicting walking at discharge. Approximately 28% of persons with AIS C and 67% with AIS D at rehabilitation admission will regain the ability to ambulate by discharge. While level of injury does not significantly affect walking, age more than 50 does have a negative effect on walking in subjects with AIS D (154).

A number of orthotic options exist to assist in ambulation, including mechanicalorthoses, functional ES (FES), and hybrid orthoses (a combination of a mechanical orthosis and FES) (155). The knee-ankle-foot orthoses (KAFOs) are most frequently prescribed for ambulation. Other devices used occasionally include the Parawalker, the reciprocal gait orthosis (RGOs), advanced RGO (ARGO), hip-knee-ankle-foot orthoses (HKAFOs), and the hip guidance orthosis (HGO), that may enable persons with thoracic level paraplegia to ambulate. The Parastep system is the simplest example of a FES system for walking (156,157). Newer FES systems are being developed for both standing and walking (158) (see Chapter 48).
The use of partial body weight-supported treadmill training (PBWSTT) for ambulation as a rehabilitation intervention has gained a lot of recent attention. PBWSTT is based on the principle of generating normative, locomotor-like sensory input to promote the recovery of the spinal cord neural circuitry (159,160). With PBWSTT, the weight of the individual is partially supported by an overhead harness, while the therapists assist the hips and the legs of the individuals to walk on the treadmill. Studies in animal models and persons with incomplete SCI have shown improvement in acute and chronically injured individuals, in their ability to learn to walk (161,162). A multi-center trial of 146 subjects with acute neurologically incomplete injuries undergoing conventional training versus PBWSTT, did not find a significant difference in subjects regaining the ability to ambulate (163,164) although there are some significant limitations of this study. There is optimism that this form of therapy is beneficial and additional studies are underway to determine the effectiveness of this modality in the rehabilitation of the patient with acute and chronic SCI. PBWSTT with automated robotic systems are also available (e.g., Lokomat, Hocoma AG), although there is no evidence that robotic PBWSTT will produce superior outcomes to conventional PBWSTT (165).

**Assisted Technology Devices (ATDs)**

ATDs serve as the control center to allow for control of electrical appliances in the patient's environment, including the radio, television, bed, computer, lights, etc. The correct ATD can enhance a person's life by giving them a sense of control, security, and independence. One can use any body part to activate a switch as long as they can perform the activity consistently. Reliable activating sites include the head, chin, mouth, shoulder, arm, or hand. Voice activation is also an option. Prior to prescription, it is important to identify the patient's capabilities (i.e., cognitive status and functional movements) and needed tasks, the patient's goals, as well as any environmental barriers that may be present (166).

The use of brain-based command signals for controlling assistive technology, robotics, or neuroprostheses is a newer area of rehabilitation research that may prove to be useful for persons with high level tetraplegia. Brain signals are collected and processed through electrodes that may be placed or implanted at different levels. Once a signal is processed, it can potentially be used to control a number of devices, including computers linked to neuroprosthetic or robotic devices for assistance with ADL (167). There is no clinical product available for brain-based command signals at present and more research is necessary for its application.

**Home Modifications for People with SCI**

A number of general guidelines and recommendations for making the home accessible are listed in Table 27-9. A home evaluation or a floor plan of the residence should be completed early as an important aspect of the rehabilitation process to allow the injured individual to return home. The key pieces of information to know when performing the home evaluation are the patient's level of injury and mobility status, their prognosis for functional recovery, social situation for return to home, and financial considerations. The main areas to be evaluated include the entrances, bedroom, bathroom, and kitchen, and general safety issues ensuring that there is safe wheelchair access and egress space to maneuver a wheelchair in the home. The home should be free from fire, health and safety hazards, and an adequate heating, cooling, and electrical supply to meet the needs of additional medical equipment that must be present.

**Driving after SCI**

Most individuals with a motor level at or below C5 have the potential to return to independent driving with the appropriate adaptive equipment. The driver rehabilitation specialist evaluates the patient, assists in choosing the proper vehicle, and recommends modifications required including proper controls, lifts, lockdown, and tie-down devices.

---

**TABLE 27.9 General Recommendations for Accessibility in the Home**

- Outdoor walkways must accommodate a slope of 20° length to every 1" height as the maximum grade.
- Ramps must recognize a 12" length for every 1" rise.
- The minimum space for turning around is 5' × 5' for a manual wheelchair, and 6' × 6' for a power wheelchair.
- Doorway widths that require a “straight shot” (no turning involved) is 32″ for a manual chair and 34″ for a power wheelchair. This space increases to 38″ if there is a turn involved.
- All thresholds should be no greater than 1” to allow the person in the wheelchair to maneuver.
- Install carbon monoxide detectors and smoke alarms throughout the home.
- Low-pile carpeting or hard surface flooring is recommended for wheelchair maneuvering.
- There should be two means of exit from the residence if possible in case of emergency.
- Eliminate throw rugs as they pose a hazard for people who are walking and are difficult for people in wheelchairs to roll over.
- Remove or rearrange furniture that will impede wheelchair access.
- Notify police/fire departments that an individual with a disability resides in the home and provide the bedroom location.
- An intercom system can be useful to allow for communication.
- Backup power should be provided if the person with a SCI is dependent on equipment for life support, such as a ventilator.
- Light switches should be at a height of no more than 36”.
- Fireplace/heater cautions: Wheelchairs are generally constructed of some type of metal which may or may not conduct heat, therefore, when a person in a wheelchair is seated near a fireplace or heater, care should be taken to cover the metal parts of the wheelchair that may contact the person's skin, to ensure that the person in the wheelchair does not get burned.
- Power door openers can be installed for people in wheelchairs, where a remote can be used, or install a push plate on the wall.
A person with complete paraplegia and no additional complications will probably require mechanical hand controls and a few additional minor pieces of equipment to operate a car with an automatic transmission. The evaluation and training for such individuals can usually occur within 3 months of injury. Individuals with tetraplegia are usually evaluated at a later time postinjury (up to 1 year) to allow for neurological recovery to plateau, since any gain in motor function can mean a difference in equipment required. Higher level injuries (C5 and some C6) require more extensive adaptations allowing for acceleration, breaking, and steering. For lower level tetraplegia, steering may be facilitated by a spinner knob or other devices. In all cases, the person should be medically stable and be psychologically ready to return to the road.

A predriver assessment should be performed and includes a current and past driving history, current medications, a vision screen, physical skills testing (i.e., ROM, strength, sensory and proprioception testing, balance, spasticity, transfer, and wheelchair loading skills), wheelchair or mobility equipment required and reaction time (168,169). If the individual has a history of brain trauma, cognitive and perceptual screens should also be included. The behind the wheel assessment involves vehicle entry and exit, and operation of primary and secondary controls. Training time can range from a few hours for mechanical hand controls and standard steering, to over 40 hours for joystick drivers.

For most persons with paraplegia an automatic transmission car is an option if there are no problems with transfers. Transfer aids are available that can raise the person up to the level of a full size truck seat, but these are costly. If transfers or chair loading are more difficult, then vans should be considered. Loading devices can assist the client in loading the wheelchair. Options include car top devices that fold the manual wheelchair and stow it in a rooftop carrier, or a lift that can stow the folding chair behind the driver’s seat. Most persons with tetraplegia choose a modified van with power door openers and a ramp or lift. Persons with a NLI above C5, will require a van to accommodate their transportation needs, but will not be able to drive independently. It is usually easier to transport the person seated in their wheelchair that is already set up to provide the proper support. A structurally modified full-size van or lowered-floor minivan will usually be required. For the dependent passenger, the lowered floor drops their eye height to a point where they may be able to see out the side windows.

MEDICAL ISSUES AFTER SCI

Thromboembolic Disorders

Thromboembolic disorders, including DVT and pulmonary embolism (PE), are common medical complications after SCI. The incidence of DVT during the acute postinjury period in older studies varied depending on the method of detection used for screening, and has been reported in approximately 64% of patients (range of 47% to 100%) (170,171). Model system data reported an incidence of 9.8% for DVT and 2.6% for PE during acute rehabilitation (172). The development of DVT occurs most frequently during the first 2 weeks (approximately 80% of cases) following injury. PE occurs in approximately 5% to 8% of patients in the first year and is the third leading cause of death in all SCI patients in the first year postinjury (21,173).

The high incidence of DVT/PE in persons with SCI is related to Virchow’s triad that includes stasis, intimal injury, and hypercoagulability; which are all sequelae of acute neurologic injury. Suspicion should be high, and diagnostic testing ordered to make the proper diagnosis if suspected. Clinical signs of a DVT may include unilateral edema, low-grade fever, and pain/tenderness in a patient with an incomplete injury. However, the physical exam is limited in SCI because edema may be present secondary to immobilization and the patient may have loss of sensation. For this reason, screening for DVT is recommended at rehabilitation admission (by duplex ultrasound) and perhaps even during the rehabilitation course (174–176). Clinical signs of a PE may include fever, tachypnea, dyspnea, tachycardia, chest pain, or hypotension.

The consortium guideline on thromboembolism in SCI recommends that patients receive both a method of mechanical prophylaxis of DVT and anticoagulant prophylaxis (177). Pneumatic compressive devices are recommended during the first 2 weeks following injury. If this measure is delayed for more than 72 hours after injury, a duplex scan to exclude the presence of lower limb clots should be performed. It is not clear if patients who are participating in therapy and not utilizing these devices full time should use them on a limited basis. While ES has been shown to be effective, this is rarely used. Anticoagulant prophylaxis should be initiated within 72 hours following SCI provided there is no active bleeding or contraindications such as evidence of head injury or a coagulopathy. Most often low molecular weight heparins (LMWH) are used, with varying opinions as to specific medications and dosages, although dose adjusted or scheduled (two to three times per day) unfractionated heparin has also been utilized (178–183). Given recent warnings with use of LMWH after spinal surgery, clearance should be obtained from the spinal surgeon before initiating. Prophylaxis should be continued for persons with an AIS D injury while they are hospitalized. For those with motor incomplete AIS C or motor complete injury for 8 weeks, and for persons with motor complete injuries and other risk factors for 8 to 12 weeks (177).

Vena cava filter placement is recommended in patients who have failed anticoagulant prophylaxis, or who have a contraindication to anticoagulation. Routine placement of prophylactic caval filters is not warranted in all SCI patients (184). Filters should also be considered in individuals with high-level motor complete tetraplegia with poor cardiopulmonary reserve, patients with other high risk factors (i.e., long bone fractures), or with thrombosis of the inferior vena cava (IVC) despite anticoagulant prophylaxis. Filter placement is not a substitute for thromboprophylaxis, and may even be a risk factor for the development of a DVT. Treatment for a DVT or PE in SCI is similar to the non-SCI population and anticoagulation is usually continued for 6 months.
PART III | MAJOR CONDITIONS

Chronically, DVT is seen less frequently; 2.1% at 1 year follow-up, 1% at 2 years, and 0.5% at 5 year follow-up (173), but still higher than the general population. Patients with complete injuries have a higher incidence. The consortium guidelines recommend prophylaxis in persons with chronic SCI who are hospitalized and require immobilization (177).

PULMONARY SYSTEM

Respiratory failure is the leading cause of death in acute and chronic SCI (6,7,185,186). Approximately two thirds of all acute SCI patients will experience a complication of the pulmonary system including atelectasis, pneumonia, and/or respiratory failure requiring mechanical ventilation (187). The primary muscle for inspiration is the diaphragm, which receives its innervation from the phrenic nerve (C3-5), and contributes to 65% of the VC in able-bodied individuals. Additional muscles involved in inspiration include the external intercostals and accessory muscles including the scalenes, sternocleidomastoid, trapezius, and pectoralis. However, these alone are insufficient to maintain adequate oxygenation. Expiration is largely a passive activity caused by recoil of the chest wall. A forceful expiration, such as that required for effective cough, requires contraction of the abdominal and thoracic musculature, innervated by the thoracic level nerve roots (T1-11). The pattern of pulmonary dysfunction most commonly seen in SCI is one of restriction rather than obstruction.

The VC in the newly injured individual with tetraplegia is reduced but improves over time due to increased strength and the development of intercostal and abdominal tone that stabilizes the rib cage and enhances the mechanical effect of the diaphragm (188–190). The typical loss of VC in persons with a complete C5 motor lesion and above is roughly 50%; for C6-8 one third; and for T1-7 only slightly below the lower limits of normal (191–193).

Epidemiologic studies indicate that 20% to 25% of acutely injured SCI patients will require mechanical ventilation. Ventilation is usually initiated at lower tidal volumes (TVs) (6 to 8 cc/kg of ideal body weight) based on general ICU management, but many specialists in SCI suggest the volumes to be adjusted upward (10 to 20 cc/kg), while closely monitoring the patient pressures. Higher TV results in faster clearing of atelectasis and a better result in weaning off the ventilator (27,194,195). Oxygen should only be used as a temporary measure, as most patients have healthy lungs. If the oxygen saturation is low, the patient usually has secretions that need to be cleared or an insufficient TV.

Once patients are intubated, forced VC is a key parameter to follow. A VC of approximately 15 to 20 mg/kg is a good predictor of successful weaning from the ventilator (191). Patients with a neurologically complete NLI of C2 and above, will have no function of the diaphragm, and will need some type of ventilatory assistance immediately. While patients with injuries at the C4 or C5 level may initially require MV, most will be successfully weaned from the ventilator (196).

Weaning is usually attempted when the patient has a VC approximating 10 mL/kg body weight (195,197). The progressive free breathing technique is the recommended weaning protocol which consists of removing the patient from ventilator support with a gradual increase in time off the ventilator (186,198,199). This allows the patient to rest in-between trials, to gradually build up their strength, and to maintain expansion of the lungs in between the weaning trials. Other methods, including the use of noninvasive means of ventilation have been utilized. A peak cough flow of 160 L/min correlates with successful decannulation in persons with neuromuscular disease (200).

For patients who remain ventilator-dependent, direct stimulation of the phrenic nerves (phrenic nerve pacing) has been effective in facilitating ventilator weaning as well as improving speech and mobility (137,201,202). This method traditionally requires neurologically intact peripheral phrenic nerves (lesions above C2) that are stimulated via electrodes placed directly over the nerve, but intercostal nerve grafting for patients with phrenic nerve damage has been described (203). More recently, diaphragmatic pacing via electrodes placed intramuscularly over the phrenic nerve motor point has also been efficacious in ventilator weaning in persons with tetraplegia (204,205). Advantages of diaphragmatic pacing over phrenic nerve stimulation include laparoscopic approach versus the traditional thoracotomy required for placement of the phrenic nerve stimulators, as well as decreased risk of iatrogenic phrenic nerve damage (204,205). Noninvasive means of ventilation including intermittent positive pressure ventilation (IPPV) via mouthpiece, intermittent abdominal pressure ventilator, continuous positive airway pressure (CPAP), and biphasic PAP (BiPAP) can also be employed for persons with SCI.

Persons requiring MV have a high mortality rate, although greater numbers survive the initial injury. The 1-year survival for initially ventilated patients was approximately half of the 1-year survival for the total admissions to the model systems facilities (25.4% vs. 49.7%) (206). For patients who survive the first year on the ventilator, the 15-year survival rate is 61.4%. Survival of persons who are ventilator-dependent has improved during the first years after injury, but there is no evidence of improved long-term survival over the last few decades (207).

In the cervical and high thoracic level injured patients, the effective management of pulmonary secretions is critical. Manually assisted cough (i.e., "quad coughing") is performed by providing an upward thrust on the abdomen while the patient attempts to exhale. Due to possible displacement of an IVC filter, it is best to not use “quad coughing” in patients who have a new IVC filter placed (208). Prior to performing a cough, secretions should be mobilized by percussion or use of devices similar to the “pneumovest” or “rocking bed,” which loosen secretions by gentle vibration to the chest wall. Routine tracheal suctioning can also be used to extract secretions, but one should observe for reflex bradycardia secondary to vagal stimulation. The use of a mechanical insufflation/exsufflation (MI-E) device is beneficial in managing secretions and can be
used via tracheostomy, face mask, or mouthpiece. The MI-E delivers a deep insufflation (positive pressure) to the airway that is immediately followed by an exsufflation (negative pressure). The rapid shift in pressure produces a high expiratory flow rate from the lungs, simulating a cough. Advantages of MI-E over suctioning include better ability to clear secretions from the left lung, larger mucous plugs can be removed, and is more comfortable and better tolerated by patients (209).

VC in tetraplegia is affected by the position of the patient, with a 15% decrease in the upright position relative to supine. The diaphragm, like other muscles, is at a mechanical disadvantage near its end range. In SCI patients with mid-thoracic level injuries and higher, the diaphragm tends to partially collapse, placing it at a mechanical disadvantage and increasing the residual volume. An abdominal binder improves the respiratory function in the sitting position by placing the diaphragm in a more efficient position (210,211). The use of inspiratory resistance muscle training, abdominal weights, and incentive spirometry can improve pulmonary function (212). Glossopharyngeal breathing (GPB) is a technique that involves rapidly taking small gulps, 6 to 9 gulps of 60 to 200 mL each, and using the tongue and pharyngeal muscles to project the air past the glottis into the lungs. Many patients can use this technique to augment VC to assist with coughing or prolong ventilator free time (213).

Diseases of the respiratory system are the leading cause of death in chronic SCI, with the majority specifically due to pneumonia. Individuals with a history of ventilator usage at disease onset, a history of repeated atelectasis or pneumonia, a VC less than 2 L, nocturnal hypercapnea, or a mean SaO2 less than 95% are at risk to develop late onset respiratory insufficiency. Immunization for pneumonia (every 5 years) and influenza (yearly) are important for persons with SCI as they are considered to be a high risk population (214).

Sleep Apnea
Sleep apnea occurs in approximately 15% to 60% of patients with SCI (215,216). Common features include loud snoring, disrupted sleep, witnessed apnea, nocturnal gasping and choking, daytime sleepiness, and fatigue. It is commonly seen in older men with a short, thick neck, and tetraplegia. Certain medications (i.e., antispasticity and antiarrhythmics) and length of time from injury may also play a role. In SCI patients, the sleep apnea is primarily obstructive, with a small percentage of patients demonstrating central sleep apnea. Complications include daytime sleepiness and cognitive changes including poor attention, concentration, complex problem solving, short-term recall, and judgment. In addition, there is an increased risk for the development of hypertension, pulmonary hypertension, congestive heart failure, depression, and mortality.

It is advisable to review questions with the patient regarding sleep disturbance, snoring, and daytime somnolence. Optimal work-up would include an overnight oximetry recording, followed by a formal sleep study, if the oximetry recording is abnormal. Treatment includes assisted ventilation at night and possible use of medications directed toward relief of upper airway symptoms.

Orthostatic Hypotension
OH occurs when there is a decrease in systolic BP (SBP) by 20 mm Hg, or a decrease in diastolic blood pressure (DBP) of greater than 10 mm Hg or more (217). Many patients with acute SCI have a baseline SBP of 90 mm Hg or less, making symptoms a more reliable parameter to follow in diagnosis and treatment. SCI patients often experience symptomatic hypotension with position changes, especially moving from supine to more upright positions. Associated symptoms include lightheadedness, dizziness, ringing of the ears, fatigue, tachycardia, and sometimes syncope. In one study, OH occurred during 74% of therapy treatments in patients with acute SCI, with accompanying signs and symptoms noted on 59% of occasions, and were perceived as limiting treatment on 43% of occasions (218). OH can last for several weeks after the injury, delaying the rehabilitative process. OH occurs more frequently in persons with cervical level and neurologically complete injuries. When bed rest is prolonged, the degree of orthostasis tends to be more severe (219). OH intensifies after eating, exposure to warm environments, defecation, and rapid bladder emptying (220). Symptoms are related to a reduction in cerebral perfusion rather than a specific peripheral BP level.

Habituation to the symptoms of OH occurs slowly although the resting BP rarely returns to preinjury values, especially in person with tetraplegia. The exact mechanism is unknown, but theories include increased sensitivity of baroreceptors and catecholamine receptors in the vessel walls, development of spasticity, improved autoregulation of cerebral vascular perfusion, and adaptations of the renin-angiotensin system (221). Patients should be cautioned to avoid rapid changes in position. Simple adjustments, such as raising the head for several minutes prior to transferring out of bed, can be effective in decreasing episodes of OH.

Physical methods, including compression wraps to the legs and an abdominal binder donned prior to sitting up help to prevent venous distension and prolonged pooling of blood in the LEs, should be tried to mitigate orthostasis. Repeated postural changes on a tilt table or a high back reclining wheelchair also lessens the drop in BP. Maintaining adequate fluid intake is important and one should not be started on fluid restriction for an IC bladder program until the symptoms have improved. Avoiding diuretics such as alcohol and caffeine, and partaking in small meals to minimize postprandial hypotension are recommended. Sleeping with the bed head raised by 10 to 20 degrees should be encouraged to increase plasma volume and orthostatic tolerance.

Pharmacologic agents are added to the treatment regimen if the above interventions do not resolve the symptoms. Sodium chloride tablets (1 g four times per day), catecholamines such as midodrine hydrochloride (2.5 mg to 10 mg three times per day) are used; and if ineffective can initiate a salt-retaining mineralcorticoid such as fludrocortisone (0.05 to 0.1 mg daily) (222–224). The medication should be given approximately
1 hour prior to activity known to cause hypotensive episodes. Patients should be monitored closely for hypertension when taking these medications.

Chronic low resting BP may interfere with participation in activities that may provoke symptoms and impact QOL. Chronic hypotension may have a deleterious affect on the patient’s long-term health, as low resting systolic BP (<110 mm Hg) is associated with fatigue and can lead to deficits in cognitive performance (225). These patients may also be at greatest risk for other cardiovascular abnormalities.

**Autonomic Dysreflexia**

AD, also known as autonomic hyperreflexia, is a composite of symptoms, most notably a sudden rise in BP, is seen in SCI patients with injuries at or above T6, and is due to autonomic dysfunction. Individuals who are neurologically complete and at higher levels of injury are apt to have more severe symptoms (226,227). The incidence in susceptible patients varies between 48% and 90%, but rarely presents within the first month after injury, and nearly all patients who develop AD will do so within the first year (226–229).

AD is caused by a noxious stimulus below the level of injury. The most common source is from the bladder, either from over distension or infection, with the second most common source being the bowel (i.e., fecal impaction). These account for over 80% of all cases. Other causes include PUs, ingrown toenails, abdominal emergencies, fractures, and body positioning. In females, AD can occur during labor and delivery (230) and in males during stimulation techniques for ejaculation. The noxious stimulus travels via the peripheral nerves to the spinal cord and ascends in the spinothalamic tracts and dorsal columns where sympathetic neurons in the intermediolateral cell columns are stimulated. Splanchnic vessels, which are innervated by nerves originating from T5-L2, constrict and raise the BP. The hypertension from AD is defined as a rise of 20 to 40 mm Hg above baseline (231). Symptoms include most commonly a headache that is pounding in nature and found in the frontal and occipital areas, followed by sweating and flushing above the level of lesion. Baroreceptors from the carotid sinus and aortic arch detect the elevated BP and relay impulses to the vasomotor center in the brainstem that attempts to compensate by slowing the heart rate via impulses through the vagus nerve. Signals meant to inhibit the sympathetic system are blocked at the spinal cord lesion (Fig. 27-4).

Signs and symptoms reflect sympathetic overflow, and the consequent compensatory reaction mediated by the PS nervous system. The sympathetic discharge causes hypertension and piloerection. The PS response causes headache, pupillary constriction, sinus congestion, and bradycardia. More serious and even life-threatening symptoms may occur including cardiac arrhythmias, seizures, intracranial hemorrhage, pulmonary edema, and myocardial infarction (232).

AD requires immediate attention, with the goal to remove the inciting stimulus and reduce the BP. The BP and pulse should be monitored every 2 to 5 minutes until stabilized. A rapid survey of the patient for the causes should be initiated. The first intervention upon diagnosis is to sit the patient upright with loosening of all tight fitting clothes. The bladder should next be checked for over distension by flushing an indwelling catheter if one is present or catheterizing the patient if without one. If the catheter is blocked, gentle irrigation with a small amount of fluid, at body temperature, is used. If available, xylocaine gel (2%) should be used prior to catheterization. If the bladder is the source, the BP should quickly

![Figure 27-4. Autonomic dysreflexia.](image-url)
return to baseline. If the pressure remains elevated (i.e., systolic BP > 150 mm Hg), pharmacological management should be initiated prior to checking for fecal impaction. If so, one should use an antihypertensive agent with rapid onset and short duration while the causes of AD are being investigated. Prior to rectal examination, xylocaine gel should be placed within the anorectal area. If impaction is not present, other possible sources should then be sought, including PUs, ingrown toenails, infections, fractures, DVT, and HO.

Medications used for AD most commonly include nitrate gels (e.g., nitroglycerine paste), and less frequently clonidine, nifedipine, terazosin, β-blockers, phenoxybenzamine, and hydralazine. The use of nitrates is beneficial as they can be easily applied, titrated, and removed immediately if a source is later found. If using nifedipine, chew and swallow is the recommended route, rather than sublingual. The BP should be monitored frequently because hypotension may occur, especially if the source is found rapidly. The individual should be monitored for recurrent symptoms for at least 2 hours after resolution of the AD episode to ensure that it does not recur (231). If there is poor response to treatment and/or if the cause of the AD has not been identified, the patient should be hospitalized for monitoring, maintenance of pharmacologic control of BP, and investigation of other possible causes of the AD.

### Immobilization Hypercalcemia

After acute immobilization, calciuria increases within 2 weeks, reaching a maximum between 1 and 6 months after injury. The incidence of Hypercalcemia (level > 10.5) in SCI is 10% to 23%. Signs and symptoms usually present between 1 and 2 months postinjury, but may occur as early as 2 weeks and up to 6 months after injury. They can be vague and could include acute onset of nausea, vomiting, anorexia, lethargy, abdominal discomfort, constipation, diffuse musculoskeletal pain, polydipsia, and polyuria (which could result in dehydration). Risk factors include multiple fractures, younger age (<21 years) because of their high rate of bone turnover, male gender, high level lesion, complete neurologic injury, prolonged immobilization, and dehydration (233). Ingesting a high calcium diet does not increase either urinary or serum calcium concentration. For diagnosis, a serum calcium level should be evaluated, with correction for a low serum albumin if present. Ionized calcium can also be measured.

Treatment consists of IV fluid (normal saline at 100 to 150 cc/h), as tolerated, to increase the calcium excretion. A Foley catheter is recommended due to the volume of fluid being administered. Once the patient is hydrated, furosemide may be given to enhance calcium excretion, but often continued aggressive hydration alone is used. Calcium-sparing diuretics such as hydrochlorothiazide are contraindicated. Bisphosphonates such as pamidronate (30 to 90 mg IV administered over 4 to 24 hours) is effective, with the advantage of only requiring one dose, and has a rapid onset (234,235). This treatment rapidly lowers serum calcium within 3 days with the symptoms improving quickly as the serum calcium level drops. The serum calcium falls to a nadir within 7 days and may remain normal for several weeks or longer. Some may require repeat treatment, so continued monitoring of calcium levels following treatment is recommended. Standing has been reported to decrease hypercalciuria (236,237). Other medications that have been used include calcitonin, etidronate, and glucocorticoids (223,238). Treatment for asymptomatic hypercalcemia is recommended as prolonged hypercalcemia may cause nephrocalcinosis.

### GI Complications

Immediately after a SCI, an adynamic ileus is usually present and typically resolves within a week. Gastric decompression via nasogastric tube with intermittent suction may be necessary for persistent abdominal distention. Parenteral nutrition should be considered when the ileus lasts longer than 3 days. In cases where the ileus lasts longer, erythromycin and/or metoclopramide can be used to stimulate peristalsis (239). Neostigmine has also been effective in refractory cases of pseudo-obstruction (240).

Individuals with SCI have a higher risk of developing peptic ulcers compared with the general trauma population (241). The incidence is approximately 9% to 7%, although an incidence as high as 24% has been reported (242,243). Most ulcers occur acutely, within a few days from injury. Higher-level and more severe injuries (neurologically complete) have the greatest risk. The use of stress ulcer prophylaxis is recommended (histamine 2 blockers or proton pump inhibitors [PPI]) usually for 4 weeks unless other risk factors are present (24). PPIs are more effective in preventing upper GI bleeding in high risk patients; however, it may increase the rate of clostridium difficile infections.

Individuals with SCI have a higher risk (25% to 30%) of developing biliary stasis and gallstone disease compared to able-bodied persons (244,245). Impaired gallbladder motility has been implicated as the cause. Patients with lesions above T10 have impaired filling of the gallbladder and lower fasting volumes, which may make the bile more lithogenic. Due to altered sensation after SCI, the diagnosis may be more challenging. However, the majority of patients still presents with traditional symptoms, often with radiation of the pain to the right shoulder, and may still have findings such as right upper quadrant abdominal pain and tenderness (244,246). A high index of suspicion, as well as laboratory and radiographic tests are essential in making a timely diagnosis.

Superior mesenteric artery (SMA) syndrome is a condition where the third segment of the duodenum is compressed between the SMA and the aorta. It occurs rarely, though more commonly in persons with tetraplegia, with an incidence of less than 0.33%. The patient experiences nausea, vomiting, abdominal tightness, and bloating which is made worse after a meal or by lying supine. The diagnosis is made definitively with an upper GI series. Risk factors that may predispose patients to this condition include weight loss, use of a spinal orthotic, and immobilization in the supine position. Patients may gain some relief by lying on the left side and by taking metoclopramide (247).
Acute abdominal emergencies are often a challenge to diagnose, especially in patients with higher level SCI because of the absence of the usual signs and symptoms of abdominal pathology. It is, therefore, important to have a high suspicion in evaluating patients with SCI presenting with fever, abdominal pain, and/or pain radiating to the shoulders, and an elevated white blood cell count. Performing specialized lab tests, such as abdominal ultrasonography or CT scan early, is often helpful to confirm the diagnosis. Pancreatitis should be considered, especially in acute patients with unresolved ileus or recurrence of an adynamic ileus (248). The clinical recognition of acute pancreatitis is hampered by diminished or lost visceral sensitivity and therefore is based on laboratory investigations. Acute pancreatitis in the setting of high-level SCI may result from a combination of locally mediated sphincter of oddi dysfunction and vagal dominant innervation of the pancreatic gland in autonomic failure (249).

**Bowel Management after SCI**

Full details regarding bowel management are discussed in Chapter 27. An effective bowel management program is important to both the physical and psychological well being of the SCI patient. More than one third of persons with paraplegia ranked the loss of bowel and bladder control as the most significant functional loss associated with injury (250). Stool incontinence can be devastating to patients, leading to social isolation, loss of income secondary to work absenteeism, depression, and decreased QOL (251–253). The goal is to establish an effective “bowel program,” a consistent and complete evacuation of the bowel at a specified time, in a relatively short time period, without incontinence between programs. The bowel program should be individualized for every patient and often requires trial and error. The level of injury, previous bowel habits, lifestyle of the patient, and availability of caregivers should be taken into account when planning a bowel program.

Bowel programs are ideally performed on a daily to every third day basis depending on the individual’s desire or preinjury patterns. Sitting upright on a padded commode can facilitate a more natural position and is preferred to a side lying position. Key components of bowel management include the use of digital stimulation, dietary fiber intake, use of oral medications, and rectal evacuants. Fluid intake should be adequate to maintain soft stools and is recommended at 2 to 3 L/d, depending upon bladder management and restrictions.

The bowel program will vary depending on the type of neurogenic bowel dysfunction, that is, UMN versus LMN bowel. Patients with UMN injuries may have decreased GI motility, especially in the descending colon, but they will have normal to increased resting rectal tone and GI-related reflexes remain intact. These include the gastro-colic and recto-colic reflexes. To utilize the gastro-colic reflex (contraction of the colon occurring with gastric distension), patients should be instructed to perform their bowel program 20 to 30 minutes after eating. Caffeine may act as a stimulant and may be used prior to a bowel program to help facilitate fecal evacuation.

Dietary or supplemental fiber acts as a bulking agent and can enhance colonic transit time. It is suggested that the daily intake of fiber be at 30 g/d. A bowel program for patients with UMN injuries is initiated—the “3-2-1” program consisting of a stool softener (i.e., docusate sodium) three times per day, two senokot tablets given 8 hours before the suppository, and one suppository started once bowel sounds are present. An integral part of the bowel program is digital stimulation which is performed by inserting a gloved lubricated finger into the rectum and slowly rotating the finger in a circular motion in a clockwise direction until relaxation of the bowel wall is felt or stool or flatus passes (approximately 1 minute). This initiates the recto-colic reflex; contraction of the colon occurring with stimulation of the rectal mucosa. This will trigger evacuation of stool in the lower rectal vault. Once the vault is empty, a suppository is inserted to stimulate contraction of the lower colon and evacuation of stool located higher in the descending colon. Repeat digital stimulation (three to five times) should be performed every 10 to 15 minutes to check for stool which may remain in the rectal vault until there is closure around the finger by the internal sphincter or there are no results after two stimulations. Ideally, the bowel program should be performed at the same time of day to facilitate “retraining” of the bowel. Patients injured for more than 1 year are more likely to perform the bowel program every other day, and in the morning. Persons with tetraplegia report requiring more assistance and more time overall for the program (254).

For patients who remain constipated, laxatives such as lactulose, polyethylene glycol, bisacodyl tablets, milk of magnesia, or cascara may replace, or be used in combination with components of the 3-2-1 program. The use of oral agents should be individualized, with the ultimate goal of minimizing medications as the time from injury increases. Only one change to the bowel program regimen should be at a time, and at least three bowel cycles should be completed to realize the effects of the change. While large volume enemas are occasionally indicated for episodes of constipation, they are not recommended for chronic use. Bisacodyl is an active ingredient in most suppository preparations. Suppositories with a water soluble base (Magic Bullet) have been shown to dissolve more quickly and significantly shorten the time to complete a bowel program compared with standard vegetable oil-based preparations (255). Docusate sodium mini-enemas may allow for the most rapid bowel evacuation (256).

Surgical interventions should be considered if diet modifications, medications, and physical techniques have been attempted but fail to produce a consistent bowel movement. Colostomies and ileostomies decrease the time required for bowel management, increase independence, and simplify the overall process. Many patients have reported that they retrospectively wished that these alternatives were offered earlier (257,258). An antegrade continence enema (ACE) procedure may be an option for adult patients with neurogenic bowel recalcitrant to a bowel program. Originally designed for children with myelomeningocele, it also significantly decreases toileting time and improves QOL (259). Patients also find...
it more cosmetically acceptable than a colostomy. Recently, studies examining a transanal irrigation system have shown improvements in constipation, QOL measures, and fecal incontinence (260).

In LMN lesions, continence is often lost due to weakness of the pelvic floor muscles with a flaccid external anal sphincter. Because spinal-mediated reflexes are absent, bacterial stimulation and contact irritant suppositories are largely ineffective, necessitating manual disimpaction. The stool is kept firm by use of bulking agents to aid manual disimpaction. Performing the disimpaction more than once per day may be required. Assisitive techniques such as valsalva maneuver, abdominal massage in a clockwise direction starting in the right lower quadrant and progressing along the course of the colon, increase in physical activity, standing, and completing the bowel program in a commode chair rather than in bed can also greatly facilitate the process.

**GU SYSTEM**

Prior to the 1970s, renal disease was the leading cause of mortality in chronic SCI. Since the advent of IC programs, the prevalence of renal disease has decreased dramatically. Acutely, patients are usually in spinal shock, with the bladder being areflexic, retaining urine. This can last from a week to 1 year, but in most cases resolves after 6 to 12 weeks. An indwelling catheter is the initial management of choice to remove urine and monitor output. Patients can be started on an IC program, when they can tolerate a fluid restriction of approximately 2 L/d. Once patients are medically stable, a discussion regarding the long-term management options should commence. The goal is to have a “balanced bladder” that includes maintaining low bladder pressures and continence while minimizing infections and the risk of upper tract deterioration (261,262).

As with the prescription of the bowel program, patient preference must be considered. Ideally, the clinician will present all available management strategies, including pros and cons of each, and the ultimate decision will be made by the patient.

For patients with an UMN injury who have adequate hand function, an IC program provides a near physiologic voiding pattern. For patients who have urinary leakage between catheterizations, or high filling pressures documented on urodynamics, anticholinergic medications can be administered. Newer techniques, including botulinum toxin to the detrusor wall, are being used. Volumes of each IC should be less than 500 cc and frequency of caths are most feasible four to five times per day; correlating with a 2 to 3 L fluid restriction. Patients must understand that failure to comply with the catheterization schedule may lead to high bladder volumes and pressures, and ultimately may progress to upper tract deterioration. The benefits of an IC program include improved self-image and conduciveness to sexual activity, and it may result in a decreased incidence of bladder stones, UTIs, and bladder cancer. IC is generally accepted to be the best and safest long-term bladder method.

Chronic indwelling catheters (i.e., urethral and suprapubic) are generally easy for patients and caregivers to manage. Negative consequences associated with indwelling catheters include increased incidence of bladder stones, UTI, urethritis, urethral erosions, prostatitis, strictures, and cancer for long-term (more than 10 years) users. Additionally, the presence of the catheter may hinder attempts at intimacy. However, when other more medically appropriate methods are not functionally feasible, this type of option is preferred (261).

Reflex voiding (with a condom catheter) is an option for men with UMN bladders, but urodynamic monitoring is recommended to assure voiding at acceptable pressures. There are no consistent external catheters for women. Careful monitoring of postvoid residuals is needed to insure adequate emptying as urinary stasis increases occurrence of UTIs. If there is an outlet obstruction, the use of α-blockers, botulinum toxin injection of the external urethral sphincter, and surgical procedures are available. In addition, there are a number of surgical approaches to assist in bladder management for the patient with chronic SCI. These include FES techniques, and approaches for continent and incontinent diversions (261,262).

For patients who have a LMN bladder, techniques such as Crede (direct pressure applied over the lower abdomen) and/or valsalva maneuvers may provide enough pressure to cause expulsion of urine from the bladder. Patients should be cautioned that this method can increase bladder pressures thereby increasing the risk of hydronephrosis. For this reason, these maneuvers are not advised for patients with known dystynergia. IC programs are otherwise the best option.

Patients with neurogenic bladder dysfunction require routine screening for development of stones, fibrosis, hydronephrosis, renal disease, and bladder cancer (261). The presence of bacteria in patients with indwelling catheters or who perform IC is quite common. A UTI is defined as: (a) significant bacteruria, (b) pyuria, (c) clinical signs and symptoms, for example, fever, malaise, increased spasticity, or neurogenic pain (263). Generally, asymptomatic bacteriuria is not treated and the use of prophylactic antibiotics to prevent UTIs after SCI is not supported (264). Exceptions include patient who are scheduled to undergo invasive procedures such as cystoscopy or urodynamic studies, or have urease-producing pathogens (e.g., proteus, pseudomonas, Klebsiella, Providentia, E. coli, Staphylococcus epidermidis), or in the presence of reflux.

Follow-up screenings are recommended, although the frequencies of these are variable across SCI centers. Renal sonograms are generally obtained every 1 to 2 years, or if there is clinical evidence suggestive of nephrolithiasis. Patients with indwelling catheters should have regularly scheduled cystoscopy evaluations at least every 2 to 3 years to screen for bladder cancer. Those at higher risk, that is, smokers, those with a history of abnormal cystoscopy findings, require more frequent examinations. Patients with persistent hematuria or constitutional symptoms consistent with an underlying malignancy should be screened immediately. Urodynamic studies are indicated whenever there is a change in the patient’s voiding pattern.
and for occasional screening to assure appropriate filling and voiding pressures.

The incidence of renal calculi in SCI is approximately 5% to 10% (173). Stones are typically associated with UTIs, especially with bacteria that are urease-producing, since urease alkalinizes the urine and promotes crystallization of struvite and calcium phosphate stones. The most common stones previously reported are infection-related stones with struvite (magnesium ammonium phosphate), followed by calcium oxalate and phosphate stones; although this trend may be changing (265). Risk factors include the presence of uroterovesicular reflux, chronic UTIs, higher level and more complete injury, indwelling catheterization, and bladder stones (266,267). Staghorn calculi are more common with struvite stones and carry a higher morbidity rate.

Clinical presentation of renal stones includes flank pain, nausea, emesis, hematuria, hyperhidrosis, presence of AD, or recurrent UTIs. Workup includes stone analysis once obtained, with serum studies (BUN, creatinine, calcium, and phosphorus); urinalysis (including urine pH); and a 24-hour urine for total volume and measuring calcium, oxalate, citrate, magnesium, uric acid, sodium, potassium, and phosphorus. Imaging such as a KUB x-ray, which can visualize calcium stones (struvite stones are not as dense), ultrasound, computerized axial tomography (CAT) scan, and intravenous pyelography (IVP) may be used. Treatment for kidney stones includes medical stabilization with fluids, treatment of infection if present, and pain management. Surgical treatment for stones includes extracorporeal shock wave lithotripsy (ESWL), percutaneous nephrolithotomy, ureteroscopy, and open nephrolithotomy. Once stones are present, there is a recurrence rate of 15% to 72% over a 5-year-period. (For more information on neurogenic bladder dysfunction, see Chapter 24.)

Osteoporosis
Following injury, there is a marked increase in osteoclastic activity leading to net bone resorption. Bone loss is most active in the first 14 months postinjury, with a slower loss over the next few years, placing patients at an increased risk early after injury of developing pathological fractures (268,269). Bone loss occurs in the LEs at a rate of 1% to 2% per month, with a more rapid loss in the distal femur and proximal tibia as compared with the proximal and midshaft femur. There is a greater loss of trabecular bone as opposed to cortical bone. There is also loss of bone mass in the UE of persons with tetraplegia, with reported normal to an increase in bone mass in the lumbar spine of all patients, possibly due to constant loading while sitting in the wheelchair, and in the forearm of athletes. The increase in the lumbar spine bone mass may just represent a spurious finding rather than being spared bone mineral loss (270). There is no difference in bone loss in the LEs between persons with paraplegia or tetraplegia (269,271). The greatest risk factors for osteoporosis, especially at the knee, include a neurologically complete injury, lower body mass index (BMI), and greater age.

Most patients with chronic SCI have T-scores on DEXA scan of the LEs of greater than 2 standard deviations below the mean, and therefore this value alone cannot determine fracture risk. In addition, the sites of measurement are not always consistent with the sites of fracture most common in SCI. Garland et al. reported that patients who experience pathologic fractures of their LEs have lost 50% of the BMD of their knees (relative to able-bodied group) and that the fracture threshold of the knee is 0.6 g/cm² (272). A study utilizing quantitative CT that calculates volumetric densities reported a fracture threshold of approximately 110 mg/cm³ in the distal femur and 70 mg/cm³ in the distal tibia (273). Below these levels, fracture due to minor trauma is common.

Treatments for bone loss have included weight bearing, FES, and pharmacological intervention (274). Therapeutic interventions including the use of standing tables, ambulation, and FES have produced mixed results (275–278). While there may be a dose-response effect in utilizing FES, the precise requirements needed to obtain clinically useful outcomes are yet to be determined. The use of FES in a person with chronic SCI and low BMD may lead to fracture, and caution should be taken (279). Possible treatments to prevent bone loss following SCI including the use of bisphosphonates, is being researched. Data have shown promising results in decreasing bone loss in the acute injury period; however, sustained benefit postinjury and its effectiveness on long-term fracture rates are unclear (280–282).

Fracture
The historical incidence of long-bone fractures ranges between 2% and 6%, and has been documented at approximately 2.5% over 20 years after SCI in model SCI system data (173). The incidence is most likely much higher when one considers that many individuals with SCI have impaired sensation, and thus may not seek medical treatment. Fractures are more common in women and persons with neurologically complete injuries, and in paraplegia relative to tetraplegia (272,283). Most fractures are due to falls during transfers, but fractures can result from minor stresses (e.g., long-sitting or ROM) or without any known etiology. Supracondylar femur fractures are the most common fractures, followed by the distal tibia, proximal tibia, femoral shaft, femoral neck, and humerus, respectively. A bone mineral density fracture threshold of 50% appears to exist for the knee, and this most likely is the bone mineral density fracture threshold for most regions in the body. Symptoms of acute fracture include fever, acute pain, swelling, or increased spasticity. A plain x-ray will usually give a definitive diagnosis.

For fractures in chronic patients with SCI who do not use their lower limbs for functional mobility, the main goals of treatment are to minimize complications, allow for healing with satisfactory alignment, and to preserve prefracture function. Most fractures are treated nonoperatively with soft padded splints (284,289). A well-padded knee-immobilizer is useful for femoral supracondylar, femoral shaft, andproximal tibial fractures; a well-padded ankle immobilizer can be used for distal tibial fractures. In nonambulatory patients, some degree of shortening and angulation is acceptable. The patient should be allowed to sit within a few days. Callus formation is
usually evident in 3 to 4 weeks; however, ROM is initiated at 6 to 8 weeks, with weight bearing delayed for a longer period. While nonunion of fractures occur (2% to 10%), this is not clinically significant in those who do not weight-bear through their lower limbs.

Surgery, circumferential casting, and external fixation are usually not indicated in the SCI population because of potential complications due to low bone mass, risk of osteomyelitis, recurrent bacteremia, and skin breakdown. Surgery may be indicated when conservative methods will not control rotational deformity, for proximal femur fractures, in patients with severe muscle spasm, if vascular supply is in danger, and in whom shortening and angulation will result in unacceptable function or cosmesis. Femoral neck and subtrochanteric fractures are the most difficult to manage and internal fixation may be considered if a minimal device such as an intramedullary rod can be used.

**Heterotopic Ossification**

HO is the formation of lamellar bone within the soft tissue surrounding a joint. The incidence ranges between 13% and 57%, and is usually found in the first 6 months (peak at 2 months) after injury. HO may occur beyond 1 year and is usually associated with a newly developed PU, DVT, or fracture. Risk factors for HO include older age (children and adolescents have a lower incidence), neurological complete lesions, male gender, spasticity, DVT, and pressure sores (285,286). These risk factors may be cumulative.

Most of the cases of HO will have only radiological findings and are not clinically significant. Up to 20% will present with a limitation of the ROM, with up to 8% progressing to joint ankylosis. Only joints below the NLI will develop heterotopic bone, with the most common location being the hips (anteromedial aspect), followed by the knees, then shoulders. The joint may appear warm and swollen and must therefore be differentiated from a septic joint, cellulitis, DVT, fractures, and inflammatory arthritis. The patient may experience pain, malaise, low-grade fever, and an increase in spasticity. In severe cases, adjacent neurovascular structures may be compromised leading to distal extremity swelling and nerve entrapment (287).

The pathogenesis of HO is not fully understood, but involves changes of soft tissue mesenchymal cells to osteogenic precursor cells by humeral, neural, and local factors. Bone morphogenic protein is thought to play a key role (288). Damage to the sympathetic tracts in the spinal cord may also promote HO by increasing the local vascularity and blood perfusion around the joint. When passive ROM is delayed more than 1 week after injury, patients are more likely to develop HO (289). It is also possible that forced ROM of contracted limbs may cause micro trauma and bleeding, leading to HO (290).

Laboratory tests are sensitive but are not specific markers for HO. Serum alkaline phosphatase levels start to increase prior to clinical and radiographic presentation, but may not exceed normal levels for several weeks. Because levels do not correlate with the amount or degree of bone activity, they should not be used to judge the maturity of the new bone or predict its recurrence. An elevation of serum creatine phosphokinase (CPK) may be a more reliable predictor of HO (291,292). While nonspecific markers of inflammation such as C-reactive protein (CRP) and the erythrocyte sedimentation rate (ESR) can be useful in following disease activity (293), any cause of systemic inflammation (e.g., infection, PUs) may result in these elevations. CRP is a more reliable predictor of disease activity, with normalization of the CRP correlating with resolution of the inflammatory phase of HO. Urinary excretion of hydroxyproline and collagen metabolites correlates with alkaline phosphatase levels and can also serve as indirect markers for the presence of HO.

A triple phase bone scan is the most sensitive imaging study in diagnosing early HO, and can detect disease activity before calcification becomes apparent on plain x-ray. The first two phases of the bone scan measures the increase in blood flow to a joint during the early inflammatory period. The third phase, or static bone phase, is more specific since it measures the incorporation of the radionuclide into the bony matrix but may take another 3 weeks before it is positive (294). Bone scans are also the most useful technique to assess maturity of the heterotopic bone. Plain x-rays become positive approximately 2 to 6 weeks after a triple phase bone scan first reveals HO or 1 to 10 weeks after clinical presentation (295). Ultrasonography may be positive early and has the advantage of being a relatively inexpensive examination without requiring radiation (296). MRI, with an increased T2 signal (edema) in muscles, fascia, and subcutaneous tissue can be helpful in diagnosing HO acutely (297). CT scan may be used to determine the volume of bone needed for planning surgical resection. Multiple grading systems for classifying HO have been proposed (285).

Treatments include ROM with gentle stretching after the acute inflammatory period is over (1 to 2 weeks), bisphosphonates, NSAIDS (e.g., indomethacin) if not contraindicated, radiation therapy, and surgical excision (298). Passive and active assistive ROM to the affected limb is necessary to prevent further loss of ROM. Aggressive ROM (beyond the initial end-point), especially during the acute inflammatory phase, is not recommended (299). Treatment with bisphosphonates has been shown to decrease the rate of new bone formation in patients with HO; however, it has no effect on bone which has already been deposited. Disodium etidronate inhibits osteoclastic activity and conversion of calcium phosphate to hydroxyapatite. Although IV administration of etidronate reportedly led to quicker resolution of edema with less rebound formation after the medication was discontinued, it is no longer available. Current recommendation is for oral administration of etidronate 20 mg/kg/d for 6 months if the CPK level is elevated at the time of diagnosis or 20 mg/kg/d for 3 months, followed by 10 mg/kg/d for an additional 3 months if the initial CPK level is normal (300). With this regimen, there was faster resolution of edema with less rebound formation after the medication was discontinued. If CPK is elevated, or CRP greater than 8, some recommend addition of a NSAID until
the CRP less than 2 or CPK normalizes. Clinical trials with newer generation bisphosphonates are ongoing. Surgical excision should be reserved for patients with severely limited ROM that causes functional limitations. Most clinicians recommend waiting until after the ectopic bone is mature by bone scan, which may take up to 12 to 18 months to occur. Complications of surgery include significant blood loss, infection, and recurrent HO. Postop treatment includes NSAIDS for at least 6 weeks, bisphosphonates for 3 to 12 months, and/or radiation. While radiation decreases the degree of recurrence of HO, complications include delayed wound healing, osteonecrosis, and the risk of developing sarcoma (295,301).

Prophylaxis of HO in SCI patients has been studied using several agents including etidronate and indomethacin (75 mg daily for 3 weeks within 5 weeks of injury), with less HO formation as compared with placebo (302,303). Warfarin may also be an effective agent by inhibiting the formation of osteocalcin (304). Despite the available therapeutic options, prophylaxis is still not routinely used because of the relatively low incidence of morbidity (286,295,305).

Thermoregulation
When the core temperature is cool and requires an adjustment, the hypothalamus, which regulates body temperature, employs shivering and vasoconstriction to increase temperature. Similarly, sweating and vasodilatation decreases temperature through increased heat loss. After a SCI above the T6 level, the ability of the hypothalamus to direct the periphery is impaired. This results in the patients being partially poikiloithermic, in that they may have difficulty maintaining a normal core temperature in response to environmental change in temperature (306,307). This is an important factor to keep in mind when a person with SCI has a high temperature while wearing appropriate clothing depending on the setting/environment they are in. Patients with chronic tetraplegia also frequently have subnormal body temperatures (<97.7°F) in a normal ambient environment (308).

Anemia
Anemia is a common finding following acute SCI, and is usually normochromic and normocytic (309). Serum iron, TIBC, and transferrin are usually low. Although the exact cause is not known, bleeding may be a factor in some cases. By 1 year postinjury, anemia improves in the majority of patients and if it persists it is usually associated with chronic inflammatory complications such as PUs or frequent UTIs (310).

Pressure Ulcers
PUs are one of the most common and potentially serious complications of SCI (see Chapter 23). Model system data reported that since 1996 approximately 24% of patients developed at least 1 PU during their initial rehabilitation (172). This is a decrease from previous studies that reported a prevalence of approximately one third of patients who developed a PU during their initial hospitalization and rehabilitation (172, 311–313), with a lower incidence for persons admitted immediately to a model SCI center. The longer time injured, the greater the risk of developing an ulcer, as up to 80% of persons with SCI develop a PU at some point in their lifetime, with a community prevalence of approximately one third (175,314–316). Skin issues are the second most common etiology for rehospitalization in chronic SCI (327). Persons with AIS A, B, or C paraplegia are more likely to be rehospitalized with skin issues than those with any level of tetraplegia or AIS D paraplegia.

PUs are classified using the National Pressure Ulcer Advisory Panel (NPUAP) staging system, with new revisions as of 2007 (317) (Table 27-10). The NPUAP includes the original four stages (318) and additional two categories on deep tissue injury and unstageable PUs. Staging should only be used when the wound bed can be visualized. Deep tissue injury refers to a purple or maroon localized area of discolored intact skin or blood-filled blister due to damage of underlying soft tissue from pressure and/or shear. Unstageable wounds occur when not enough slough and/or eschar is removed to expose the base of the wound, and therefore the true depth cannot be determined and therefore staged.

PUs develop over bony prominences. The most common location in persons with SCI within the first 2 years is the sacrum, followed by the ischium, heels, and trochanters. After 2 years, the ITIs are the most common site of development. In children up to age 13 years, the occiput is the site of most frequent development.

The most important risk factors for the development of PUs are pressure and shear. Other risk factors significant in SCI include level and severity of the injury, mobility status, gender, ethnicity, marital status, employment status, educational achievement, tobacco and alcohol use, history of a prior PU, nutritional status, anemia, incontinence, smoking, and possibly psychosocial issues (i.e., depression).

Secondary effects of PUs include further medical morbidity, pain (especially in incomplete lesions), increased spasticity, and wound infections. In addition, psychological complications from prolonged bed rest (often required for management of the sore) often occur, along with a diminished QOL secondary to

<table>
<thead>
<tr>
<th>Stage</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Non-blanchable erythema of a localized area usually over a bony prominence.</td>
</tr>
<tr>
<td>II</td>
<td>Partial-thickness loss of dermis or dry shallow ulcer.</td>
</tr>
<tr>
<td>III</td>
<td>Full-thickness destruction through dermis into subcutaneous tissue. Subcutaneous fat may be visible but bone, tendon, or muscle is not exposed. May include undermining and tunneling.</td>
</tr>
<tr>
<td>IV</td>
<td>Full-thickness tissue loss with exposed bone, tendon, or muscle.</td>
</tr>
</tbody>
</table>


### TABLE 27.10 NPUAP Staging of PUs

<table>
<thead>
<tr>
<th>Stage</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Non-blanchable erythema of a localized area usually over a bony prominence.</td>
</tr>
<tr>
<td>II</td>
<td>Partial-thickness loss of dermis or dry shallow ulcer.</td>
</tr>
<tr>
<td>III</td>
<td>Full-thickness destruction through dermis into subcutaneous tissue. Subcutaneous fat may be visible but bone, tendon, or muscle is not exposed. May include undermining and tunneling.</td>
</tr>
<tr>
<td>IV</td>
<td>Full-thickness tissue loss with exposed bone, tendon, or muscle.</td>
</tr>
</tbody>
</table>

Deep tissue injury.
Unstageable wounds.

decreased recreational and vocational activities and a significant burden financially. Costs can be staggering in terms of lost workdays, increased attendant or skilled care, specialized bedding, supplies, hospitalization, and surgery (319). Previous estimates of costs (1990s costs) range from $20 to 30,000 for less serious ulcers to $70,000 to heal a full thickness ulcer (313). These do not take into account all of the hospital and postoperative costs as well as the indirect costs of the ulcer.

PU prevention is critical, and education regarding this should begin as early as possible after injury. This includes continued observation of skin (teaching patients to use a mirror), avoidance of excessive pressure or shearing, proper turning frequency in bed and weight relief techniques when seated, proper equipment (e.g., mattress, wheelchairs, cushions, and seating system), and early recognition and treatment. In addition, maintaining proper nutrition and discontinuing smoking are recommendations that will help prevent and heal ulcers if they develop.

The key to treatment of PUs is to relieve the pressure. Pressure mapping can be helpful in establishing an appropriate seating system for the patient to assure that appropriate pressure distribution occurs while sitting, and that appropriate weight relief is produced during the pressure relief. There are a great deal of dressings, gels, and ointments available on the market, but the most important aspect is to remove the etiology (e.g., sitting for prolonged periods in cases of an ischial wound). Wound management is discussed in Chapter 23. A VAC® system, a controlled application of sub-atmospheric pressure to the wound, can be used to keep the wound clean and help promote wound healing (313). Surgical consultation should be considered in most cases of stage 3 and 4 wounds for potential surgical intervention. There are a number of surgical options available (see Chapter 23 for details) but it should be recognized that surgery does not treat the underlying cause of the ulcer. Postoperatively, this still needs to be addressed. An appropriate sitting program should be followed after these procedures.

**Dual Diagnosis: TBI in Person with SCI**

The incidence of concomitant TBI in those with a primary SCI is reported between 24% and 60% (320,321). Historical factors such as mechanism of injury (i.e., high velocity impact), loss of consciousness, prolonged extrication and/or intubation at the scene, higher NLI, presence of PTA, and impaired initial Glasgow Coma Scale (GCS) score, should alert medical personnel of the possibility of a concomitant TBI.

The individual with a dual diagnosis presents a challenge to the rehabilitation team (322). Deficits may be seen in attention, concentration, and memory as well as interfere with new learning and problem solving. The patient may exhibit agitation, aggression, disinhibition, and depression. This is problematic as SCI rehabilitation requires intensive new learning, with the ability to master new skills in terms of mobility and self-care, and adapt to a new lifestyle to be able to integrate into the community. Recognizing that the TBI exists and then developing a comprehensive treatment program to deal with the dual injury is necessary in order for the individual to meet their maximum potential. Before a patient is diagnosed as having a dual disability based on behavioral issues, the clinician should seek other causes for such behavior including the possibility of seizures, post traumatic hydrocephalus, intracranial pathology, neuroendocrine disturbance, metabolic disturbance, side effects of centrally acting medications, hypoxemia, and infectious issues.

Medical management of problems common to the SCI population such as pain, DVT prophylaxis, spasticity, and neurogenic bladder requires special consideration in the dual diagnosis patient. In addition, the physician needs to take into account some medications routinely used in SCI that may have an impact on the recovering brain (i.e., baclofen and benzodiazepines). Special care should be taken to utilize medications with minimal cognitive impact. Peripherally acting medicines are generally preferred to those with central mechanisms of action. Clinical findings in SCI individuals may have different etiologies and as such require different treatments in the patient with dual diagnosis. Fever may represent infection, poikilothermia, or “central fever” which occurs in TBI. Once the former two etiologies have been excluded, treatment with a β-blocker can be considered. Elevated BP may signal the presence of AD in those with SCI at or above the T6 level. However, those with concomitant TBI may experience transient elevation of BP secondary to a centrally driven “sympathetic storm.” If initial interventions to treat AD do not improve the patient’s BP, pharmacotherapy should be initiated. Agitation is common in those with TBI and should be managed similarly in those with dual diagnosis. Behavioral interventions such as redirection and decreasing environmental stimulation are first-line treatments. Pharmacotherapy is appropriate when behavioral modifications are ineffective.

In those with TBI, acquisition of new information is often challenging, and may benefit from simple task repetition in therapy until individual skills can be completed with minimal prompting from the therapist. The therapeutic environment itself may also differ, as persons with TBI may perform better in less stimulating environments (i.e., with fewer visual and auditory distractions, with close supervision by the treating therapist). Cognitive therapy strategies should be reinforced by all members of the treating team such as reminding patients to utilize memory books and orientation cues with each clinical encounter. Neuropsychological evaluation is extremely important.

While few studies have examined the effects of TBI on outcome among individuals with SCI, greater adjustment difficulties and unmet needs were noted among persons with dual diagnosis (323–325). Smaller functional gains (as measured by the motor FIM) in those with dual diagnosis have been reported, however, no significant difference in rehabilitation LOS has been observed (326). Medical morbidities may be increased. For example, there could be an increased risk of PU due to the inability of the patient to remember to perform a pressure relief, and bladder and bowel difficulties due to the inability to perform these programs adequately.

**Spasticity**

The incidence of spasticity in UMN-related SCI is approximately 70%, with roughly half the patients requiring...
pharmacological intervention (327,328). While spasticity may occasionally contribute to improved function (i.e., transfers, standing, ambulation, and assisting in ADL), it more often leads to various complications including contractures, pain, impaired function, and decreased QOL. Spasticity occurs more frequently in persons with cervical and upper thoracic SCI than in those with lower thoracic and lumbosacral SCI, and is usually more significant in persons with certain incomplete injuries, AIS grades B and C, than in persons with grades A or D. Details regarding the basic definitions of spasticity, diagnostic procedures, and treatment options, are in Chapter 60.

Clinical examination is important, however, the degree of spasticity may vary during the course of the day. Therefore, the patient’s description of secondary problems from spasticity becomes extremely important in determining whether to treat. The most common clinical scales used to assess spasticity in SCI include the Ashworth scale (AS) (329), modified Ashworth scale (MAS) (330), the Penn spasm frequency scale (PSFS) (331), Spinal Cord Assessment Tool for Spasticity (SCATS) (332), visual analog scale (VAS), self-rated scale of spasticity (333), and the Wartenberg Pendulum Test (334). Different scales measure different aspects of spasticity and individual tools correlate weakly with each other (335).

Before initiating treatment, potential nociceptive sources should be evaluated and if present, they should be treated. Common causes include UTI, bladder calculi, PUs, abdominal pathology, ingrown toenails, hemorrhoids, and bowel impaction. Selective serotonin reuptake inhibitors (SSRIs) have been reported to increase spasticity (336). Changes in spasticity in an otherwise stable patient may also be an important presenting sign in syringomyelia (337). Spasticity-related interventions should be aimed at what matters most to the patient, improving comfort and function, and allowing the individual to participate in life activities.

Stretching is a mainstay of treatment for SCI-related spasticity (338) and should be done twice a day or more. Standing activities, including use of tilt table or standing frame, provide prolonged stretch to joints and may reduce spasticity as well. Modalities have been used including cold and ES, usually with short-term benefits. Posture and positioning are extremely important for tone reduction. Adequate low back support in the wheelchair to maintain lumbar lordosis and a positive seat plane angle or “dump,” with a reduction in seat-to-back angle, encourages proper upright posture and may reduce extensor tone. Inhibitive casting, splinting, and orthotic management of muscles and joints at risk for contracture may be helpful in reducing tone (338,339).

There are numerous medications for persons with SCI-related spasticity. A systematic review indicated that there is insufficient evidence to assist clinicians in a rational approach to antispasticity treatment in SCI (340), and therefore some amount of trial-and-error is required. The person’s age, comorbidities, and cognitive status should be carefully considered when choosing a medication for spasticity.

Baclofen exerts GABA agonist activity by binding presynaptically at the GABA B receptor. The dosage of baclofen used in SCI is often higher than that described in the Physician Desk Reference (PDR) (341). If baclofen is to be discontinued, it should be tapered as abrupt discontinuation can result in seizures, visual disturbances, and hallucinations. Benzodiazepines may affect cognitive performance measures such as attention, concentration, and memory, and are generally not used in persons with concomitant brain injury. A single night-time dose of diazepam or clonazepam is often used to treat nocturnal spasms that interfere with sleep. Alpha 2 adrenergic agonists, including clonidine and tizanidine, bind to presynaptic α-2 receptors on interneurons in the dorsal horn of the spinal cord, resulting in depression of polysynaptic reflexes. Clonidine can be delivered either orally, via transdermal patch, or intrathecally (342,343). Tizanidine has a shorter half-life and lower incidence of hypotension than clonidine, and has been shown to be effective in patients with SCI and MS (344,345).

Gabapentin (especially at higher doses of >1,800 mg/d) and pregabalin, both anticonvulsants, have antispasticity benefits in persons with SCI (346–348). Dantrolene sodium prevents the release of calcium ions from the sarcoplasmic reticulum along muscle fibers, and acts on normal as well as spastic muscle, potentially resulting in weakness. Dantrolene may be of special benefit in persons with concurrent SCI and TBI, since its effects are peripheral rather than in the central nervous system (349). The most worrisome side effect is hepatotoxicity that occurs in approximately 1.8% of patients. Cyproheptadine is a nonselective serotonin (5-HT) antagonist, which has antihistamine quality, and reduces clonus and spontaneous spasms in patients with SCI either alone (350) or in combination with clonidine (351). Most common side effects include fatigue and excessive weight gain. Opiates exhibit potent antispasticity activity (352) by suppressing polysynaptic reflexes to a greater extent than monosynaptic reflexes, but are not considered a primary treatment option for SCI-related spasticity. Four-aminopyridine (4-AP), a potassium channel blocking agent has shown some effect on decreasing spasticity and improving function in persons with incomplete SCI (353). Although not approved for clinical use, cannabis has been beneficial for persons with SCI-related spasticity (354).

Intrathecal baclofen (ITB) is of benefit for persons with SCI, with effective control of spasticity in the lower limbs being greater than in the trunk and upper limbs (355). Intrathecal clonidine may also be effective either alone (356) or in combination with baclofen (357), although this is not FDA approved. Careful patient selection is critical with close monitoring for pump failure or other complications.

Peripheral nerve and motor point blocks with alcohol or phenol allow for a focused delivery of drug to manage focal areas of spasticity. These drugs can have effects of up to 1 year by causing destruction of the axons. Botulinum toxin, which can be injected into the affected muscle, blocks neuromuscular transmission by inhibiting the release of acetylcholine into the synapse. It has an onset of action of 4 to 10 days, the effect lasts on an average for 3 months, and it has shown efficacy in improving UE and LE functions (358,359). Follow-up stretching, which may include dynamic splinting or serial casting, is recommended to optimize function.
Surgical interventions, including cordectomy and myelotomy, have been performed to reduce spasticity, but with limited long-term success (360). More selective procedures, dorsal rhizotomy, and dorsal root entry zone (DREZ) procedures, have been proposed to control spasticity by decreasing the afferent component of spasticity, although these should only be considered if all previous treatments have failed.

**PAIN AFTER SCI**

Pain after SCI is common enough to be considered an expected condition (361). In the acute phase, pain is usually related to damage to the soft tissue and skeletal system from the initial trauma. After the injuries are appropriately treated, this type of pain usually subsides. However, between 47% and 96% of patients will develop pain thereafter (362–366). Model system data indicate that pain prevalence is up to 81% at 1 year after injury and 82.7% at 25 years. Pain is rated as severe in 20% to 33% of persons with pain after SCI. The presence of chronic pain is a significant contributor to decreased QOL, lower levels of psychological functioning and social integration, and an increased interference with ADL (367). The impact of pain on QOL may be greater than that of the injury itself (368).

Despite numerous classifications proposed, there is no single widely accepted scheme for SCI pain. Three main classifications include the International Association for the Study of Pain (IASP) (369), Bryce-Ragnarsson (370), and Cardenas classifications (371). Most share two basic categories of pain after SCI: musculoskeletal/nociceptive and neuropathic/neurologic. As a major category, musculoskeletal pain is more common, occurring in 50% to 60% with SCI greater than 5 years, while neuropathic being prevalent in 30% to 40% of persons 5 or more years after SCI. Although the classifications are similar in some ways there are differences in the organization, level of detail, and terminology used for particular types of pain. As none have achieved universal acceptance, a joint initiative between ASIA and ISCOS is developing a universally accepted assessment tool for pain in SCI.

Musculoskeletal pain refers to the pain originating from damage to tissue and bone structures and may include overuse syndromes, fractures, and compression syndromes. Neuropathic pain (also called neurologic or central pain) is directly attributable to spinal cord damage and can be divided into multiple subtypes by the level of injury (at the zone of injury or below) and etiology. It is often described as burning, tingling, cramping, shock-like, or a sensation of intolerable coldness. Neuropathic pain is more refractory to treatment than musculoskeletal pain, can occur at any time following injury, even years following their initial injury, with a low frequency of spontaneous recovery.

There are several proposed mechanisms for the origin of neuropathic pain after SCI. Pain may arise from a combination of generators: peripheral, spinal, and supraspinal (372). Peripheral sources may include impingement of nerve roots, resulting in radicular-at-level neuropathic pain. Spinal pain may be from an “irritated focus” or “neural pain generator” located at the injury site, as there are cases of spinal blockade with anesthetics abolishing pain. In some cases spinal anesthetic blockade does not always relieve pain and thereby supraspinal pain may be from the presence of thalamic neurons with abnormal patterns of activity. GSWs, neurologically incomplete, and CE injuries may be predisposing factors that increase the risk of developing neuropathic pain. Psychosocial deficiencies may predict subjective pain complaints. Individuals who are depressed, anxious, or experience severe stress are more likely to have pain (373), while those who are active and working are less prone to complain of pain.

Unfortunately, although many treatments for neuropathic pain have been tried, few have been found to be effective in controlling SCI neuropathic pain. The focus should be symptomatic relief and helping the patient manage with the pain rather than complete elimination of the pain. Treatments used include medications, such as anticonvulsants, antidepressants, local anesthetics, and NMDA antagonists and invasive procedures such as intrathecal drug administration and surgery. Satisfactory relief, defined as a 50% reduction in pain, is at best found in one third of subjects in controlled trials.

Anticonvulsants, most notably gabapentin and pregabalin, are increasingly being used as first-line agents and have shown effectiveness in patients with SCI-related neuropathic pain (374,375). Carbamazepine and oxcarbamazepine have been used although controlled trials have not been performed (372). Tricyclic antidepressants (TCAs) have been commonly prescribed. A recent randomized trial found amitriptyline (maximum dose of 50 mg TID) to be more effective than gabapentin (maximum dosage of 1,200 mg TID) (376). TCAs combined with anticonvulsants may be more effective than either drug alone. TCAs are contraindicated in patients with ischemic heart disease, heart failure, cardiac conduction disturbances, and a history of seizures. Sedation and anticholinergic effects such as dry mouth and constipation are of concern in the population with SCI. SSRIs are effective antidepressants and have a better side effect profile, but have not shown improvements in pain. Randomized controlled studies involving trazodone, mexilitine, and valproate have not shown a consistent benefit. Opiates are effective, but not without controversy, especially in SCI where side effects such as constipation are notable (377).

The efficacy of intrathecal medications are confined to case series and should only be considered if conservative measures have failed. Morphine and clonidine have shown benefit and the combination of the two (ITB) can also be effective if pain is spasticity-related.

Surgical treatments including spine fusions or nerve decompression are useful for treating pain originating from spine instability or syringomyelia. Ablative surgeries such as cordotomy, cordectomy, and myelotomy are controversial. Several studies have not found any advantage to spinal cord stimulation (378,379). DREZ lesions are thought to be indicated for some cases of neuropathic pain at the level of injury, while those with below level neuropathic pain had poor results (380). ES of the brain has shown some positive impact on pain in SCI (381).
Nonpharmacological and nontraditional approaches to neuropathic pain in SCI are frequently used. These include massage, marijuana, acupuncture, chiropractics, biofeedback, magnets, and hypnosis. Among these, massage reportedly gave the most relief followed by marijuana in a sample of persons with SCI (363). Exercise may increase circulating beta-endorphin levels, a substance known to be associated with analgesia. Psychological counseling reportedly provided years of relief for 20% of the sample (363).

**UE NeuroMusculoskeletal Pain**

The upper limbs in persons with SCI are used for weight-bearing activities including weight shifts, transfers, and wheelchair propulsion as well as for ADL, thereby increasing the chances of overuse syndromes. Shoulder pain is the most commonly reported painful joint after SCI (382). Approximately 30% to 50% of patients complain of shoulder pain severe enough to interfere with function, with the prevalence increasing with time from injury (366,383). Pain during the first year after injury is more common in tetraplegics, while in later years it is more common in paraplegics. Since the person with SCI relies extensively on their upper limbs, any further loss of upper limb function because of pain may have adverse effects on functional independence (384,385).

Two thirds of shoulder pain is due to chronic impingement syndrome and approximately half involves rotator cuff pathology. Bicipital tendonitis, subacromial bursitis, adhesive capsulitis, acromioclavicular osteoarthritis, and cervical radiculopathy are other common causes of shoulder pain in chronic SCI. Other causes that are specific to SCI include muscle imbalance, spasticity, contractures, HO, and the presence of a syrinx. If pain develops acutely, then referred pain should be excluded, including UTI, peptic ulcer disease, angina, and an acute abdominal problem. Pain associated with neurologic changes (i.e., weakness, sensory loss, and/or reflex changes) may be due to peripheral nerve entrapment, radiculopathy, or a post-traumatic syrinx.

Range of motion, strength and sensory testing, provocative tests for specific abnormalities (e.g., Tinel, Phalen, Adson), and a functional assessment should be performed including an evaluation of the patient’s posture, function (i.e., pressure relief, wheelchair mobility, ADL, transfers), and home/work environment. Diagnostic tests should be ordered as needed. The Wheelchair User’s Shoulder Pain Index (WUSPI) is a 15-item self-report instrument, with a range of 0 to 15 that measures shoulder pain intensity during various ADL utilizing a visual analog scale (385,386). This is often used to monitor the effectiveness of treatment.

Relief of acute pain includes rest, icing, and other modalities along with pharmacologic intervention (NSAIDs, acetaminophen, muscle relaxants) and injections, as needed. Acupuncture may also be helpful (387,388). Once overuse problems arise, they can become a chronic problem if treatment is not focused on correcting the cause. Complete rest of a painful shoulder is difficult but decreasing stress on the joints should be encouraged by utilizing compensatory techniques (e.g., using a transfer board rather than performing lateral transfers, or performing lateral or anterior weight shifts rather than push-up weight shifts).

After the pain has subsided, emphasis should be focused toward preventing further injury by obtaining a balance in strength and flexibility of the musculature surrounding the shoulder (389,390). The most often described muscle imbalance is the tightening of anterior muscles with development of proper strength in posterior stabilizing muscles (366). A proper wheelchair and back support will also play a role in preventing shoulder pain. Surgery is not typically performed for patients with rotator cuff tears, as patients require prolonged immobilization postoperatively, creating a period of extreme disability and loss of independent function (366,391). However, there may be some benefit for pain and a small study reported improved strength and ROM with surgery, and therefore can be considered for those who fail conservative management, who are compliant, and in whom postoperative immobilization is not a limiting factor (392).

Overuse injuries may also affect the elbow (32%), wrist, and hand (45%), including the development of carpal tunnel syndrome (CTS), ulnar nerve entrapment at the elbow and wrist, de Quervain syndrome or other tenosynovitis, osteoarthritis, and stress fractures (384). The most common source of elbow pain is overuse of the extensor and flexor tendons as they attach to the medial epicondyle. Ulnar neuropathy can lead to intrinsic hand weakness and medial hand numbness, and can impact ADL (366). The incidence of CTS is between 21% and 65%, with persons with paraplegia more affected than persons with tetraplegia (380), and with a higher incidence at longer times postinjury (393). This is due to recurrent stress from transfers, wheelchair propulsion, and pressure reliefs. Treatment for CTS includes analgesics, NSAIDs, splinting (especially at night), injections (anesthetic and/or corticosteroid), physical modalities (ultrasound, friction massage, etc.), and education regarding transfers to avoid end-range stress. The use of padded gloves may decrease the trauma of wheelchair propulsion. Surgical release may be required, with the postoperative recovery time being weighed against the long-term benefits of the procedure.

Since much of the chronic upper limb musculoskeletal problems are due to the overuse from functional activities, it is vital that proper exercise techniques and shoulder protection, and conservation should be learned early during acute rehabilitation and reinforced frequently. Often one hears the phrase “no pain, no gain”; rather the SCI team should teach their patients to “conserve it, to preserve it” (394).

Spine pain is not uncommon. In patients with incomplete SCI, mechanical pain of the spine may be the result of muscle imbalances, however, the source in persons with complete SCI are less well understood. Considerations should include spasticity, vertebral osteomyelitis, and Charcot spondyloarthropathy (395). A comprehensive workup should be initiated.

**Post-Traumatic Syringomyelia**

The most common cause of progressive myelopathy after a SCI is post-traumatic syringomyelia (PTS). PTS may develop at any
time, from 2 months to decades postinjury (337,396–399). PTS presents as neurologic decline in up to 8% of patients, but is more frequently first seen on MRI as an elongated cavity in a much higher percentage of cases. The pathogenesis is unknown, but the cavity begins at the level of the cord injury in the gray matter between the dorsal horns and posterior columns. Initiating factors include cord hematoma, residual spinal canal stenosis or compression, and spinal kyphotic deformity. The cyst may extend rostrally and/or caudally, compressing the cord, by dissecting through the intermediate gray matter, and may result from increases in subarachnoid fluid pressure due to intra-abdominal or intra-thoracic pressure increases (i.e., cough, sneeze, straining or valsava, weight-lifting, forward-lean pressure release, and quad coughing).

Early signs and symptoms of PTS are often nonspecific and variable; some patients with a markedly elongated syrinx may have minimal symptoms. The most common presenting symptom is pain, usually located at the site of the original injury or may radiate to the neck or upper limbs. The pain is described as aching or burning, often worse with coughing, sneezing, straining, and in the sitting rather than in the supine position. The earliest sign is an ascending loss of deep tendon reflexes. An ascending sensory level is common, typically with a dissociated sensory loss with impaired pain and temperature sensation but intact touch, position, and vibration sense. Loss of pain sensation can lead to a Charcot joint. Weakness occurs, but rarely in isolation. Additional findings may include increased or decreased spasticity, hyperhidrosis, AD, neck muscle fatigue with prolonged sitting, loss of reflex bladder emptying, worsening OH, scoliosis, central and/or obstructive sleep apnea, new Horner’s syndrome, reduced respiratory drive, diaphragmatic paralysis, cranial nerve dysfunction, impaired vagal cardiovascular reflexes, and sudden death (337,396,398,400).

MRI with gadolinium is the gold standard for diagnosing PTS. MRI characteristics that are often associated with neurologic decline include those that are longer and wider, those with poorly demarcated T2-weighted signal hyperintensity at the rostral extent, a flow void sign on T2-weighted images suggesting high pressure, and those associated with spinal stenosis (399). If an MRI cannot be performed, a CT myelogram with delayed imaging can be utilized. Electrodiagnostic findings are occasionally used to aid in the diagnosis (396,401,402).

A syrinx may spontaneously resolve, progress, and then plateau or progress continuously. Neurologic monitoring is essential, and includes clinical examination, serial electrophysiologic tests, and/or MRI. Conservative treatment includes close monitoring, pain control, activity restrictions, and providing rehabilitation interventions as needed (i.e., functional training and adaptive equipment). Activity restrictions include avoiding maneuvers that increase intra-thoracic/abdominal pressure, that is, high-force exercise; valsava, Crede, and quad coughing with direct compression over the IVC; avoiding anterior weight reliefs and weight lifting, especially if these activities exacerbate symptoms. Surgical treatment is usually indicated if there is an ongoing neurologic decline or severe intractable pain. Surgical treatments include shunting (syringo-subarachnoid, syringo-pleural, or syringo-peritoneal), reconstructing the subarachnoid space with dissection of arachnoiditis/meningeal scarring and duraplasty, and cordectomy (403–405). Surgery yields improved strength and improved pain control in some but not all, whereas sensory recovery is not usually as favorable. Reduction of syrinx size on postoperative MRI usually predicts a good surgical result, however, complete resolution of the syrinx is not necessary for a good clinical outcome. Recurrence of neurologic symptoms is common (up to 50%).

### Surgical Interventions of the UE in Tetraplegia

Surgical interventions can improve the functional mobility in appropriately selected persons with tetraplegia. Tendon transfers, arthrodesis of the interphalangeal joints of the thumb, and implantation of an UE neuroprosthesis system are procedures that can be utilized alone, or more commonly, in combination, to improve motor function by one level (406). As a general rule, delay of surgery for 1 year allows adequate time for neurologic stabilization. Components of the preoperative evaluation include measurements of UE strength, sensation, ROM, and spasticity. Vocational and recreational activities should also be noted (407). Discussing the anticipated outcomes and the postoperative recovery plan prior to the procedure allows patients and clinicians to establish realistic expectations.

The International Classification of the upper limb in tetraplegia has been utilized to standardize communication regarding tendon transfer surgery (264) (Table 27-11). In

<table>
<thead>
<tr>
<th>Motor Group</th>
<th>Functional Musclesa</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Weak or absent BR (grade 3 or less)</td>
</tr>
<tr>
<td>1</td>
<td>BR</td>
</tr>
<tr>
<td>2</td>
<td>BR, ECRL</td>
</tr>
<tr>
<td>3</td>
<td>BR, ECRL, ECRB</td>
</tr>
<tr>
<td>4</td>
<td>BR, ECR, ECRB, PT</td>
</tr>
<tr>
<td>5</td>
<td>BR, ECR, ECRB, PT, FCR</td>
</tr>
<tr>
<td>6</td>
<td>BR, ECR, ECRB, PT, FCR, finger extensors</td>
</tr>
<tr>
<td>7</td>
<td>BR, ECR, ECRB, PT, FCR, finger extensors, thumb extensors</td>
</tr>
<tr>
<td>8</td>
<td>BR, ECR, ECRB, PT, FCR, finger extensors, thumb extensors, finger flexors</td>
</tr>
<tr>
<td>9</td>
<td>Lacks intrinsics only</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sensory</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Two-point discrimination in thumb &gt;10 mm</td>
</tr>
<tr>
<td>Cu</td>
<td>Two-point discrimination in thumb &lt;10 mm</td>
</tr>
</tbody>
</table>

*aFunctional muscle: Grade 4 or 5.
BR, brachioradialis; ECRL, extensor carpi radialis longus; ECRB, extensor carpi radialis brevis; PT, pronator teres; FCR, flexor carpi radialis.

contrast to the ASIA classification, functional muscle strength is considered grade 4 or 5 rather than 3 or greater (408). This is because muscles used for transfer generally lose one strength grade and therefore, those with a grade of 3 or less prior to surgery may not have sufficient strength for functional tasks following the transfer. Spasticity in the proposed transferred muscle should be minimal or absent. Shoulder ROM and strength should also be quantified as both are needed for ultimate placement of the hand. ROM of the wrist and fingers should be full for all procedures involving the hand, and for the posterior deltoid or biceps to triceps transfer there should be functional ROM at the shoulder and no more than a 30-degree elbow flexion contracture. The sensory evaluation consists of the Weber two-point discrimination test at the pulp of the thumb (409). Cutaneous sensation is considered present if the patient can distinguish objects less than 10 mm apart and a grade of ocular-cutaneous (O-Cu) is given as this is predictive of sufficient proprioception to allow for hand function without visual cues, as this will allow for ease with ADL. Those who are unable to distinguish objects less than 10 mm are placed in the O-group. If cutaneous afferents are present, both extremities may undergo surgical reconstruction, but it is advisable to perform surgery on only one extremity at a time.

Behavioral evaluation of potential surgical candidates is extremely important. Because the recovery phase will involve a period of immobilization, activity restriction and rehabilitation, patients must be able to adhere to all phases of the recovery plan. During this process, patients will require more assistance from caregivers, modification of work/leisure activities, and adaptive strategies for ADL/mobility. Each surgery is limited to one extremity, with the stronger extremity being operated first. Facilitating hand opening and closing via tendon transfer may require two separate procedures.

The most common procedures for persons with a C5 level include the transfer of the BR to the extensor carpi radialis brevis (ECRB) to restore wrist extension (improvements in lifting objects, feeding, grooming, and hygiene is gained) and deltoid to triceps to provide elbow extension (410,411). This latter transfer enables patients to stabilize themselves for sitting and transferring, and to reach overhead against gravity, improving grooming, personal hygiene, pressure relief, writing speed and clarity, and self-feeding (412). The goal, however, should not be to improve the ability to perform transfers, since the latissimus dorsi, the prime shoulder depressor muscle, is lacking. A biceps to triceps transfer can also be performed (413,414). At the C6 functional level, the Moberg “key grip” procedure may be performed to restore the lateral or “key” grip in patients without a natural tenodesis (415,416). Improvements in functional activities include grooming, eating, writing, and desktop skills and these are maintained over time (417). Transfer of an active muscle, like the BR or others, to the flexor pollicis longus (FPL) to provide lateral pinch and to the finger flexors to provide grasp, provides better function and is preferable to Moberg’s simple FPL tenodesis (418–420). Posterior deltoid to triceps transfer may also be performed in C6 patients. This procedure is recommended prior to hand reconstruction, or simultaneously without diminishing outcomes. A supination contracture of the forearm that may occur in individuals with C5 and C6 motor injuries can be corrected by rerouting of the biceps brachii around the radial neck (421).

At the C7 level, the goal is to restore active grasp and improve hand control. Transfer of the BR to FPL can achieve restoration of thumb flexion. Restoration of finger flexion is achieved through transfer of ECRU or flexor carpi ulnaris to the flexor digitorum profundus (FDP). At the C8 level, an intrinsic minus or “claw hand” posture results. A lumbrical bar that prevents hyperextension of the metacarpal phalangeal may improve a patient’s function, and surgery is rarely indicated.

Postoperative immobilization of the extremity in a cast is usually between 1 and 3 months. The patient should be instructed in proper elevation techniques to control edema. Once the patient is cleared for ROM, a removable thermoplastic splint is fabricated to protect the tendon transfers from over stretching. Scar management may include breaking down adhesions and desensitization techniques. Biofeedback and therapeutic ES can help reeducate muscles following tendon transfers (422,423). Activities are graded progressively by performance time and resistance applied, eliminating any that create heavy resistance to the tendons until after 3 months. Data suggests that benefits from transfer procedures are maintained and patient satisfaction remains high (424,425).

The Freehand system is a neuroprosthesis that was previously implanted in patients with C5 or C6 tetraplegia. Movement of the contralateral shoulder facilitates control of the opposite hand to allow for grasp. The procedure can be performed in association with tendon transfers. Home use of the device and satisfaction with the neuroprosthesis is high (426–428), however, at the time of writing this chapter, this system is not available. Newer techniques of surgery with FES include advanced controls to allow individuals with C4 and higher lesions or C7 lesions to take better advantage of the neuroprosthesis, and nerve grafting for persons with C4-5 level of injury with LMN injuries. The nerve graft is attempted to restore volitional control or to reverse flaccid paralysis into spastic paralysis for application of FES (429).

**PSYCHOLOGICAL ISSUES**

SCI is a life-altering event for patients and their families. It is, therefore, not surprising that patients have a myriad of adjustment issues postinjury.

A person who sustains a SCI is at risk for the “four D syndrome”: dependency, depression, drug addiction, and, if married, divorce (430). Psychological issues such as depression will have an impact on functional capabilities (431).

Depressive disorders are the most common form of psychological distress after SCI, estimated to affect 20% to 45% of those injured, and usually occurs within the first month (432–435). All members of the rehabilitation team should be aware of signs of evolving depression in this population and react appropriately. Risk factors for depression include a prior history or
family history of depression, pain, female gender, lack of social support, multiplicity of life stresses, concurrent medical illness, and alcohol or substance abuse. Additional factors include having a complete neurological injury, medical comorbidity with TBI, a low level of autonomy, poor education, unemployment, having a poor social network and family support, few financial resources, architectural barriers, vocational difficulties, and the need for personal and transportation assistance.

The suicide rate for individuals with SCI is approximately five times the age–sex specific suicide rate in the United States. It is the leading cause of death in individuals with SCI in the youngest age groups, and is highest 1 to 5 years postinjury (22,436–437). The suicide rate is higher for those “marginally” injured (incomplete injuries) who have a near complete recovery (438). In addition to depression, lack of social support, history of suicide attempt, and a feasible plan for inflicting harm to self are risk factors for suicide. Substance abuse, which is increased in the SCI population, is also a major risk factor for suicide. It is imperative that physicians and other health care personnel continue to monitor patients for depression well after the acute phase of injury.

Anxiety and post-traumatic stress disorder (PTSD) have been reported in up to 20% of persons with SCI (439). Symptoms of PTSD may be higher in those with features of depression and/or anxiety (440).

The treatment for psychological disturbances after SCI includes counseling and pharmacological intervention. Medications should be considered for individuals who present biological, somatic, and/or mood-related symptoms of sufficient severity to disrupt the person's life and ADL. The medications should continue for at least 4 months (441,442), and be tapered slowly as discontinuation early is associated with a high rate of relapse (443). After this initial period, medication may need to be continued to prevent recurrence of symptoms. A psychiatrist should be consulted if the patient has suicidal ideation or presents with psychotic features, and/or the patient does not respond to one or two of the trials of antidepressant medications. Recurrence rates for depression are common.

There is a direct relationship between QOL scores and length of injury, suggesting that people may adapt to their injury with time. Individuals injured at a younger age tend to experience successful adjustment and a better QOL than older individuals (444). This may be due to their flexible use of coping strategies, and the development of a self-concept that includes the knowledge of future limitations (445). Additional variables associated with lower QOL scores included presence of neurogenic pain, spasticity, and bowel and bladder problems (446). Conversely, those who are married, currently employed, more highly educated, female, and Caucasian are more likely to have higher QOL scores (447,448). Interestingly, the QOL for persons with a high level of injury have been underestimated by health care professionals (449).

Substance Abuse
Substance abuse is frequently encountered in SCI. The CAGE, a four-question screening device for alcoholism, is a reliable instrument with the SCI population (450). At-risk drinkers and substance users tend to be younger, single, male, and less educated. Those with higher CAGE scores have a higher incidence of medical complications, including more pain, PUs, and lower satisfaction with life (451,452).

Alcohol and drug misuse contributes to SCI by increasing risk taking while intoxicated, hampers learning and rehabilitation gains, interferes with self-care, places persons at risk for complications, contributes to depression, mortality and morbidity, and limits long-term outcomes, and capacity for independent living (451–454). Heavy drinkers prior to injury spend less time in educational and vocational activities while in a rehabilitation hospital, and alcohol users are less motivated to participate in their rehabilitation than nondrinkers (455). Substance abuse prevention and treatment programs should be included as part of the SCI rehabilitation.

**SEXUALITY AND FERTILITY**

Sexuality issues should be discussed with the individual soon after the injury as part of the SCI education during inpatient rehabilitation. Aspects for discussion for men include the options for erectile function and capability to father children. Greater than 90% of men with complete and incomplete UMN lesions can achieve reflexogenic erections (456). Patients with complete UMN injuries have a less than 10% chance of achieving psychogenic erections, but approximately 50% of those with an incomplete UMN lesion may. While the majority of men with UMN lesions are able to obtain reflex erections, these erections are often poorly sustained and often are inadequate for intercourse. For persons with a complete LMN lesion, up to 12% can achieve reflexogenic erections, with approximately 25% being able to achieve psychogenic erections (456).

There are several treatment options available for erectile dysfunction. The phosphodiesterase class of medications, including sildenafil and vardenafil, has been used with success in the SCI population with UMN lesions (457–459). Hypotension is a potential adverse side effect with these medications and should not be used in patients taking nitrates (460). Patients at risk for AD should be so cautioned. Intracorporeal injections with prostaglandin E1 can induce an erection in those with UMN and LMN injuries. This should not be used in persons with sickle cell disease and the patient should be aware of the risk of priapism. Additional methods include penile implants, and vacuum and ring devices. Penile implants are effective, but have a relatively high failure rate and may cause infections or penile erosion.

Male fertility is variably affected by SCI. Retrograde ejaculation is reported to occur in 37% to 100% of men with SCI (461). Approximately 5% of men with complete UMN lesions and 18% of those with LMN lesions have ejaculations. The percentage is higher in those with incomplete injuries. Achieving ejaculation, however, does not insure successful reproduction. Less than 10% of couples will have successful spontaneous
pregnancies (462). Therefore, patients who experience infertility should be evaluated by a reproductive specialist soon after the decision to attempt pregnancy is made. Semen analysis in men with SCI reveals decreased sperm count and decreased sperm motility. To obtain sperm in men who do not ejaculate, penile vibratory stimulation (PVS) and if unsuccessful, electroejaculation (EEJ) can be attempted. As a last resort, direct testicular extraction of sperm can be performed. Couples have had success with in-vitro fertilization (IVF), gamete intrafallopian transfer (GIFT), intracytoplasmic sperm injection (ICSI), and intrauterine insemination procedures (463).

For women, SCI does not affect female fertility once menarche returns. Immediately following SCI, amenorrhea occurs in 85% of woman with cervical and high thoracic injuries and 50% to 60% of woman overall. Within 6 months and 1 year postinjury, 50% and 90% of women have return of menstruation. The completeness of injury does not appear to influence the menstrual cycle. With menstruation, women with SCI report increased spasticity, bladder spasm, and dysautonomia. Women with SCI experience menopause at similar ages to women without SCI.

Libido is generally preserved after injury and physiologic parameters of sexual function in women with SCI have been described (465). Preservation of psychogenic genital vascongestion is associated with the preservation of sensation in the T11-L2 dermatomes (464). Reflexogenic vascongestion can be induced via manual genital stimulation in women with UMN injuries. Approximately 50% of all women with SCI are able to achieve orgasm, although time to orgasm is prolonged compared to women without SCI (466). Women with LMN injuries affecting the S2-5 levels have reduced ability to achieve orgasm. Sildenafil has been shown to improve subjective arousal when used with manual stimulation, and may improve genital vascongestion as well (466).

Pregnancy presents a unique set of potential problems for women with SCI. Women may develop PUs, recurrent UTIs, increased spasticity, or decreased pulmonary function during pregnancy. There is a slightly increased incidence of preterm labor in SCI women (467). AD may develop in susceptible women during labor. Pre-eclampsia can be difficult to distinguish from AD, however, once the diagnosis of AD has been made, epidural anesthesia is the treatment of choice. The epidural should continue at least 12 hours after delivery or until the dysreflexia resolves. Patient should be followed closely by the SCI physician in cooperation with the obstetrician (see Chapter 51 for more details on the above).

Cardiovascular Disease, Obesity, and Diabetes After SCI

Cardiovascular disease (CVD) is a leading cause of mortality and morbidity in chronic SCI. In contrast to the able-bodied population, CVD occurs more frequently and at earlier ages in the SCI population (468,469). Several factors contribute to this increased incidence, including relatively sedentary lifestyle, poor dietary habits, dyslipidemia, increased prevalence of obesity and diabetes mellitus. Those with tetraplegia and neurologically complete injuries are at highest risk (470). The prevalence of CVD has been reported between 25% and 50% (468,469,471).

The pathophysiology of CVD in persons with SCI is multi-factorial. Changes in body composition occur and include decreased fat-free body mass and a relative increase in adipose tissue. The relative decrease of fat-free body mass contributes to lowered high density lipoprotein (HDL) levels, impaired glucose metabolism, and progressive insulin resistance (472,473). Given the prevalence of CVD in the SCI population, early recognition of risk factors is crucial to initiate early interventions which can modify these risks. The American Heart Association (AHA) identifies age (men >45 years, women >55 years, or post-menopausal), hypertension (SBP ≥ 140 or DBP ≥ 90), diabetes, smoking, HDL less than 35 mg/dL, family history of CVD in first degree relative, and obesity as significant risk factors (474). These guidelines have been established for the able-bodied population and may not be as reliable in predicting CVD risk in the SCI population. For example, body composition characteristics unique to those with SCI make traditional BMI measures less accurate in predicting body fat (475–478). Individuals with SCI also often have lowered BPs; therefore, the definition of hypertension established by the AHA may need to be modified for this population. Hypertension in the SCI population has been defined as increase of at least 20 points above the individual’s baseline.

Interventions to decrease the risk of cardiovascular events should initially focus on modifiable risk factor reduction including smoking cessation, weight loss, dietary modification, and increased physical activity levels. Those with hypertension, dyslipidemia, or hyperglycemia that persists after attempted behavior modification should receive appropriate pharmacotherapy (479). A structured exercise regimen is recommended (480). Arm ergometry (high intensity with a target HR of 70% to 80% of max predicted) and FES have shown efficacy in improving glucose tolerance and lipid profiles in persons with SCI (481–483).

Acute exercise responses and capacity for exercise conditioning are related to the level and completeness of the spinal lesion. Patients with complete SCI at or above T4 have diminished cardiac acceleration with maximal heart rates less than 130 beats/min (484). The work capacity of these persons is limited by reductions in cardiac output and circulation to the exercising musculature. Persons with paraplegia also have reduced exercise capacity and increased heart rate responses (compared with the nondisabled), which have been associated with circulatory limitations within the paralysed tissues.

The true incidence of obesity within SCI is likely underreported given the measures used to typically measure adiposity. Estimates (using % body fat) conclude that two thirds of persons with SCI are obese (485). Unfortunately, the presence of paralysis makes effective weight loss more difficult for wheelchair users. Recommendations for CV exercise include exercise at least three times per week (more to facilitate weight loss) for up to 90 minutes/d incorporating cardiovascular conditioning and strength training. Caution should be taken to avoid overuse
injuries to the UEs commonly seen in this population (486). Obesity has been associated with breast and colon cancer, CTS, stroke, coronary artery disease, diabetes mellitus, hypertension, dyslipidemia, obstructive sleep apnea, and PUs (485,487).

Diabetes mellitus has been reported to occur in approximately 20% of individuals with chronic SCI. Studies have shown that oral glucose tolerance testing may detect insulin resistance earlier than fasting blood glucose measures (488,489). Early detection is helpful in initiating early treatment, a key component to minimize the long-term complications associated with diabetes mellitus including chronic skin ulcers, peripheral vascular disease, heart attacks, and strokes. Symptoms of hyperglycemia include polydipsia and polyuria. In the SCI patient population, polyuria may increase if insulin is not taken appropriately, new onset urinary incontinence, or AD. The measurement of the hemoglobin A1C is also helpful in making the diagnosis of diabetes mellitus and monitoring the efficacy of treatment.

Long-Term Follow-Up
Long-term follow-up with a SCI specialist is extremely important, especially with shorter LOS in acute rehabilitation after injury. Initially, visits should occur on a monthly basis, especially while the person is on outpatient therapy. This allows for monitoring of medical issues, reevaluation of the therapy program, and updating goals and equipment prescriptions. As patients go home from inpatient hospitalization earlier, medical issues that previously were experienced in the hospital may develop while at home. This includes bowel, bladder, and spasticity changes, and the possible development of HO, hypercalcemia, and AD. After medical issues have stabilized and outpatient therapy has concluded, visits are recommended every 3 to 6 months throughout the first year.

For those who are medically stable, at least yearly visits are recommended. Persons with SCI require regular and comprehensive health care throughout their lifetime. This includes routine health monitoring and care for nonSCI medical issues as well as for SCI-specific problems. The altered physiology and the absence of many typical symptoms for common problems following SCI pose a unique challenge for health care providers. The importance of these visits is also to review any medical changes, monitor the neurological examination for any changes, ensure the maintenance of the equipment, and prescribe any additional equipment that may be needed. Deterioration of the neurologic status may be secondary to a tethered cord, syringomyelia, or peripheral problems, such as median or ulnar nerve entrapment, or other musculoskeletal complications. As a greater proportion of persons with SCI are surviving after their injury, the importance of their follow-up visits to maintain their QOL cannot be overemphasized.

Guidelines for routine screening and preventive health care should be used as a starting point for people with SCI (490). Yearly evaluations should include the patient’s current condition and concerns, medications, review of systems (ROS), and age-appropriate screening (491). Specifics include dental, vision, and hearing screens, evaluating for cognitive slowing, depression, and suicide. Questions related to smoking, alcohol and drug use, exercise, and current employment status also should be included. Areas of focus should also include the unique issues of SCI and aging with a disability. An expanded ROS should address the specific problems common in SCI, including bladder and bowel control, appetite, sleep, BP, skin integrity, pain, spasticity, sexual function, equipment needs, and changes in strength, sensation, and functional ability.

In addition to a general physical examination, areas to examine include the skin, oral cavity in smokers, digital rectal examination for men over age 50, the SCI-specific neurological examination, and a musculoskeletal examination to identify problems related to aging with SCI and pain. A gynecologic examination is recommended for a clinical breast exam for women over 50 and a pelvic examination.

Counseling is an important aspect of follow-up visits, both for medical as well as physical and social issues. Lifestyle issues to review include diet and nutrition, smoking, physical activity, wheelchair and other equipment maintenance and safety, PU prevention, compliance with the individual’s bladder management program, and advice against alcohol abuse.

Recommendations for the general medical screening should be followed in persons with SCI, including screening for colon cancer with a flexible sigmoidoscopy for those over age 50, and colonoscopy if at high risk. Mammography is recommended for women every 1 to 2 years between the ages of 50 and 69. Fasting glucose every 3 years for those over age 40 or annually for persons with risk factors for type-II diabetes, screening for nutritional deficiencies with complete blood count, albumin and iron levels (if at risk), fasting lipid profile for men over 40 and for women over 50, and prostate-specific antigen testing in men between age 50 and 70 or after age 45 if at increased risk, are also recommended (491). There have been few studies formally evaluating these recommendations in the SCI population (see Chapter 57 for more details of primary care management for persons with disability).

At times, readmission to the rehabilitation hospital for medical (e.g., PU, AD) or rehabilitation issues may be needed. For medical issues, the SCI physician and team are able to care for the patient with consultation of other specialists, as required. This includes the importance of maintaining the proper bowel and bladder programs, skin issues, and awareness and knowledge of treatment for AD. In addition, many times the person with SCI may require a “refresher course” in rehabilitation techniques that can best be taught in an intensive inpatient setting.

Secondary medical complications are extremely common in patients with chronic SCI. The most common reasons for rehospitalization include GU complications, PUs, and respiratory complications (173,317). Pneumonia is more common in persons with tetraplegia while pressure sores are more common in persons with paraplegia. Rehospitalization rates have remained static over the last 10 years (317). Patients from skilled nursing facilities, with lower motor FIM scores, or using state or federal health insurance plans, have higher rehospitalization rates.
THE FUTURE

The quest for cure is as strong as ever, with many new advances constantly being made. Over the last decade, there is increased knowledge regarding the basic pathophysiology of SCI and new pharmacological strategies in the acute and chronic treatment to enhance neurological and functional recovery. In animal models, scientists have successfully used antibodies to block growth inhibitors, applied growth factors, implanted cells and peripheral nerve bridges, and used gene therapies to stimulate spinal cord regeneration. Some advances are nearly ready for application in the clinical setting. The key factors for cure include minimizing secondary effects of injury, neutralizing the effects of substrates that inhibit CNS regeneration, delivery of regeneration-promoting substances to the injured spinal cord, allowing a bridge to which the spinal cord axons can attach and grow along after injury, and determining the genes that may allow for “turning on” axonal growth after injury (see Chapters 81 and 82 for more details).

Basic science research continues to make progress with the ultimate goal of a cure for SCI. Rehabilitation is a crucial part of any cure treatment strategy and is what we currently have available to maximize functional potential. Rehabilitation professionals should be involved in the research but remain committed to the patients and their needs. Caring for the acute and long-term medical issues, as well as assisting the patient to be active in all domains including social, recreational, and vocational activities, is what rehabilitation is all about!

REFERENCES


PART III | MAJOR CONDITIONS


INTRODUCTION

Adult motor neuron disease (MND) is often considered synonymous with amyotrophic lateral sclerosis (ALS). In the United States, the term ALS is frequently used to describe all forms of adult-onset MND. However, it is also used to refer specifically to the most common form of adult MND, which is sporadic, acquired ALS. In the United Kingdom, the converse is true where the generic term for all forms of ALS is MND (1). In reality, adult MND actually encompasses a group of disorders that include ALS; primary lateral sclerosis (PLS); progressive muscular atrophy (PMA); progressive bulbar palsy (PBP); adult-onset, progressive spinal muscular atrophy (SMA); and X-linked, recessive spinobulbar muscular atrophy (SBMA). A presentation with pure upper motor neuron (UMN) signs may be called PLS, whereas pure bulbar presentation may be called PBP, and pure lower motor neuron (LMN) presentation called PMA. Whether these conditions exist as distinct diseases or rather represent part of the spectrum of ALS is still debated. This is represented schematically in Figure 28-1.

At least one form of PLS with a benign course and autosomal dominant inheritance has been reported (2,3). ALS is also referred to as Lou Gehrig’s disease, named after perhaps the most famous person yet afflicted with the disease. SBMA is commonly referred to as Kennedy’s disease, named after the physician who first described this disorder in 1968 (4).

For the purposes of this chapter, we will use the terms ALS, PLS, PMA, PLS, SMA, and SBMA. In the adult population, ALS is far more common than the other disorders. Thus, ALS will constitute most of the focus of this chapter, which starts first with a description of the diseases, including epidemiology and genetics. This is followed by diagnostic workup, including electrodiagnosis, pharmacological management, and rehabilitation strategies, most of which may be applied to any of the adult MNDs.

OVERVIEW OF THE MAJOR ADULT MNDs

Amyotrophic Lateral Sclerosis

ALS is a rapidly progressive neuromuscular disease that destroys both UMNs and LMNs, resulting in spasticity and diffuse muscular atrophy and weakness. The vast majority of ALS cases are presumably acquired and occur sporadically. However, approximately 10% of all ALS cases are familial amyotrophic lateral sclerosis (FALS) and usually inherited as an autosomal dominant trait. About 15% of these cases result from a gene defect on chromosome 21q12.1, which leads to a toxic gain of function in the antioxidant enzyme Cu/Zn superoxide dismutase (SOD1) (5,6). Over 100 unique SOD1 mutations have been identified (5–7). Emerging evidence suggests that these mutations result in increased oxidative stress for the motor neurons, leading to cell death, which is felt to be related to free radical toxicity (6,7).

The etiology of sporadic amyotrophic lateral sclerosis (SALS) and the other 85% of FALS is unknown. Increasing data suggest that excessive glutamate activity in the brain and spinal cord may play an important role. Glutamate is one of the main central nervous system (CNS) excitatory neurotransmitters in the brain, and excess levels of this chemical have been demonstrated in the serum, spinal fluid, and brain tissue of ALS patients (8,9). There appears to be reduced clearance of glutamate from critical motor control areas in ALS as well as decreased levels of glutamate transport protein (10,11).

Epidemiology of ALS

ALS most commonly strikes people between 40 and 60 years of age with a mean age of onset of 58 years (12–14). The overall
prevalence rate in the worldwide population is somewhere between 5 and 7 per 100,000, making it one of the most common neuromuscular diseases worldwide (15). Further, population studies suggest that the incidence of ALS is increasing, although this is probably due, in large part, to people living longer and better recognition of the diagnosis (16,17). There appears to be a higher incidence in urban areas, felt to be related to environmental factors (17–19). The association of nutrient intake with the risk of ALS was investigated in a population-based case-control study conducted in three counties of western Washington State from 1990 to 1994 (20,21). The authors found that alcohol consumption was not associated with increased risk of ALS. Ever having smoked cigarettes was associated with a twofold increase in risk and a greater than threelfold increased risk was observed for current smokers. Further, dietary fat intake was associated with an increased risk of ALS, while dietary fiber intake was associated with a decreased risk. Interestingly, consumption of antioxidant vitamins from diet or supplement sources did not alter the risk but glutamate intake was associated with an increased risk of ALS. The finding that cigarette smoking and glutamate consumption are risk factors for ALS is consistent with current etiologic theories that implicate glutamate excitotoxicity and oxidative stress in the pathogenesis of ALS. The associations with fat and fiber intake warrant further study and biologic explanation.

Considerable clustering has been demonstrated in the Western Pacific region of the world (15–17). Other sporadic cluster cases have been reported but without obvious environmental or causal factors (15). Men appear to be more commonly affected than women with a male-to-female ratio of about 1.5:1.0 (12).

Poor prognostic factors include older age at time of onset, bulbar and/or pulmonary dysfunction early in the clinical course of the disease, short time period from symptom onset to diagnosis, and predominance of LMN findings at the time of diagnosis (12,14–16).

More women than men present with bulbar symptoms, and the progression of bulbar palsy appears to be more rapid in women (20,21). Young males with ALS may have a longer life expectancy but overall the median 50% survival rate is 2.5 years postdiagnosis, except in patients with primary bulbar symptoms, where the 50% survival rate is only 1 year (22). Survival rates will obviously vary to a degree depending on the patient’s decision to use or not use mechanical ventilation and a feeding tube (23). Nonetheless, by 5 years postdiagnosis the overall survival rate is only 28% (12,14,15).

Atypical, “ALS-like,” MNDs have been reported infrequently as a remote complication of several malignancies, including lymphoma and small cell carcinoma of the lung (24,25). These likely represent paraneoplastic syndromes and not a true manifestation of ALS (26). Regardless, patients with atypical MND should be screened for malignancy.

Spinal and Spinobulbar Muscular Atrophy

There are many forms of SMA, all of which involve selective destruction of anterior horn cells. The various forms of SMA are clinically dissimilar, with some rare forms affecting distal or bulbar muscles only. The most common forms are often referred to as types I, II, and III (27). These are mostly disorders of childhood and are usually inherited as autosomal recessive traits. SMA I, also known as Werdnig-Hoffman disease (WHD) or acute, infantile-onset SMA, is a severe disorder resulting in death before age 2 years. SMA II, also referred to as early-onset, intermediate SMA or chronic WHD, is less severe, with signs and symptoms becoming apparent in the first 6 to 18 months of life. SMA III, also known as Kugelberg-Welander disease (KWD), is a chronic, later onset disorder, associated with significantly less morbidity. Signs and symptoms of SMA III usually become apparent between ages 5 and 15 years. In prior studies looking at SMA II and III over a 10 year period, SMA II subjects showed marked weakness and progressive decline of strength while SMA III subjects had a relatively static or very slowly progressive course and were far stronger. In both SMA II and SMA III, proximal weakness was greater than distal. Joint contractures, progressive scoliosis, and restrictive lung disease (RLD) were present in most of the SMA II individuals but these complications were rare in SMA III (27). Scoliosis in a patient with SMA II is shown in Figure 28-2.

There are two forms of SMA that have onset in the adult age group. One is an adult-onset form of SMA with age of onset of 17 to 55 years with either recessive or dominant forms of inheritance (28,29). The disease clinically appears much like SMA III, although it may be more progressive. The other form...
is SBMA, or Kennedy's disease, a sex-linked, recessive MND characterized by progressive spinal and bulbar muscular atrophy, gynecomastia, and reduced fertility (4,30).

Adult-onset SMA, SBMA, and SMA III patients can have normal life spans, and many of the rehabilitative modalities discussed in this chapter are applicable to this population. Further, with the rapid advancement of rehabilitation technology, many SMA II patients are now living well into adulthood and successful pregnancies have been reported in this disease (31).

Genetics of SMA
A detailed analysis of the 5q13 region revealed that this chromosomal region in humans contained a large inverted duplication with at least two genes present in telomeric and centromeric copies. Further studies have identified the SMA causative gene as the survival motor neuron 1 gene (SMN1, telomeric copy), along with a disease modifying gene (SMN2, centromeric copy) (32–35). Briefly, the two SMN genes are nearly identical except for a difference of only five nucleotides in their 3′ regions, without any alteration of the amino acid sequence of the protein. However, the critical difference between SMN1 and SMN2 genes is a C-T transition located within the exon splicing region of the SMN2 that affects the splicing of exon 7. This change results in frequent exon 7 skipping during the splicing of SMN2 transcripts (32,33). It is thought that the resulting truncated SMN protein without its exon 7 contribution is a less stable form of SMN protein and is therefore rapidly degraded. In about 95% of SMA patients, both copies of SMN1 exon 7 are absent due to mutations. In the remaining SMA affected patients, other small or subtle mutations have been identified (33,34).

Genetic studies have now established that SMA is caused by mutations in the telomeric SMN1 gene, with all patients having at least one copy of the centromeric SMN2 gene. At least one copy of the SMN2 must be present in the setting of homozygous SMN1 mutations; otherwise, embryonic lethality occurs. The copy number of SMN2 varies in the population, and this variation appears to have some important modifying effects on SMA disease severity (34). It appears that higher number of SMN2 copies in the setting of SMN1 mutations results in less severe clinical SMA phenotype. However, substantial variations in SMA phenotype and disease severity can exist with a given SMN2 copy number, so it is not recommended to predict disease severity based on SMN2 copy numbers. Although we now know that SMN protein is expressed widely in many tissues throughout the body, its function is still not completely understood at this time (35).

Genetics of SBMA (Kennedy's Disease)
SBMA is a hereditary adult-onset disease that causes preferential degeneration of LMNs leading to weakness and atrophy of bulbar, facial, and limb muscles. It is clinically similar to ALS. The main clinical distinction between the two is that ALS involves degeneration of both UMN and LMNs, whereas the affected cell types in SBMA are LMNs.

Another interesting difference to note is that Onuf’s nucleus, an androgen-sensitive spinal cord motor neuron nucleus, is spared in ALS, although it degenerates in SBMA (30,35). Degeneration of sensory neurons of the dorsal root ganglia is also a typical sign associated with SBMA, often preceding the onset of motor dysfunction. In addition to the neurological phenotype, SBMA patients display some of the characteristic signs of androgen insensitivity syndromes including testicular atrophy, decreased fertility, gynecomastia, and elevated androgen levels (36). SBMA is caused by a novel mutation, the expansion of a trinucleotide CAG repeat, in the first exon of the androgen receptor gene (32). Unaffected individuals have a CAG repeat size that ranges between 5 and 35 glutamines, while symptomatic individuals always have a repeat size of at least 37 glutamines (32).

SBMA has some clinical variability; however, phenotypic expression does not correlate with the length of CAG repeats. This is in contrast to myotonic muscular dystrophy and fragile X syndrome, where increased numbers of tandem triplet repeats correlate directly with disease severity (30). Commercially available blood tests (DNA analysis) are now available for SMA and SBMA. SBMA can occur without any family history or gynecomastia and all males with atypical ALS should be tested for SBMA. Prevalence rates for SMA types II and III are estimated to range from as high as 40 per million among children to around 12 per million in the general population, with adult-onset SMA and SBMA being far less common (27).

DIAGNOSTIC EVALUATION OF MND
The diagnosis of ALS and other forms of adult MND is primarily a process of exclusion. If, based on the history and physical examination, clinical signs and symptoms of MND are detected, one must generate a differential diagnosis and then work to exclude processes mimicking MND. Only in FALS with known SOD1 mutations, Kennedy's disease, and the few adult-onset SMA cases in which SMN mutations are detected is a definitive diagnostic test available. For most patients with ALS or its variants (PMA and PLS), electrodiagnostic testing (EDX), laboratory testing, neuroimaging studies, and, occasionally, a muscle biopsy are used to exclude other diagnoses. The El Escorial criteria (Fig. 28-3) are used to assess the certainty of a diagnosis of ALS once other disease processes have been excluded.

Clinical Presentation
Patients with ALS most often seek evaluation complaining of focal weakness (60%), rarely of generalized weakness or cramps, and very rarely of generalized fasciculations or respiratory failure. Symptom onset may be anywhere within the motor system, including areas of the brain and brainstem as shown in Figure 28-4. Although fasciculations are a prominent feature in most patients with ALS, patients who complain of fasciculations only and have an otherwise normal neurologic
examination usually have benign fasciculation syndrome and are unlikely to develop ALS. Symptom onset may be anywhere within the motor system.

The evaluation of a patient suspected of having MND begins with a detailed history, general physical examination, and neurologic examination. On neurologic examination, one is looking for evidence of UMN and LMN dysfunction. The mental status, nonmotor cranial nerve function, sensory examination, and cerebellar examinations should be normal. The symptoms may be quite insidious and develop slowly or in unusual motor patterns. Loss of function may not be initially obviously disabling but performance may be affected. This is demonstrated by the noted decline in Lou Gehrig’s batting average over the last 10 years of his professional baseball career (Fig. 28-5).

Patients with UMN pathology often complain of loss of dexterity or a feeling of stiffness in the limbs. They may note weakness which is caused by spasticity resulting from disinhibition of brainstem control of the vestibulospinal and reticulospinal tracts. Findings on examination include spasticity and hyperreflexia, indicated by abnormal spread of reflexes and clonus or by the presence of brisk reflexes despite muscle atrophy due to LMN loss. The gold standard used to diagnose UMN pathology is the presence of pathologic reflexes, such as the Babinski’s sign, Hoffman’s sign, and jaw jerk. If the toe extensors are paralyzed, visualization of contraction of the tensor fascia lata when an attempt is made to elicit a Babinski response has the same significance as great toe extension. Recently, it has been suggested that the corneomandibular reflex may be a more sensitive and specific indicator than the jaw jerk of UMN pathology in the bulbar region (37).

Patients with LMN pathology usually present complaining of muscle weakness. In addition, they may note muscle atrophy, fasciculations, and muscle cramping. Cramping may occur anywhere in the body, including the thighs, arms, and abdomen. Cramping of abdominal or other trunk muscles raises a red flag urging the clinician to consider a diagnosis of ALS. Findings on examination include weakness, atrophy, hypotonia, hyoporeflexia, and fasciculations. Head drop is a manifestation of muscle weakness often seen in ALS although it can be seen in other neuromuscular disorders; ALS and myasthenia gravis are the two most common causes of head drop. Atrophy often appears first in the hand intrinsic muscles. Although fasciculations are not a necessary criteria for the diagnosis of ALS, one should question the diagnosis when none are observed. A recent study identified the following hierarchy of initial symptoms: leg weakness in nearly half of ALS patients; followed by arms, bulbar muscles, and then generalized weakness (38). Presentation with only respiratory muscle weakness is extremely rare.
Signs and symptoms suggesting bulbar muscle weakness include dysarthria, dysphagia, drooling, and aspiration. These signs and symptoms may be caused by UMN and/or LMN dysfunction involving the bulbar muscles. Signs of spastic dysarthria, indicating UMN pathology, include a strained and strangled quality of speech, reduced rate, low pitch, imprecise consonant pronunciation, vowel distortion, and breaks in pitch. LMN dysfunction creates a flaccid dysarthria in which speech has a nasal and/or wet quality; pitch and intensity are monotone, phrases abnormally short, and inspiration audible. Complaints of difficulty chewing and swallowing, nasal regurgitation, or coughing when drinking liquids, may all indicate dysphagia. On physical examination, the following tests may be used to assess facial and bulbar muscle function: ability to bury the eyelashes, pocket air in the cheeks, whistle; jaw opening and lip closure strength, phonation of a variety of syllables such as "puh," "kuh," "tuh," and "ah." The tongue should be examined for fasciculations and atrophy and tongue strength and range of motion assessed. The gag reflex and jaw jerk should be assessed to look for UMN dysfunction. Pseudobulbar affect is a symptom of pseudobulbar palsy which refers to an UMN syndrome caused by motor neuron loss in the corticobulbar tracts. Patients experience inappropriate laughter or crying which is not concordant with their mood and can be embarrassing. Disinhibition of limbic motor control produces pseudobulbar affect, also sometimes called "emotional incontinence."

In patients who present with respiratory failure, the earliest signs are often nocturnal and include poor sleep with frequent awakening, early morning headaches, excessive daytime fatigue and sleepiness, nightmares, and orthopnea. Frequent sighing, a weak cough, and difficulty clearing bronchial or pulmonary secretions are other signs or respiratory muscle weakness. Later signs of respiratory dysfunction are dyspnea with exertion, truncated speech, respiratory paradox, dyspnea when eating, rapid shallow breathing, visible accessory muscle contraction, and flaring of the nasal alae. With advanced, untreated respiratory failure, patients may have an elevated hematocrit, low serum chloride, respiratory acidosis with a compensatory metabolic alkalosis, hypertension, and cor pulmonale.

Other signs and symptoms frequently associated with ALS are cachexia, fatigue, and musculoskeletal complaints. The term "ALS cachexia" refers to a phenomenon experienced by some patients in which weight loss occurs in excess of that caused by muscle atrophy and reduced caloric intake. Both subcutaneous fat and peritoneal fat are lost, presumably because of acceleration of the basal metabolic rate (39). In patients with ALS cachexia, greater than 20% of body weight is typically lost over a 6-month period. Many patients with ALS feel an overwhelming sense of muscle fatigue which is probably due to a combination of blocking of neuromuscular transmission in reinnervated nerve terminal sprouts and impairment of excitation contraction coupling (39). Some patients seek initial medical attention because of fractures or sprains that do not heal. In reality, these patients probably sustained their initial injury because of a fall or other injury (e.g., sprained ankle) that occurred because of underlying muscle weakness; they were then unable to recover to their premorbid level of function because of that weakness. Other common musculoskeletal complaints include neck and back pain, shoulder pain due to a frozen shoulder, elbow flexion and ankle plantar flexion contractures, and claw hand. Patients may experience osteoporotic fractures and/or stress fractures because of immobilization-induced bone density loss.

Rare signs and symptoms which usually occur only in advanced ALS include sensory impairment, autonomic dysfunction, bowel and bladder dysfunction, extraocular muscle paralysis, pressure ulcer formation, and severe dementia. Although ALS is discussed as a pure motor disorder, some patients complain of paresthesias. These may be due to compression or entrapment neuropathies, but subclinical abnormalities in somatosensory evoked potentials and quantitative sensory testing have been reported (40).

**Differential Diagnosis**

After obtaining a history and examining the patient, the clinician is able to generate a differential diagnosis which guides further diagnostic testing. The differential diagnosis differs depending on whether the presentation is primarily LMN, UMN, bulbar or mixed LMN and UMN.

**El Escorial Criteria**

The El Escorial criteria (see Fig. 28-3) for diagnosing ALS were developed by a task force of the World Federation of Neurology in 1990 to ensure inclusion of more homogeneous patient populations in ALS clinical trials (41). These criteria have been used to enroll patients in most of the recent clinical trials. The criteria were revised in 1998 to improve the speed and certainty of diagnosis (42). The criteria classify the certainty level of the diagnosis of ALS as falling into one of five categories: definite, probable, probable with laboratory support, possible, and suspected. In brief, the motor system is divided into four regions: bulbar, cervical, thoracic, and lumbosacral. Clinical evidence of UMN and LMN pathology is sought in each region. The certainty level of diagnosis depends on how many regions reveal UMN and/or LMN pathology. Figure 28-3 summarizes the
PART III | MAJOR CONDITIONS

schema for placing patients in the five diagnostic categories. Clinical weakness, atrophy, and fasciculations are considered evidence of LMN pathology. Pathologic spread of reflexes, clonus, and pseudobulbar features are considered evidence of UMN pathology. Electrophysiologic findings can be used to both confirm LMN dysfunction in clinically affected regions and to detect LMN dysfunction in clinically uninvolved regions. Neuroimaging and clinical laboratory studies are used to exclude other conditions that may mimic ALS.

Electrodiagnostic Testing
The various forms of MND, including SMA, Kennedy’s disease, PMA, SALS, and FALS share several electrodiagnostic features but also differ in some aspects due to varying rates of disease progression (43). General EDX characteristics of MND include normal sensory nerve conduction studies (NCs), normal or low motor amplitudes depending on disease stage, and normal distal motor latencies and conduction velocities. However, with profound loss of motor amplitude, conduction velocities may drop as low as 25% below the lower limit of normal because of loss of the fastest conducting fibers. The needle electrode examination (NEE) reveals a decreased recruitment pattern, either normal size or large motor unit action potentials (MUAPs) with or without evidence of remodeling depending on the specific disease process, and abnormal spontaneous activity including positive sharp waves (PSWs), fibrillation potentials, fasciculations, and complex repetitive discharges (CRDs). The prominence of the various forms of spontaneous activity varies with the different forms of MND.

Spinal Muscular Atrophies
The EDX portion of the autosomal recessive SMAs I–IV are determined by the rate of anterior horn cell degeneration and the stage in the course of the disease. Sensory NCs are normal in all forms of SMA. Compound motor action potentials (CMAPs) are decreased in proportion to the degree of muscle atrophy. Motor velocities are most likely to be abnormally slow in SMA I because of the extensive loss of motor axons.

The most profound loss of MUAPs is seen in SMA I. With maximal effort, only a few MUAPs may fire at a rapid rate. Small MUAPs are common because reinnervation cannot compensate for the rapid loss of anterior horn cells. Myopathic appearing low amplitude, polyphasic, short duration units may also be seen because of muscle fiber degeneration. In the other types of SMA, one sees large amplitude MUAPs (up to 10 or 15 mV) because the number of fibers per motor unit increases as motor unit remodeling occurs. These large units also tend to be polyphasic with increased duration. Satellite potentials appear as remodeling occurs. Myopathic appearing MUAPs are also seen in some older patients with SMA III, and their etiology is not well understood.

On NEE in SMA I, fibrillation potentials and PSWs are diffuse and seen in many muscles, including the paraspinals. In the more chronic forms of SMA, fibrillation potentials and PSWs are even more common and increase in frequency as age increases. CRDs are often seen in SMA II and III, and fasciculations are more common than in type I (44–46).

**Kennedy’s Disease**
Motor NCs abnormalities are similar to those seen in other forms of MND. Although patients generally do not have sensory complaints, absence or reduction of sensory nerve action potential (SNAPS) is a common finding (47,48). NEE shows large amplitude and duration MUAPs consistent with the rather indolent disease course. Fibrillation potentials and PSWs may be very prominent and present in all muscles examined. Fasciculation potentials are also abundant in limb, facial, and tongue muscles.

**Adult Nonhereditary MND**
For many years, Lambert’s criteria were the standard for the electromyographic (EMG) diagnosis of ALS (49,50). The following four criteria were required to make a definite diagnosis of ALS: (a) PSWs and/or fibrillation potentials in three of five limbs, counting the head as a limb. For a limb to be considered affected, at least two muscles innervated by different peripheral nerves and roots should show active denervation. (b) Normal sensory NCs. (c) Normal motor conduction studies; however, if the CMAP amplitude is very low, conduction velocity may drop as low as 70% of the lower limit of normal. (d) Reduced recruitment of MUAPs on needle exam. More recently, MND signs on physical examination. By definition, the EDX examination is normal in PLS. In PBP, active denervation is found only in muscles of the head and neck.

The EDX portion of the El Escorial criteria differs somewhat from Lambert’s criteria and is generally more liberal. The revised El Escorial criteria allow EDX findings to be used to upgrade the certainty of a diagnosis from clinically possible ALS to probable ALS; this upgrading of the diagnosis is important because it often allows additional patients to participate in clinical trials which generally require a diagnosis of probable or definite ALS. The El Escorial EDX criteria state that active denervation must be present in two of the four spinal regions (bulbar, cervical, thoracic, and lumbar) to support a diagnosis of ALS. For the cervical or lumbosacral region to be counted, at least two muscles innervated by different nerve roots and peripheral nerves must have EMG changes. In the bulbar and thoracic regions, changes in one muscle are sufficient. Thus, a patient with active denervation in the left arm and thoracic paraspinals would meet the El Escorial criteria for an EDX diagnosis of ALS but not the Lambert criteria because only one limb is involved. On the other hand, a patient with denervation in the tongue and both arms would fulfill Lambert’s criteria and the El Escorial criteria for ALS because three limbs and two regions are involved (bulbar and cervical). Early in the
progression of ALS, many patients with a suspected clinical
diagnosis do not meet EDX criteria for a definite diagnosis.
A repeat study several months later will often fulfill the EDX
criteria for diagnosis.

NCS in ALS are characterized primarily by decreased
CMAP amplitudes. The mild slowing of motor conduction
velocity and the prolongation of F wave latencies are attributed
to loss of the fastest conducting axons. An interesting phenomenon observed in many patients is that of the “split hand”;
CMAP amplitudes are decreased to a greater extent on the
radial side of the hand than on the ulnar side. CMAPs obtained
from the abductor pollicis brevis and first dorsal interosseous
are much lower than those obtained from the abductor digitii
eminendi. More than two stimulation sites should be used in the
evaluation of motor nerves to exclude the presence of conduc-
tion block since multifocal motor neuropathy with conduction
block is occasionally misdiagnosed as ALS. The ulnar nerve
easily can be stimulated at the wrist, below the elbow, above the
elbow, in the axilla, and in the supraclavicular fossa. In limbs
with UMN signs, H-reflexes may be elicited from muscles in which
they cannot normally be obtained. A few patients do have SNAP
abnormalities, and the sympathetic skin response is absent in 40%, suggesting subclinical autonomic nervous system involvement (52). Repetitive stimulation studies may show a decrement in CMAP with stimulation at 3 Hz which is similar to that seen in myasthenia gravis. A decrement is especially likely to be detected in patients with rapidly pro-
gressing disease and in muscles with an abundance of fascicu-
lations (43).

The NEE is the most important part of the EDX exami-
nation in cases of suspected ALS. Fasculation potentials are
seen in most patients with ALS, but they are not necessary to
meet diagnostic criteria, and the presence of only fasciculations
is inadequate as evidence of LMN involvement of a particular
limb or region. The significance of fasciculations depends on
the company they keep; they are pathologic only when accom-
panied by fibrillation potentials, PSWs, or recruitment pattern
or MUAP size changes. In patients with advanced ALS, fibrillation potentials and PSWs are prominent in most mus-
cles, but they may be sparse early in the course of the disease.
Occasionally, CRDs and doublets or triplets are seen in patients
with ALS, but these are not typical EDX findings in ALS. The
thoracic paraspinals should be examined with a needle; they
are not involved in tandem cervical and lumbar stenosis and
can help exclude this as a diagnostic possibility. In addition,
when the El Escorial criteria are employed, the finding of den-
ervation in the thoracic and in either the cervical or the lumbar
region is sufficient for a definite diagnosis, making examina-
tion of the tongue or facial muscles, which many patients find
unpleasant, unnecessary. Although fasciculations and denerv-
ation of the tongue are considered almost pathognomonic
for ALS, they are seldom found in patients who do not have
clinical evidence of bulbar muscle involvement. The recruit-
ment pattern is decreased in involved muscles. If the disease is
progressing relatively slowly, MUAP amplitudes and durations
become increased; but if the course is very rapid, denervation
outpaces reinnervation and enlarged MUAPs do not have time
to develop. The density and distribution of fasciculations and
fibrillations do not correlate with disease course or prognosis,
and serial EDX examinations are not useful for monitoring
disease progression once a definite diagnosis has been made.

Neurophysiologists have begun to explore the use of tran
scranial magnetic stimulation as a method of identifying
subclinical UMN dysfunction. Results are contradictory with
respect to the sensitivity and specificity of various findings as
evidence of UMN dysfunction, and these techniques must be
considered experimental at present. Abnormalities suggesting
UMN pathology include a motor evoked potential (MEP)
much lower in amplitude than the CMAP recorded from the
same muscle, prolonged central motor conduction time,
decreased MEP thresholds and silent periods early in disease,
increased MEP thresholds in advanced disease, and decreased
cortical representation of individual muscles (53–56).

Neuroimaging
Imaging studies are used to exclude possibilities other than
MND from the differential diagnosis. Magnetic resonance
imaging (MRI) is the primary imaging modality used in the
evaluation of patients with suspected ALS. Almost all patients
should have an MRI of the cervical spine to rule out cord com-
pression, a syrinx or other spinal cord pathology. The loca-
tion of symptoms will dictate whether or not other regions of
the spinal cord should be imaged. In patients presenting with
the PMA phenotype, an MRI of the involved region of the
spinal cord with gadolinium should be considered to look for
a metastatic polyradiculopathy. In those presenting with bul-
bar symptoms, a brain MRI should be performed to rule out
stroke, tumor, syringobulbia, etc.

Although MRI is generally not performed to confirm a
diagnosis of ALS, a few associated abnormalities have been
reported. Rarely, spinal cord and motor cortex atrophy is
apparent. Corticospinal tract hyperintensity with T2 imaging
has been observed in a few younger patients with a predomi-
nance of UMN signs (53).

Laboratory Evaluation and Other Diagnostic Tests
In most neuromuscular clinics, a routine panel of laboratory
tests is performed for all patients suspected of having ALS. The
rationale behind performing this battery of tests is to assess the
general health of the patient and exclude treatable conditions.
The differential diagnosis, developed following the history and
physical exam, may suggest that more specialized testing be per-
formed. Additional tests may be warranted when the presenta-
tion is with the PMA, PLS, or PBP phenotype. When there is a
family history of MND, genetic testing for FALS is performed.

ALS CLINICAL TRIALS
Clinical trials play a crucial role in the development of new
therapeutic agents. It is important to begin thinking about
what an appropriate trial would look like far in advance of actual study initiation. Decisions regarding dose, outcomes to be measured, and duration of treatment all may have a critical impact on whether a new agent is found to be efficacious. Many of these decisions depend on adequate preclinical data. In the absence of good preclinical data, trials may fail to show efficacy or even demonstrate harm to patients that could have been avoided. In recent years, much has been learned about pathogenic mechanisms of ALS, leading to a proliferation of new targets for disease modification. Mitochondrial dysfunction, glutamate toxicity, protein misfolding, and microglial activation are just a few mechanisms that have been proposed. For each proposed mechanism, pharmacological manipulation is possible. Targeted drug discovery programs can lead to new compounds, and reevaluation of existing drugs may lead to recognition of properties not previously investigated. Recently, a collaborative effort jointly funded by the National Institute of Neurological Disorders and Stroke (NINDS) and the Amyotrophic Lateral Sclerosis Association (ALSA) tested more than 1,000 available compounds in 29 different assays to determine activity against a variety of different aspects of neurodegeneration (57). Although the full extent of this effort has not yet been published, individual laboratories have further investigated drugs identified by this screening program, with the first of these (ceftriaxone) having entered clinical trials in 2006. At this writing, there are at least nine different compounds either in human trials or about to be tested in humans. All of these compounds target different aspects of the neurodegenerative process.

Assuming that the maximum tolerated dose (MTD) has been established in experimental models, as well as a range of doses that show activity against the target disease mechanism, dose ranges in the initial studies on humans can be appropriately chosen. In ALS, this step has been problematic. For a number of compounds that have been previously studied in efficacy trials, the MTD has not been established. Thus, negative results have been reported for creatine and celecoxib, yet the lack of a known MTD leads to the question of whether higher doses of either creatine or celecoxib could have demonstrated efficacy (58,59). In other studies, attempts were made to study compounds at doses close to MTD, but lower doses were not studied as well. This may have contributed to the fact that patients treated with topiramate at 800 mg/day progressed faster than placebo patients, and may also account for similar results in the recently reported minocycline trial (60,61).

ALS presents some unusual challenges in terms of clinical trials. In diseases associated with markers of activity (i.e., CD4 counts or viral load in HIV), the effect of differing doses on these markers can be used to determine the dose choices for a phase III trial. In ALS, no such markers have been identified, so that attempts to gauge efficacy must be based on the outcomes typically employed in larger phase III trials. For this reason, the line between phase II and phase III trials is often blurred in ALS. Thus, decisions about dose are often made after phase II trials, so it is essential that multiple doses be evaluated. This has often not been done in phase II ALS trials, and when dose ranging is done, it is often inadequate. For example, topiramate was tested in a phase II study at a dose of 800 mg/day (62). Although there was no statistically significant effect on mortality, treated patients lost an average of 10 lb more weight than placebo-treated patients and performed more poorly on functional and respiratory measurements (62). From assessment of adverse events, it was clear that at this dose, topiramate was quite difficult to tolerate, and the results reported could easily have been a function of patient’s weight loss and other events.

Outcome Measures

Given that there are no tissue-based biomarkers currently existing to determine drug activity in ALS, clinical assessment of efficacy is based on measurement of a variety of aspects of disease. The gold standard outcome for ALS trials currently remains survival. Survival is obviously clinically meaningful and straightforward to measure. There are several reasons, however, why other measures are being sought and why many current trials use outcomes other than survival. First, survival can be manipulated by many interventions that do not clearly alter the progression of underlying disease. Good nutrition and early use of percutaneous endoscopic gastrostomy (PEG) clearly prolong life (63). Respiratory support with noninvasive positive pressure ventilation (NIPPV) has been less well studied, but likely also prolongs life (64,65). Beyond these clearly defined interventions, there is emerging evidence that patients cared for at multidisciplinary ALS clinics have prolonged survival as compared to community-based controls (66,67). As these interventions may not be applied uniformly across all sites in a clinical trial, conclusions based on survival may be confounded by these variables. Many trials stratify along certain treatment variables, but stratification can reduce the power of a trial to find a significant drug benefit.

PHARMACOLOGIC MANAGEMENT OF MND

Riluzole

Despite clinical use for more than 14 years, riluzole, a 2-amino-6-(trifluoromethoxy) benzothiozole, remains the only Food and Drug Administration (FDA)-approved medication proven to slow the progression of ALS. Pharmacological mechanisms of riluzole include interference with N-methyl-D-aspartate (NMDA) receptor–mediated responses, stabilization of the activated state of voltage-dependent sodium channels, inhibition of glutamate release from synaptic terminals, and activation of extracellular glutamate uptake. Riluzole has demonstrated neuroprotective effect in motor neuron cultures and SOD1G93A transgenic mice (68–79). A recent Cochrane Database Review concluded Riluzole 100 mg daily prolongs median survival by about 2 to 3 months based on analysis of four randomized controlled trials (75). Recent studies using large registries suggest a greater benefit, ranging from 4 to 20 months. Although American Academy of Neurology practice guideline recommends the use of Riluzole for nonventilated ALS patients, analysis of the
ALS CARE database found that 41% of the cohort was not prescribed this medication, largely due to the expense (80,81). The drug is generally well tolerated, with asthenia, nausea, and an increase in serum alanine aminotransferase the most common side effects (82). Liver function should be monitored during therapy.

Over the past 15 years, over 20 agents have been tested in Phase II and III clinical trials, and several additional trials are ongoing. None of these agents tested thus far has shown enough efficacy to warrant FDA approval for treatment of ALS. Some of these agents and trials are briefly discussed below.

**Experimental Therapies and Clinical Trials**

**Growth Factors**

Growth factors represent a large, heterogeneous group of endogenous polypeptides with varying physiological activity including cell signaling, cellular growth and differentiation, angiogenesis, regulation of inflammation, and antiapoptotic effect. Growth factor clinical trials to date have been disappointing. Treatment with subcutaneous recombinant human insulin-like growth factor-1 (rhIGF-1 or myotrophin) for 9 months slowed deterioration on the Appel ALS rating scale in a multicenter, North American trial but not in a similarly designed European study which may have been statistically underpowered (83,84). A 2007 Cochrane Database Review concluded that available data were insufficient to render definitive assessment of rhIGF-1 as a clinical therapy for treatment of ALS (85). A third phase III study is currently complete and results are pending. Neurotrophins, including brain-derived neurotrophic factor (BDNF), ciliary neurotrophic factor (CNTF), and glial cell line–derived neurotrophic factor (GDNF) and oral xaliproden which has neurotrophic-like activity, have failed to demonstrate benefit in human clinical trials (86–90). Vascular endothelial growth factor (VEGF), erythropoietin (EPO), and hepatocyte growth factor (HGF) slow motor neuron deterioration in vitro and prolong survival in transgenic ALS rodent models but, to date, no human trial data has been reported (91–100).

**Antiexcitotoxic and Glutamate Reducing**

The cephalosporin antibiotics were identified via an National Institutes of Health (NIH) sponsored high throughput drug screening program as potentially beneficial for slowing progression of ALS. Ceftriaxone increased both brain expression of astroglial glutamate transporter GLT1 and its biochemical and functional activity, and delayed loss of neurons and muscle strength, associated with increased mouse survival. Its CNS penetration and long half-life are well known, obviating the need for extensive safety trials (101). A large multicenter clinical trial is currently ongoing.

Memantine is a noncompetitive NMDA receptor antagonist. It has been shown to protect neurons against NMDA- or glutamate-induced toxicity in vitro. Treatment of SOD1G93A mice significantly delayed disease progression and increased lifespan (102). Safety studies are completed, and efficacy studies are enrolling patients (Clinicaltrials.gov NCT00409721 & NCT00353665).

**Antioxidants**

Overexpression of heat shock protein conferred protection from ischemic injury in mammalian brain (114). Arimoclomol, a coinducer of heat shock proteins, delayed progression of ALS in a mouse model (115). A dose ranging phase II study is completed and a phase III study is planned. Mitochondrial dysfunction in ALS may be aided by CoQ10 (116). In SOD1G93A
transgenic mice, low-dose coenzyme Q10 prolonged median survival by 4.4% (117). Phase II studies in humans suggested that doses as high as 2,700 mg/day had no greater efficacy than placebo (118–120).

A study of vitamin E versus placebo in patients on riluzole yielded no difference in survival, although there was a tendency to remain longer in the milder states of ALS with vitamin E (121). Five grams per day of vitamin E was well tolerated in a phase III trial. However, no significant difference in survival was noted between treatment groups (121,122).

**Immunomodulatory/Anti-inflammatory**
The anti-inflammatory Celecoxib, a selective cyclooxygenase-2 (Cox-2) inhibitor which demonstrated neuroprotective effects and prolonged longevity in SOD1G93A mice, failed to show benefit on strength measures or survival at 800 mg/day dosing in a large, randomized, placebo-controlled trial (123).

Creatine has neuroprotective effects through blockade of the mitochondrial membrane pore and as an antioxidant, but failed to show benefit in two large, randomized controlled trials (124,125).

Copolymer-1 (copaxone) induces a neuroprotective T cell-mediated response. ALS mice treated with copolymer-1 experienced delayed disease onset, improved motor function, and extended survival (126). Thalidomide and its analog lenalidomide, both inhibit the expression of TNF-α and other cytokines by destabilizing their mRNA, thus reducing inflammation. Treatment with either thalidomide or lenalidomide attenuated weight loss, enhanced motor performance, decreased motor neuron cell death, and significantly increased the life span in G93A transgenic ALS mice (127).

**Neurodegeneration**
Lithium prevents neurodegeneration by promoting autophagy, through inhibition of inositol monophosphatase, rescues spinal cord mitochondria, and facilitates the clearance of α-synuclein, ubiquitin and SOD1. It delayed disease onset and progression in G93A transgenic mice and increased survival and slowed progression in humans over 15 months compared to controls in a small study with somewhat atypical slowly progressing patients (128). Larger trials are currently underway.

ALS progression is complex and likely due to cellular dysfunction at multiple levels, including mitochondrial dysfunction, glutamate excitotoxicity, oxidative stress, axonal dysfunction, reactive astrogliosis, and mutant SOD1 expression, therefore, treatment must provide neuronal protection from multiple insults. A significant amount of ALS research focuses on growth factor–based therapies. Growth factors including insulin-like growth factor-I, VEGF, BDNF, and GDNF exhibit robust neuroprotective effects on motor neurons in ALS models. Issues concerning growth factor delivery, stability, and unwanted side effects slow the transfer of these treatments to human ALS patients. Stem cells represent a new therapeutic approach offering both cellular replacement and trophic support for the existing population. Combination therapy consisting of stem cells expressing beneficial growth factors may provide a comprehensive treatment for ALS and there are a number of animal studies which support this (129–137).

**Innovative Approaches**
Extensive ineffective clinical trials in ALS involving various subcutaneously and orally administered medications have been disappointing, which has led to novel approaches to drug delivery and nonpharmacologic forms of treatment. Gene therapy approaches and RNA interference are being explored for patients with SOD1 mutations. There is also a great deal of interest in stem cell transplantation, but it is not yet available in scientifically sound human trials. In feasibility and safety studies of intraspinal cord implantation of autologous mesenchymal stem cells (MSCs), there were no convincing clinical benefits (138,139).

**Cannabinoids (Cannabis)**
Studies done in animal models of glutamate-induced neurotoxicity have shown that cannabinoids (the active ingredients in cannabis, aka, marijuana) afford protection against oxidative damage induced by free radicals produced by glutamate (140–148). Administration of delta-9-THC both before and after the onset of ALS symptoms slowed disease progression and prolonged survival in animals compared to untreated controls (149). Other trials in animal models of ALS have also shown that naturally occurring and synthetic cannabinoids slow down the progression of ALS (150–152). In addition to the neuroprotective effect, patients also report that cannabis helps in treating symptoms of the disease, including alleviating pain and muscle spams, improving appetite, diminishing depression, and helping to manage sialorrhea (excessive drooling) by drying up saliva in the mouth (153,154). Further investigation at both the basic science and clinical levels into the usefulness of cannabinoids in treating ALS is warranted.

**REHABILITATION AND PALLIATIVE CARE**
Although currently incurable, adult MNDs are not untreatable. The goals of rehabilitation and palliative care in these patients are to maximize functional capacities, prolong or maintain independent function and locomotion, inhibit or prevent physical deformity, and provide access to full community integration with good quality of life. In ALS, this also includes addressing end-of-life issues and assuring that the patient has a comfortable death.

The comprehensive management of all of the varied clinical problems associated with adult MNDs is an arduous task. For this reason, the multidisciplinary approach is much more effective and takes advantage of the expertise of many clinicians, rather than placing the burden on one. Management is best carried out by a team consisting of physicians, physical, occupational, and speech therapists, social workers, vocational counselors, and psychologists, among others. Ideally, due to the significant mobility problems associated with these diseases, the physician and all the key clinic personnel should be
The warning signs of overwork weakness, which include feeling participating in an exercise program should be cautioned of overwork weakness is great and exercise should be prescribed of ALS, and to a lesser extent in SMA and SBMA, the risk for due to the active, ongoing muscle degeneration in most cases overwork weakness in some of the subjects (157). However, to the moderate resistance program, and there was evidence of program showed no further added beneficial effect compared to the maximum weight a subject could lift 12 times) exercise the same population, a 12-week high-resistance (training at the maximum weight a subject could lift 12 times) exercise trials met the inclusion criteria. The first examined the effects of a twice-daily exercise program of moderate load, endurance exercise or excessive muscle soreness 24 to 48 hours following exercise. Other warning signs include severe muscle cramping, heaviness in the extremities, and prolonged shortness of breath (158).

Given the lack of any apparent contraindication, aerobic exercise training is recommended for patients with ALS as long as it can be performed safely without a risk of falling or injury. In addition to the physical benefits, this form of exercise often has a beneficial effect on mood, psychological well-being, appetite, and sleep. Pool therapy is often an ideal place for patients with ALS to do aerobic exercise, which can be as simple as walking in the water, with the water at midchest height. This is best done in a therapy pool with a flat, uniform depth floor that is heated to 92°F to 95°F. The warmth of the water will help reduce spasticity and facilitate movement. Other interventions might include low impact aerobic exercise such as walking, or stationary bicycling to improve cardiovascular performance and increase muscle efficiency, and thus help fight fatigue (158). Fatigue in ALS is multifactorial and is due, in part, to impaired muscular activation (159,160). Other contributing factors include generalized deconditioning from immobility and clinical depression. Aerobic exercise not only improves physical functioning but is beneficial in fighting depression and improving pain tolerance.

Although there have been few well-controlled studies looking at exercise-induced strength gains in the ALS population, a recent Cochrane review looked at this. Studies in randomized or quasi-randomized controlled trials of people with a diagnosis of definite, probable, probable with laboratory support, or possible ALS, as defined by the El Escorial criteria were recently published (161). This included progressive resistance or strengthening exercise and endurance or aerobic exercise. The control condition was no exercise or standard rehabilitation management and the primary outcome measure was improvement in functional ability, decrease in disability, or reduction in the rate of decline as measured by a validated outcome tool at 3 months. Secondary outcome measures were improvement in psychological status or quality of life, decrease in fatigue, increase in or reduction in the rate of decline of muscle strength (strengthening or resistance studies), increase in or reduction in the rate of decline of aerobic endurance (aerobic or endurance studies) at 3 months, and frequency of adverse effects. Two randomized controlled trials met the inclusion criteria. The first examined the effects of a twice-daily exercise program of moderate load, endurance exercise versus “usual activities” in 25 people with ALS (162). The second examined the effects of thrice weekly moderate load and moderate intensity resistance exercises compared to usual care (stretching exercises) in 27 people with ALS (163). After 3 months, when the results of the two trials were combined, there was a significant weighted mean improvement in the ALSFRS measure of function in the exercise compared with the control groups (3.21, 95% confidence interval 0.46 to 5.96) in favor of the exercise group. No statistically significant differences in quality of life, fatigue, or muscle strength were found.
Respiratory System Involvement in ALS and SMA

Neither ALS nor SMA has any direct effect on the lung. However, the mechanical respiratory system is significantly involved in all of these MNDs, although least often in later onset SMA III. For the purposes of this discussion, we focus primarily on ALS, although all of the modalities discussed herein are applicable to all of these disorders. ALS affects all of the major muscle groups of the mechanical respiratory system: (a) upper airway muscles (abnormal swallowing and cough); (b) expiratory muscles (inadequate cough) and inspiratory muscles (inadequate maintenance of ventilation). Therefore, all patients with ALS are at significant risk of respiratory complications, and the leading cause of death in this population is respiratory failure.

Cough is an essential airway protection reflex. Particles are expelled from the airway through a complex set of nerve and muscle responses to cough stimulation through receptors located predominantly in the upper airway (164–167). Cough receptor stimulation results in inhalation to approximately 60% of maximum vital capacity (inspiratory phase). The glottis then closes and the abdominal muscles contract, resulting in markedly elevated intrathoracic pressures without airflow (compressive phase). The glottis opens shortly thereafter and gas is propelled through the airways at very high velocities, resulting in airway clearance (expiratory phase). The individual with ALS may experience cough impairment in any one or all three of the stages of cough including reduction in the inspired volume due to diaphragm weakness, inability to close the glottis completely during the compressive phase due to bulbar muscle dysfunction, and inability to compress and expel intrathoracic gas because of expiratory muscle weakness. There appears to be a threshold of expiratory muscle strength loss necessary to reduce cough force. Maximal expiratory pressure (MEP), a commonly used clinical measure of expiratory muscle strength, does not correlate well with the presence or absence of cough generation (164). Inspiratory muscle strength also does not correlate well with cough generation (164). Endoscopic evaluation of ALS patients with respiratory symptoms reveals obvious glottic dysfunction (168). Some have suggested that measurement of peak cough flow (PCF) is an effective noninvasive assessment of cough function (169). In our experience, a measured PCF of less than 160 L/minute during acute respiratory illness and 270 L/minute in the absence of respiratory illness is associated with poor cough and a high risk of respiratory infection.

ALS often affects the inspiratory muscles including the diaphragm and external intercostal muscles. This leads to a reduction in respiratory muscle strength, RLD, and ultimately, carbon dioxide retention and frank respiratory failure. In rare cases, respiratory muscle dysfunction leading to respiratory failure may be the presenting clinical picture for the ALS patient. Usually, the symptoms of respiratory muscle insufficiency such as dyspnea occur gradually over time and may defy diagnosis. Pulmonary function testing is invaluable in assessing the level of respiratory impairment, following disease progression as well as assessing prognosis in ALS. Respiratory muscle strength measures do appear to correlate with dyspnea in ALS patients even when there is near normal vital capacity (169–171). In patients where oral bulbar weakness may limit the ability to accurately measure maximum inspiratory pressure by mouth, sniff nasal inspiratory pressure (SNIP) has been found to be a reliable alternative (171). Both maximum inspiratory pressure (MIP) and SNIP may be inaccurate measures of inspiratory muscle strength when significant bulbar weakness effects the test maneuver due to inadequate oral seal or upper airway collapse.

Nocturnal hypoventilation and sleep disordered breathing are a common problem for patients with ALS. This can occur even when respiratory muscle function is only mildly affected and daytime gas exchange remains normal (172,173). Neural output to the respiratory muscles decreases during sleep. Even mild muscle weakness coupled with the normal decreases in ventilatory drive can result in nocturnal hypoventilation and disturbed sleep architecture. Symptoms and signs of nocturnal hypoventilation can manifest both at night and during the day. Nighttime symptoms include air hunger, observed apneas, orthopnea, cyanosis, restlessness, nightmares, and insomnia. Daytime findings include excessive sleepiness, morning headaches or drowsiness, and later in the course, polycythemia and pulmonary hypertension. The health care provider should be vigilant for these symptoms. Sleep studies can be very helpful in elucidating sleep disturbed breathing in these patients if doubt remains.

Hypercarbia as well as atelectasis can lead to lower than expected arterial oxygen levels, but primary oxygenation problems are not common in ALS, except in the final stages when pneumonia intervenes. Oxygen as a primary therapy for respiratory insufficiency is not recommended. However, oxygen therapy in an attempt to relieve dyspnea in the hospice setting may be appropriate.

Managing Respiratory Problems in ALS

The respiratory specialist is integral to the care of the ALS patient because of the multiple effects of the disease on the respiratory system. Frequent monitoring of pulmonary function gives valuable information on prognosis as well as input into the timing of interventions and discussions of long-term mechanical ventilation. We have found that the most efficient way of delivering health care to these individuals at our institution is through a multidisciplinary clinic that involves care providers from rehabilitation medicine, neurology, pulmonary medicine, respiratory therapy, as well as speech-language pathology. The following discussion focuses on the monitoring and management of respiratory issues related to ALS, as outlined in Figure 28-6.

The risk of aspiration and development of pneumonia in patients with ALS is due primarily to problems with upper airway function and cough. Pharyngeal and laryngeal muscle dysfunction can lead directly to aspiration of oral contents into the lungs. Other than surgical diversion of the airway, no treatment directly aimed at the laryngeal and glottic function
is available. However, reduction of the amount of salivary secretions is possible through the use of a number of medications and modalities, including anticholinergic medicines and dietary manipulation such as avoiding dairy products which thicken the secretions. Teaching proper swallowing technique to avoid aspiration is also helpful. This involves performing a chin tuck and using straws, thickening liquids, concentrating on eating during mealtime (no TV, reading, etc.), and maintaining hydration with at least two quarts of water per day. In addition, when ALS patients develop significant dysphagia and aspiration with solids or liquids, many experts recommend early placement of a PEG tube (174). This may prevent large volume aspiration and is associated with improvement in hydration and nutritional status.

Cough function, which is dependent in larger part in expiratory muscles, can be mechanically assisted when adequate bulbar function exists. When PCF drops below 270 L/minute, we carefully evaluate potential interventions to improve cough function. Those interventions may include teaching the caregivers manually assisted cough and the Heimlich maneuver. Inspiratory-related cough weakness can be supported by the use of manual insufflation. A one-way valve mouthpiece
circuit combined with a self-inflating resuscitator bag can be used to insufflate the lungs by applying a series of breath stacking maneuvers. When bulbular function is good but the patient has significant expiratory muscle weakness, mechanical in- exsufflation (Cough Assist Device, Respironics Inc., Murrysville, PA) can be used to augment cough function. This device mimics the normal cough and has been shown to be helpful in patients with ALS and other neuromuscular diseases. Both mechanical pressure targeted insufflation and manual hyperinflation may also be beneficial to neuromuscular patients in maintaining lung compliance, decreasing the work of breathing, and managing atelectasis. Hyperinflation maneuvers can be administered by using either manual hyperinflation or by mechanical insufflation using the cough assist device.

Progressive inspiratory muscle weakness in ALS inevitably leads to carbon dioxide retention and hypercapnic respiratory failure. In fact, the major cause of death in ALS is respiratory failure. Predicting when respiratory failure will occur in the patient with ALS is important in order to plan appropriate clinical interventions and to help patients and their families address crucial decisions concerning long-term mechanical ventilation and end-of-life issues. Unfortunately, accurately predicting impending respiratory failure is a difficult task. Assessing symptoms of respiratory insufficiency such as dyspnea and orthopnea on each visit is important. Objective measurements of pulmonary function can be helpful but are not entirely predictive of either impending respiratory failure or death. We rarely obtain arterial blood gases as PaCO₂ can be maintained until immediately prior to respiratory failure. Although most authors agree that although it is impossible to accurately predict the lifespan of any given individual with ALS, severe restrictive disease with an forced vital capacity (FVC) of less than 50% should prompt careful discussions with the patient concerning medical interventions in the event of respiratory failure (164). The frequency at which FVC measurements should be taken has not been established but every 3 months appears to be a reasonable time frame. This measurement can easily be performed in clinic with portable spirometry equipment. FVC should be measured in both the seated and the supine positions as it is often lower supine, which contributes to nocturnal hypoventilation. Patients with oral muscle weakness may need to perform spirometry maneuver through an air cushion facemask in order to obtain reliable measurements. Bulbar-related upper airway obstruction that results in an uneven forced expiratory flow volume pattern can portend a poor prognosis related to airway protection, cough effectiveness, and tolerance of and benefit from noninvasive respiratory therapies.

Mechanical ventilator support with NIPPV has been shown to be effective in improving both quality and duration of life (165,172,175). Improvements in cognitive function have been shown in ALS patients receiving nocturnal NIPPV (176). Recently, a randomized controlled trial of noninvasive ventilation was done in a cohort of ALS patients, measuring both survival and quality of life (175). Ninety-two patients were assessed every 2 months and randomly assigned to noninvasive ventilation or standard care when either orthopnea developed with a maximum inspiratory pressure of less than 60% of predicted, or when symptomatic hypercarbia occurred. NIPPV improved quality of life and survival in all patients without bulbar symptoms as well as in a subset of patients with mild bulbar symptoms. In patients with more severe bulbar symptoms, NIPPV produced some improvement in quality of life but did not improve survival.

The Practice Parameters of the American Academy of Neurology suggests that all patients with ALS and respiratory symptoms or an FVC less than 50% predicted should be offered the use of NIPPV (177). NIPPV is usually initiated at night because of high frequency of sleep-disordered breathing. Sleep studies may be helpful if symptoms are unclear and FVC greater than 50%, although they are not necessary to initiate treatment. Patients will often progress to using NIPPV during the day, as their disease progresses, and we have had several patients who have used this modality for 24 hours/day. Portable daytime NIPPV can be most easily provided by using a bilevel pressure ventilator in conjunction with a less obtrusive interface, either a nasal cannula or nasal pillow interface. A small subset of patients with slow progressing limb onset disease and no bulbar symptoms may benefit from portable daytime mouthpiece ventilation (MPV). The benefits of MPV include progressive ventilatory support as well as sigh and cough augmentation by the use of breath stacking maneuvers. Unfortunately, NIPPV is only a temporizing measure. Most patients with ALS will at some point develop bulbar symptoms that are severe enough that they will be unable to continue use of NIPPV because of aspiration pneumonia or failure of NIPPV to ventilate the patient effectively despite 24 hour/day use. At this point, consideration for invasive ventilation may be the only option for continued survival.

Invasive ventilation involves placement of tracheostomy or laryngeal diversion for direct airway access and use of a small usually volume-cycled home ventilator. A laryngeal diversion (laryngotracheal separation) is an alternative procedure where the proximal trachea is either oversewn or hooked side on end into the esophagus and the distal trachea is brought out through a stoma in the neck. This procedure has the advantage of completely preventing aspiration although phonation will no longer be possible (167). Tracheostomy is clearly a life-prolonging intervention and patient survival may extend indefinitely. Unfortunately, invasive ventilation has no effect on the progression of the disease and patients can develop complete paresis of all muscles including the extraocular muscles and develop a “locked-in” syndrome in which no communication is possible. The costs to patients and families both in financial and emotional terms are significant (178). Family members provide much of the care for these patients at home and may have to relinquish employment outside the home to do so. Despite this, many patients report a good quality of life while receiving mechanical ventilation with ALS (164). Most patients who undergo invasive mechanical ventilation do so in
The setting of emergent hospitalization without having planned in advance for this eventuality. Openheimer reviewed experience with 50 ALS patients on invasive mechanical ventilation and found that only four (8%) of the patients had chosen tracheostomy in advance, prior to acute respiratory failure and emergent intubation (179). Few ALS patients have advanced directives or living wills in place at the time when respiratory failure occurs. Benditt et al. have developed an instructive advanced directive that we have found particularly helpful in assisting discussions of mechanical ventilation and end-of-life issues in patients with ALS (178). Few patients in the United States seem to choose invasive mechanical ventilation as an option. This may be due to a number of factors including patient expectation of reduced quality of life, physicians not offering this as a medical option, the exorbitant costs of mechanical ventilation (estimates of $15,000/month), and unavailability of family members to assist in patient care. The frequency of mechanical ventilation use in ALS varies worldwide from almost never (United Kingdom) to up to 48% (Japan) (180). This appears to be due to both cultural views of ALS as well as financial and structural differences in healthcare systems.

**Dysphagia and Nutrition**

Patients with ALS frequently develop bulbar muscle dysfunction due to motor neuron involvement in the brainstem. Dysfunction of the lips, tongue, and pharyngeal and laryngeal muscles can result in an increased risk of aspiration as well as difficulty with generating adequate glottic closure for effective cough function. Swallowing may be impaired and ingesting adequate nutrition can be trying for the patient and family alike. Choking episodes are common and may even be triggered by saliva. Secretion management is a particularly difficult issue, as secretions may become viscous due to inadequate hydration. Salivorrhrea (drooling) is due to inadequate handling of secretions rather than the amount of secretions as, in fact, salivary secretions in ALS appear to be less than in unaffected individuals (177). Malnutrition due to inadequate protein-calorie intake can occur and rapid weight loss should signal the clinician to carefully assess the swallowing mechanism (177). We have found referral to a speech and swallowing clinic to be very helpful in diagnosing swallowing and airway protection problems as well as instructing patients and their families in steps that they can take to reduce the risk of aspiration.

Most ALS patients without supplemental feeding through a percutaneous feeding tube will be deficient in energy intake (181,182). Both muscle and fat mass tend to decrease with disease progression. In addition, some patients become hypermetabolic; the etiology of this change in metabolism is not entirely clear, but it may be related to increased respiratory effort. Inadequate nutrition is concerning because it fosters greater muscle catabolism and may increase fatigue. Causes of malnutrition include arm weakness and inability to feed oneself, the time required to eat, and depression, in addition to bulbar muscle dysfunction and dysphagia. In a prospective study of French patients with ALS, those who were malnourished had a 7.7-fold increased risk of death (183). An algorithm for the management of dysphagia is shown in Figure 28-7. Dysphagia can initially be managed by instruction in compensatory strategies such as a double swallow, chin tuck, and head turning. Dietary modification such as thickening of liquids, moistening solids, and modifying temperature and texture may be helpful. Patients can often maintain nutrition orally by adding high-calorie liquid supplements to their diet.

When nutrition cannot be maintained orally, either partial or complete nutritional support may be provided via a feeding tube. In the past, most centers used PEG tubes, but many centers are now using radiologically inserted gastrostomy (RIG) tubes. RIG tubes seem to be easier and safer to insert in patients with low vital capacities (181,182). Indications for PEG or RIG include aspiration pneumonia, loss of greater than 10% of body weight, and/or the impairment of quality of life due to the time required to maintain nutrition orally. Feeding tube placement stabilizes weight and prolongs survival (177).

**Dysarthria**

Dysarthria in ALS does not respond well to conventional articulation training. However, some adaptive strategies such as maintaining a slow speaking rate with an emphasis on increasing the precision of speech production may be helpful and can be taught by a speech-language pathologist. As the disease progresses, dysarthria should be approached by prescribing communicative aids, rather than traditional ongoing speech therapy. An alphabet supplementation or word board works well early on when patients still have reasonable arm function. After that, developing yes/no or other binary commands with eye-gaze systems may be used, particularly if the patient is using mechanical ventilation. There have been major recent advancements made in devices such as speech synthesizers or multipurpose, multiaccess, computer-based augmentative communication systems. Although expensive, these devices greatly enhance the patient’s ability to communicate when they can no longer phonate. These types of devices may often be borrowed or rented from assistive technology centers which are often found at tertiary care medical centers.

**Spasticity**

Spasticity in ALS is probably induced both at the motor cortex and at the spinal cord level. The gamma-aminobutyric acid (GABA) analog baclofen acts to facilitate motor neuron inhibition at spinal levels and is the agent of choice. Initial doses are 5 to 10 mg two to three times a day, titrating up to doses of 20 mg four times a day. Occasionally, higher doses (up to 160 mg/day) are more effective but caution is advised. Side effects include weakness, fatigue, and sedation. An intrathecal baclofen pump may be beneficial to some patients with MND. Tizanidine, an alpha-two agonist similar to clonidine, inhibits excitatory interneurons and may also be helpful. Dosing range is 2 to 8 mg three to four times a day, with a
similar side effect profile to baclofen. Benzodiazepines may also be helpful but can cause respiratory depression and somnolence. Dantrolene blocks Ca\(^{2+}\) release in the sarcoplasmic reticulum; is effective at reducing muscle tone but may also cause generalized muscle weakness and is not recommended. Slow (30 second sustained), static muscle stretching may be helpful, particularly in the more symptomatic muscle groups such as the gastrocnemius and may be done in bed. Positional splinting is also a helpful adjunctive modality, but skin must be monitored frequently for pressure areas. Frequent turning is critical in patients on full-time mechanical ventilation, often done by family members, particularly in cultures outside the United States (184).

**Depression**

Reactive clinical depression is expected in ALS (185). Good family, social, and religious support systems, as well as participation in a support groups, are all helpful. Once the diagnosis is confirmed, the patient should be counseled with respect to the prognosis. This allows time for grieving, anger, and ultimately, acceptance of their fate and is important for the mental well-being of patients and their family (186,187). Antidepressant medicine should be offered to every patient, since it may provide assistance with mood elevation, appetite stimulation, and sleep. As previously mentioned, serotonin-mediated tricyclic antidepressants with significant anticholinergic activity will also help dry up oral secretions and minimize drooling. Further, they may help control the symptoms of pseudobulbar affect in ALS. Amitriptyline, starting at a dose of 25 mg 1 hour before bed, is a good choice. Families of ALS patients with emotional lability should be reassured that the underlying mood state may be normal. Rather it is their emotional expression that is abnormal. Nonetheless, referral to a psychiatrist or clinical psychologist with experience in treating depression associated with terminal disease may be required. Depression in the spouse or significant other, family, or friends should not be overlooked, and group/family counseling may be helpful.

**Pain and Immobility**

Although not frequently characterized as a major component of ALS, the majority of patients do experience significant pain (187). The pain is due largely to immobility, which can cause adhesive capsulitis, mechanical back pain, pressure areas on the skin, and more rarely, neuropathic pain (187). Frequently severe weakness in the neck flexors and extensors will cause a “floppy head” associated with severe neck pain and tightness.
This may be helped by a soft cervical collar, a Freeman or Headmaster type collar, which is a wire frame collar with padding over the pressure points. Wheelchairs should have adequate lumbar support and good cushioning (gelfoam). The chair should be properly fitted to avoid pressure ulcers and inadequate support for the spine. Simply giving the patient a prescription for a wheelchair often results in the patient receiving a standard manual chair that does not fit properly; therefore, a team approach with physical and/or occupational therapists and the durable medical equipment (DME) provider is a much more beneficial approach. A power wheelchair, although expensive, can be justified since it will help prolong independent mobility and thus markedly improve quality of life. A good pressure relieving mattress (air or dense foam) should be used on the bed at home, along with foam wedges to facilitate proper positioning. This will help prevent pressure ulcers and contractures. Daily passive and active-assisted range of motion is critical. Maintaining mobility and functional independence as long as possible will have positive physical as well as psychological benefits. Ankle-foot orthoses molded in the neutral position may prolong ambulation and decrease the risk of injury if there is unilateral or bilateral foot drop. Wheeled walkers (rollator style in particular) or quad (four point) canes may also help, depending on the pattern of weakness. Other useful equipment includes hand-held showers, bath tub benches, grab bars, raised toilet seat, hospital bed, commode chair, activity of daily living (ADL) aids (sock aid, grabbers, etc.), and wheel chair ramps. An occupational therapist will help define which, if any, of these devices will be useful to the patient. Other simple suggestions such as moving the patient’s bedroom to the first floor, removing any loose rugs, or covering slippery floors, are helpful and can be made during an in-home evaluation by the therapist.

Pharmacological management of pain in ALS includes the use of nonsteroidal anti-inflammatory (NSAID) medication, particularly if there is evidence of an active inflammatory process such as tenosynovitis or arthritis. Regular dosing of acetaminophen (1,000 mg every 6 hours) may be used along with an NSAID or alone if NSAIDs are not tolerated. Tricyclic antidepressants and anticonvulsant (membrane stabilizing) drugs such as Lyrica (pregabalin) or Neurontin (gabapentin) can sometimes be helpful for pain, particularly if there is a neuropathic component. Neurontin also has the potential added benefit of antispasticity properties. Narcotic medicine should be reserved for refractory pain. Concern for narcotic addiction is pointless in a terminal disease and the medications should be given on a regular dosing schedule and titrated to the point of comfort (182). Concomitant use of the antiemetic, antihistamine hydroxyzine (Vistaril) given along with the narcotic will enhance the effectiveness (i.e., 30 mg codeine plus 50 mg hydroxyzine every 6 hours p.r.n or scheduled depending on pain severity). Unlike narcotic medications, hydroxyzine is not a cortical depressant but does have direct skeletal muscle relaxant and analgesic properties and is known to potentiate the analgesic effect of narcotic medication, although the exact mechanism is unknown (188). Combination elixirs can be prepared by the pharmacy for ease of administration. Oral or sublingual morphine (Roxanol), 10 to 30 mg every 4 hours, is also effective for comfort care and may help relieve “air hunger” in the terminal stages of the disease. Another option is taking the total dose of immediate release morphine required to alleviate pain and giving half of that every 12 hours in a controlled-release preparation such as MS Contin. Intramuscular delivery route should be avoided due to muscle wasting. Fentanyl or morphine patches may deliver inconsistent dosing, particularly if there is excessive perspiration. A patient-controlled analgesic (PCA) pump mechanism may not work in advanced stages of ALS due to inability of the patient to control the delivery. The main problems with narcotic medication in ALS are respiratory depression and constipation. These side effects may be quite acceptable in the final phases of life when respiratory insufficiency or severe pain requires increased doses of narcotics and/or benzodiazepines. In states where it is legal to do so, cannabis (marijuana) should also be considered as it has been shown to be effective in managing many of the symptoms of ALS, including pain, muscles spasms, loss of appetite, secretion control, and depression (153,154). Cannabis has an advantage of opiates in that it does not cause respiratory depression, constipation, or loss of appetite (154).

End-of-Life Issues
ALS is a disease that poses some unusual ethical and humanitarian considerations. Although it is considered a fatal condition, unlike most cancers or other grave, incurable illnesses, it may take years to die from it, even though the disease continues to debilitate the person in the process. Thus, the patient with ALS has much time to think about the inevitability of the disease and what choices he or she wants to make in the terminal stages of the disease.

Despite the most aggressive treatment available, ALS will progress. Early in the disease, a social worker should be consulted to help arrange durable power of attorney to a responsible family member, usually the spouse. In most states, this can be done by a paralegal for a nominal fee. A Living Will should then be drafted which clearly outlines the patient’s wishes regarding extent of medical intervention desired (178). This is particularly important with respect to entering hospice level care. Presumably, by the time hospice level care is being considered, the patient has had ample time for grieving, anger, and ultimately, acceptance of his or her fate. However, in our experience, many ALS patients still are hesitant about enrolling in hospice since it implies that the disease has reached end stage (189,190).

The patient should also be referred to a support group. The most prominent consumer-driven organizations facilitating support groups for people with ALS are the MDA and the ALSA. Local branches may be contacted to locate the nearest ALS support group. Support groups are often a great resource, not only for psychological support but for problem solving and equipment recycling.
The majority of this section of this chapter has centered around comprehensive clinical care of the patient with adult MND, emphasizing “everything that modern medicine has to offer.” The physician must consider that the patient with advanced ALS may not want all of this. Life sustaining therapy, defined as any artificial device or intervention that compensates for the failure of an organ system which would normally result in death, is the patient's choice, not the physician's. The most obvious example of this would be mechanical ventilation, but this also includes artificial hydration and nutrition. Legally and ethically speaking, a mentally competent patient can refuse any prescribed treatment. It is the physician’s responsibility to make sure that the patient understands the consequences of this. The physician should always respect and foster the

HEALTH-CARE DIRECTIVE FOR THE INDIVIDUAL WITH AMYOTROPHIC LATERAL SCLEROSIS.

Recognizing that respiratory failure and/or inability to take in adequate nutrition is frequently the cause of death in cases of ALS, I hereby wish to state in advance my preference regarding invasive mechanical ventilation and feeding gastrostomy tube placement. It is my desire that these preferences guide the decision making of my family and my physician(s) in the event that I am unable to participate in a meaningful way in discussions regarding my health care. I understand that none of the choices made here will be put into effect without my agreement as long as I retain the capacity for decision making and the ability to communicate, in some form, those decisions.

I. With regard to invasive mechanical ventilation requiring endotracheal intubation, it is my preference that:
(Choose one of the following three main options):

A. ___ Invasive mechanical ventilation not be instituted under any circumstances. I understand that such a choice will almost certainly mean that my death will occur earlier than if such support is instituted. I also understand that some processes that might precipitate respiratory failure may be readily reversible and that, therefore, mechanical ventilation may not necessarily be long-term, yet I still do not wish to undergo mechanical ventilation even in such circumstances.

B. ___ Invasive mechanical ventilation be used only when, in the judgment of appropriate medical personnel, the acute cause of respiratory failure is believed to be likely reversible, for example, in the case of choking. If, on the other hand, respiratory failure is a result of the irreversible deterioration from ALS, I do not wish to undergo mechanical ventilation, knowing that such a choice will almost certainly mean that my death will occur earlier than if such support is instituted.

If invasive mechanical ventilation is used and it becomes evident that long-term mechanical ventilation is required, then (choose none, one, or more of the following):

1. ___ I wish for mechanical ventilation to be discontinued regardless of the circumstances, knowing that this will result in my death.
2. ___ I wish for mechanical ventilation to be discontinued if I should be diagnosed in writing by two physicians to be in a permanent unconscious condition.
3. ___ I wish for mechanical ventilation to be discontinued if I become permanently unable to effectively communicate (“locked-in”).
4. ___ I wish for mechanical ventilation to be discontinued if I am unable to return to living at home.
5. ___ I wish for mechanical ventilation to be discontinued if my care results in major financial hardship or other burden on my family.

C. ___ Invasive mechanical ventilation should be instituted in all circumstances for respiratory failure not treatable by other measures, and long-term mechanical ventilation with tracheostomy should be continued with the following exceptions (choose none, one, or more of the following):

1. ___ I wish for mechanical ventilation to be discontinued if I should be diagnosed in writing by two physicians to be in a permanent unconscious condition.
2. ___ I wish for mechanical ventilation to be discontinued if I become permanently unable to effectively communicate (“locked-in”).
3. ___ I wish for mechanical ventilation to be discontinued if I am unable to return to living at home.
4. ___ I wish for mechanical ventilation to be discontinued if my care results in major financial hardship or other burden on my family.

II. With regard to nutrition provided by feeding percutaneous gastrostomy tube (PEG), it is my preference that:
(Choose one of the following two main options):

A. ___ I do not wish placement of a feeding gastrostomy tube at any time during the course of my illness.

B. ___ I wish for placement of a feeding gastrostomy tube at a time when it is necessary to provide me with nutrition and medications, as determined by my physician, regardless of my choice concerning invasive ventilation. It should be continued with the following exceptions (choose none, one, or more of the following):

1. ___ I wish for gastrostomy tube feeding to be discontinued regardless of the circumstances, knowing that this will result in my death.
2. ___ I wish for gastrostomy tube feeding to be discontinued if I should be diagnosed in writing by two physicians to be in a permanent unconscious condition.
3. ___ I wish for gastrostomy tube feeding to be discontinued if I become permanently unable to effectively communicate (“locked-in”).
4. ___ I wish for gastrostomy tube feeding to be discontinued if I am unable to return to living at home.
5. ___ I wish for gastrostomy tube feeding to be discontinued if my care results in major financial hardship or other burden on my family.

In all cases where I choose not to start or to discontinue mechanical ventilation or nutrition via gastrostomy tube, I instruct my physician to provide me with adequate medication to relieve anxiety and discomfort that may occur during the final course of my disease.

FIGURE 28-8. Example of an advanced health care directive for patients with ALS (178).
patient’s autonomy and self-direction with respect to these types of interventions (191–193). An advanced health care directive is absolutely necessary. An excellent example of this was developed by Benditt et al., and is shown in Figure 28-8 (178). This does not extend to the point of physician-assisted suicide, where the physician takes active steps to end the patient’s life (194). This is an illegal act that carries ethical concerns which are beyond the scope of this chapter. Despite this, a recent study documented that approximately 56% of ALS patients surveyed in Washington and Oregon states would consider assisted suicide (194). This stunningly high percentage of ALS patients who would consider this strongly implies that the quality of care in advanced ALS is inadequate. If the patient is requesting this, then the physician should reassess the situation, making sure that everything has been done to maximize patient comfort and quality of life.

Further, quality of life studies have identified a lack of adequate communication between physician and patient and a poor perception (both positive and negative) on the part of physician of the level of quality of life in these patients (187). It takes a great deal of time to explain all of the end-of-life issues, including the available treatment options and choices. Without this investment of time on the clinician’s part, the patient is unaware what services may be available to ease his suffering.

The most appropriate level of care for the ALS patient may change frequently and patients should be followed closely. Unfortunately, an advanced ALS patient is often told “there is nothing that can be done,” when in fact optimizing in-home care with hospice can maximize quality of life for these patients and provide for a comfortable, painless passing. Krivickas et al. documented that most ALS patients probably do not receive enough in-home care (195). Of 98 advanced ALS patients studied, only nine received hospice home care, 24 received nonhospice home care, and seven both hospice and nonhospice home care. The remaining 58 patients received no in-home care at all. Even among those having home care assistance, non-medical ALS primary caregivers spent an average of 11 hours/day caring for patients with tracheostomy tubes; among ALS primary caregivers of patients with tracheostomy, 42% and 48% felt physically and psychologically unwell, respectively. The authors concluded that home and hospice care received by ALS patients is inadequate because it starts too late to relieve the burden placed on family caregivers. Because the focus of care in hospice is the family, however defined by the patient, this problem could be easily resolved. Hospice provides an interdisciplinary team of professionals whose mission is to support the patient and the family through their remaining days together. Support is given for physical, psychological, emotional, and spiritual needs of the family unit in the home setting, bypassing the need for laborious trips to clinics.

The National Hospice Organization does have some guidelines for entry of an ALS patient into hospice (Table 28-1), which are somewhat arduous but would allow for early entry into hospice of most ALS patients in the advanced stages of the disease (196). These guidelines require physicians to make some estimate of life expectancy, which is very difficult to do in ALS and is something for which most physicians are probably ill-prepared. Compared to terminal cancer patients, ALS has a relatively slow progression with respect to the actual dying process, which decreases the clinician’s awareness that hospice care may be appropriate. Most clinicians likely perceive that hospice is for “near terminal” patients, which is correct, except

<table>
<thead>
<tr>
<th>TABLE 28.1 Criteria for Hospice Admission in ALS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hospice is appropriate when there has been an overall rapid progression of ALS (a critical factor), e.g., disability has progressed significantly in the past 12 months. The patient/family desires no further aggressive treatment or cardiopulmonary resuscitation.</td>
</tr>
<tr>
<td>In addition, at least one of the following must also apply:</td>
</tr>
<tr>
<td>1. Increased respiratory distress</td>
</tr>
<tr>
<td>a. Vital capacity (VC) less than 30% of predicted</td>
</tr>
<tr>
<td>b. Significant dyspnea at rest</td>
</tr>
<tr>
<td>c. Supplemental oxygen required at rest</td>
</tr>
<tr>
<td>d. Patient has refused intubation, tracheostomy, and mechanical ventilation</td>
</tr>
<tr>
<td>2. Severely impaired nutrition</td>
</tr>
<tr>
<td>a. Tube feeding not elected or discontinued</td>
</tr>
<tr>
<td>b. Oral intake insufficient/dysphagia</td>
</tr>
<tr>
<td>c. Continued weight loss in spite of tube feedings</td>
</tr>
<tr>
<td>d. Dehydration or hypovolemia</td>
</tr>
<tr>
<td>3. Life-threatening complications</td>
</tr>
<tr>
<td>a. Recurrent aspiration pneumonia</td>
</tr>
<tr>
<td>b. Decubitus ulcers, multiple, stage 3–4, particularly if infected</td>
</tr>
<tr>
<td>c. Upper urinary tract infection, e.g., pyelonephritis</td>
</tr>
<tr>
<td>d. Sepsis</td>
</tr>
<tr>
<td>e. Fever recurrent after antibiotics</td>
</tr>
</tbody>
</table>
that ALS patients may be in that state for a prolonged period of time. During this time, hospice care could ease suffering considerably. Lack of physician knowledge of the services provided by hospice is widespread (197,198). Physicians not familiar with the care of terminal patients may not be comfortable with the aggressive use of opiates and benzodiazepines advocated by hospice clinicians for the control of symptoms in ALS. The physician may find it difficult to give carte blanche orders for effective titration of these types of medications, which will ease air hunger and anxiety in the end-stage ALS patient (Table 28-1).

Irrespective, in the final stages of the disease it is medically appropriate to involve a home hospice team. Regular home visits by hospice nurses will assure proper medication delivery, pain control, skin and bowel care, as well as providing the physician with a progress report without having to bring the patient into the clinic. They can also provide counseling to avoid panic calls to 911 by family members and unnecessary nighttime visits to the emergency room. Most patients wish to die at home, and in most cases with a supportive family and the help of hospice, this is a feasible and worthwhile goal.

An informed patient and family will welcome the comprehensive level of terminal care that hospice offers, consloed with the knowledge that dying with dignity in the serenity and security of one's own home is, in some modest but meaningful way, a measure of victory over this otherwise insufferable illness. Finally, we urge clinicians to attend memorial services since it is a healing and rewarding experience.

REFERENCES

1. Rowland LP. What’s in a name? amyotrophic lateral sclerosis, motor neuron disease, and allelic heterogeneity. 
5. Siddique T, Deng HX. Genetics of ALS. 
7. Hosler BA, Brown RH. Copper/zinc superoxide dismutase mutations and free radical damage in amyotrophic lateral sclerosis. 
9. Canu W, Billiard M, Baldy-Mouliner M. Fasting plasma and CSF amino acid levels in ALS. 
10. Rothstein JD, Martin JI, Kuncl RW. Decreased glutamate transport by the brain and spinal cord in amyotrophic lateral sclerosis. 
17. Eisen A. Amyotrophic lateral sclerosis is a multifactorial disease. 
27. Fishbeck KH, Ionasecu V, Ritter AW. Localization of the gene for spinal muscular atrophy: a linkage analysis-based assessment. 
31. Carter GT, Bonekate HW, Millio L. Successful pregnancies in the presence of spinal muscular atrophy: two case reports. 
33. Wirth B, Brichra L, Schrank B, et al. Mildly affected patients with spinal muscular atrophy are partially protected by an increased SMN2 copy number. 


PART III | MAJOR CONDITIONS


Peripheral Neuropathy

INCIDENCE, PREVALENCE, DISABILITY, AND COST

The overall incidence, prevalence, disability rates, and cost of peripheral neuropathy in the United States are unknown. There are studies, however, that hint at the scope of the problem. For diabetic peripheral neuropathy alone the estimated prevalence is more than 20 million, with an annual cost of 10.9 billion dollars (1). The annual cost of patient care and disability payments for acute idiopathic demyelinating polyneuropathy (Guillain-Barre syndrome) is approximately $1.7 billion. An estimated 11% of patients have permanent disability (2). Of patients with vasculitic disease, approximately 60% to 70% suffer from peripheral neuropathy, and of these, 65% have mild to moderate disability, 13% moderately severe disability, and 4% severe disability (3).

PERIPHERAL NEUROANATOMY AND NEUROPHYSIOLOGY

The peripheral nerve is vulnerable to a wide variety of insults but has a great capacity for repair and regeneration. The peripheral nerve includes the cell body, axons and dendrites, the cell membrane (neurolemma), the endoneurium, perineurium, mesoneurium, epineurium, and the Schwann cell. The specifics of these structures help understand the classification, pathophysiology, and treatment of peripheral neuropathy.

The cell body of motor nerve fibers is the anterior horn cell. The cell body of sensory nerves is located in the dorsal root ganglion. The autonomic nerve fibers in contrast to the somatic fibers have both preganglionic and postganglionic neurons, with the cell body of the postganglionic neuron lying in the periphery and extending through unmyelinated C fibers.

The axon is contained by the cell membrane, the axolemma. It contains the axoplasm which has tracts for both antegrade and retrograde flow, with the tracts maintained by electrical polarity. Axoplasmic flow occurs at 1 to 3 mm per day and is one rate-limiting aspect of nerve regeneration, although the flow can occur more rapidly in response to injury.

The axolemma of myelinated nerve fibers is enclosed by the Schwann cell, which elaborates the lipoprotein myelin. This is why lipid lowering medications are a risk factor for peripheral neuropathy (4).

The Schwann cell internode space, or node of Ranvier, is the site of membrane depolarization, saltatory conduction, and any axonal branching. In unmyelinated nerve fibers, the relationship of nerve fibers is less complex, and several fibers may be contained in the Schwann cell trough. These fibers propagate signal conduction by continuous depolarization along the axolemma (eddy depolarization).

Axon fascicles divide and fuse with others in the epineurium. Communication between fascicles occurs every 0.5 to 15 mm, more frequently in the proximal peripheral nerve. They branch between 20 and 100 times before reaching the motor endplate.

The mesoneurium supports the capillary network that supplies the nerve fibers. The mesoneurium is easily compromised and this accounts for the peripheral nerve's susceptibility to ischemia. The endoneurium contains the axon and Schwann cell in a grouped arrangement. These fascicles are in turn contained by the perineurium which maintains a positive intrafascicular pressure. This structure presents a barrier to diffusion, the blood-nerve barrier. When the perineurium is compromised, diffusion produces axonal swelling which impairs signal conduction (Table 29-1). The epineurium is a loose collection of collagen and elastin fibers that support the fascicles of a peripheral nerve. This arrangement is elaborate in the proximal nerve and becomes progressively less complicated in more distal nerve segments. The external epineurium surrounds the peripheral nerve and is largely responsible for its resistance to mechanical disruption. Its elastic properties allow a degree of deformation beyond which rupture occurs. This is why a stretched nerve is more readily injured than when at its resting length.

A number of age-related physiologic changes occur in the peripheral nervous system (5). There is a decrease in the number of anterior horn cells, a decreased capacity for neuronal sprouting, biosynthesis, transport and proliferation, and decreased Schwann cell synthesis of trophic factors. This results in slowed protective reflexes and decreased proprioception, vibratory sense, and stretch reflexes. The pain and temperature recognition thresholds increase.

CLASSIFICATION OF PERIPHERAL NEUROPATHY

Peripheral neuropathy can be classified by etiology, pathology (including genetic error), location of the lesion, time since insult,
or clinical presentation. It may be generalized, proximal, or distal in location. Possible etiologies include trauma, metabolic disease, malnutrition, response to infection, other autoimmune disease, collagen-vascular disease, genetic error, toxin (including medication) exposure, thermal injury, or ischemia (Table 29-2).

The neuropathy can affect the axon and/or the myelin sheath. The neuropathy can result in neurapraxia, axonotmesis, or neurotmesis. The disease may be confined to a single nerve (mononeuropathy), involve multiple nerves (polyneuropathy), and be symmetric or asymmetric. The neuropathy can be acute or chronic although this may be unclear when there is an insidious onset and delayed diagnosis. Depending upon the nerve involved, the clinical presentation includes weakness, paresthesia, hypesthesia, anesthesia, or changes in autonomic function such as changes in circulation or hidrosis.

The Seddon classification of peripheral nerve pathology is clinically relevant. It can be used to predict functional outcome and suggest appropriate care (Table 29-3). There are three degrees of nerve pathology: neurapraxia, axonotmesis, and neurotmesis.

Neurapraxia is a local conduction block due to transient demyelination and rarely affects sensory or autonomic fibers. Thick myelinated nerves are most affected. Neurapraxia commonly results from compression of the peripheral nerve. These lesions heal by Schwann cell repair, and normal conduction is generally resumed in 1 to 2 months (6,7).

Axonotmesis is a more significant injury and results in Wallerian degeneration. With an axonotmetic lesion, the endoneurial tube remains intact. This lesion follows a traction injury or a severe nerve compression. The prognosis for regeneration is good, particularly for shorter injured or distal nerve segments. The distance of regeneration is the primary limiting factor, so recovery is less certain with proximal injury.

Neurotmesis is complete severance of the peripheral nerve trunk and is the most severe lesion. Recovery is unlikely unless neurorrhaphy is performed (6). Healing of a neurotmetic lesion often results in the misconnection of nerve fibers and incomplete reinnervation.

### TRAUMATIC PERIPHERAL NERVE INJURY AND SUBSEQUENT DEGENERATION AND REGENERATION

Traumatic peripheral nerve injury can result from compression, crush injury, laceration, stretch/traction, ischemia, thermal injury, or high-velocity trauma (Table 29-4). Further injury of the peripheral nerve may result from associated infection, scar tissue formation, fracture callus, or vasculopathy.
The most common mechanism of traumatic nerve injury is transection due to blunt or penetrating trauma. In such instances, there may be a delay in diagnosis due to adjacent tissue injury.

Fracture and fracture/dislocation carry a high risk of associated nerve damage. Nerve injury after shoulder dislocation occurs in 48% of cases. The incidence of radial nerve damage following humeral fracture is 11% (8). Ulnar neurapraxia is the most commonly identified nerve lesion associated with fracture dislocation at the elbow. Dislocation of the hip is associated with a nerve injury rate of 3% and the rate associated with knee dislocation is 18% (8).

Iatrogenic injury can also occur. Plating of forearm fractures results in a reported nerve injury in 1% to 10% of cases. Damage has also been reported during elbow and shoulder arthroscopy (9,10).

Compression injury of a peripheral nerve generally results in focal demyelination. This causes a conduction block. Recovery depends upon remyelination. With all peripheral nerve lesions that leave the axon intact, there is axonal transport of tumor necrosis factor alpha (TNF-α) to the lesion and a concomitant reorganization of peripheral nerve TNF receptors (11).

A crush injury provokes segmental demyelination but the Schwann cell tube is commonly preserved and recovery can occur. A laceration injury due to blunt or penetrating trauma produces a well-localized lesion, usually millimeters in size. A stretch of the peripheral nerve beyond 10% to 20% of its resting length increases the risk of axonotmesis (9). This is the common mechanism of injury during joint dislocations. Stretch alone may provoke a mild conduction block which recovers in hours. A more severe stretch will interrupt axons and connective tissue, cause hemorrhage and might require surgical repair.

Cold injury can cause necrosis of all tissues. Large myelinated fibers are most susceptible to cold injury. Damage to the blood-nerve barrier results in endoneurial edema and increased intraneurial pressure, with a resulting focal conduction block. If the pathology is progressive, axonal transport ceases and the axon degenerates within a few days.

Peripheral nerves vary in their vulnerability to compressive, thermal, or traumatic injury. Fiber-type composition, size of the nerve, number of nerve fascicles, amount of soft-tissue protective cushioning, course of the nerve (on bone, through fascia or muscle), and tethering all affect the nerves’ ability to sustain and spontaneously recover from injury. For example, the peroneal division of the sciatic nerve is tethered at the fibular neck and at the sciatic notch and is more vulnerable than the tibial division, which is tethered at the sciatic notch alone. Scar formation, heterotopic ossification, and fracture callus can tether the peripheral nerve.

**Degeneration**

Primary, or retrograde, degeneration is a consequence of trauma and is less common than secondary (Wallerian) degeneration. The degenerative process proceeds from the site of injury to the next proximal node of Ranvier (6).

Secondary or Wallerian degeneration is antegrade, progressing distally from the point of injury (7). Wallerian degeneration begins on the second or third day after injury, with retraction of myelin. Nerve fragmentation on day 2 to 3 precedes neurofibrillar degeneration. The nerve body swells. Neuron edema continues for 10 to 20 days. These changes are more pronounced and longer lasting with proximal nerve lesions. The Schwann cells at the site of injury activate and, by the end of the first week after injury, participate in the removal of myelin debris.

**Regeneration**

Axonal regeneration and remyelination progress in a sequence that follows degeneration of the injured nerve segment beginning with the activation of Schwann cells in the empty endoneurial tube. Axonal sprouts appear and progress down the endoneurial tube. These regenerating axons are guided along the perineurium by neutropins. They are directed toward the largest surviving distal fascicles (9).

The peripheral nerve can form a neuroma during the repair process. This may be a nerve stump neuroma (neuroma in continuity) which is usually located lateral to the nerve trunk. It forms as axonal/fascicular continuity is reestablished. A laterally located neuroma indicates partial neurotmesis with preserved ability to conduct signals. When the neuroma is imbedded in scar tissue, the prognosis for recovery is worse. A fusiform-shaped neuroma is likely to be in continuity, while a bulbous or dumbbell-shaped neuroma is indicative of widespread neurotmesis. This should be treated by excision and neurolysis. (Surgery following trauma is discussed in the treatment section of this chapter.) Similarly, if more than 50% of the nerve trunk is involved, function will be impaired and the neuroma should be resected (6).

Initially, unmyelinated axonal sprouts unite with the distal peripheral nerve remnant, then remyelination begins. Both sheath and axon increase in diameter. If the gap in continuity is greater than 2 mm, reconnection is much less likely. In such instances, the immature neurite (sprout) dies back or forms a neuroma.

Later, a shrinking area of sensory loss with an enlarging area of partial sensation occurs as anastomotic branches with other nerves form (12). Involvement of autonomic fibers causes anhidrosis and impaired pilomotor and vasomotor activity.

---

**TABLE 29.4 Traumatic Causes of Peripheral Neuropathy**

<table>
<thead>
<tr>
<th>Compression</th>
<th>Crush injury</th>
<th>Laceration</th>
<th>Stretch/traction</th>
<th>Thermal injury</th>
<th>High-velocity injury</th>
</tr>
</thead>
</table>

---

**References:**

(6)...

(11)...

(9,10)...

(7)...

(12)...

(8)...

(5)...

(4)...

(3)...

(2)...

(1)...

---

**Further Reading:**

For a comprehensive understanding of peripheral neuropathy, consult specific texts or resources that provide in-depth analysis and clinical guidance.
If the skin wrinkles on immersion in water, or if sweating is present, the peripheral nerve damage is incomplete.

As recovery progresses, there is a return of pain and temperature sensation, a return of sudomotor function, and later a return of light touch, vibratory sensation, and stereognosis. Perhaps the best predictor of outcome is two-point discrimination which positively correlates with a return of function (6).

A nerve percussion sign is indicative of demyelination/remyelination. When the nerve percussion sign progresses from a proximal location in the nerve to more distal segments, healing is taking place by regenerating sprouts and incomplete myelination (13).

### Nontraumatic Peripheral Neuropathies

The peripheral nerve axon is particularly vulnerable to toxic, metabolic, endocrine, and genetic insults. The myelin sheath is most commonly affected by autoimmune, nutritional, or genetic disorders and toxic or metabolic disease. In the United States, diabetes and alcohol abuse are the most frequent agents of peripheral nerve injury. Fifteen to twenty percent of nontraumatic peripheral nerve injuries are idiopathic (14).

Injury results in an increase in the neuropeptides galanin and pituitary adenylate cyclase–activating peptide with a net neurite outgrowth that can result in excessive branching during repair (15).

### Hereditary Peripheral Neuropathies

The hereditary peripheral neuropathies include the hereditary sensory motor neuropathies (HSMNs) types I and II (Charcot-Marie-Tooth [CMT] disease), type III (Dejerine-Sottas disease), type IV (Refsum's disease), and HSMN types V–VII (Table 29-5). Other inherited peripheral neuropathies include Friedreich's ataxia, pressure-sensitive hereditary neuropathy, and various diseases that include altered structure or function of the peripheral nerve such as acute intermittent porphyria, Roussy-Levy syndrome, Riley-Day syndrome, Fabry disease, and Pelizaeus-Merzbacher disease. These neuropathies demonstrate segmental demyelination and remyelination of the peripheral nerve, resulting in a slowing of signal conduction (16). Large myelinated motor fibers are the most severely affected (17). The distribution of weakness and atrophy includes the peroneal and distal leg muscles, and the peak strength loss is between 60% and 80% (18). Sensory loss and areflexia are notable. Atrophy and weakness in the upper extremities are less prominent.

Inheritance of a HSMN is usually autosomal dominant with variable penetration. Autosomal recessive and X-linked diseases occur less frequently and generally have a poorer prognosis.

The mutations in HSMN affect the genes that encode myelin proteins. Several abnormalities have been identified. These include duplication of chromosome 17p11.2, resulting in abnormalities of peripheral myelin protein 22 (19,20). The mutation produces abnormal endoplasmic proteins that lead to Schwann cell apoptosis (21). The specific of myelin gene mutation determine disease severity. Deletion of myelin protein zero results in the most severe disease. Deletion of the 17p11.2 chromosome results in a hereditary neuropathy with susceptibility to pressure palsy. Point mutation of the PO gene and defects of the connexin 32 gene (which encodes a gap junction protein) occur in X-linked forms of HSMN (22). In 30% of CMT type II disease, there is a defect in mitofusin 2 genes with a decoupling of mitochondria in the axon, leading to decreased oxidative phosphorylation (23).

CMT type I can also be defined by the genetic error. Type Ia CMT disease results from a defect in chromosome 17. In type Ib CMT disease, the defect is located in chromosome 1.

The prevalence of the most common neuropathy, HMSN type I and II CMT disease, ranges from 1 per 50,000 to 1 per 250,000 (24–27). The clinical manifestations are variable. The slowly progressive weakness is symmetric and more pronounced in distal musculature. In type I disease, the myelin is affected and onset is within the first decade of life. In type II disease, the axon is most affected and onset is usually within the second decade. For both type I and type II, the onset is often insidious and the patient asymptomatic until much later. The distribution of sensory deficit parallels that of the motor deficit. Loss of balance and tripping due to foot drop are often noted, as are deformities such as equinovarus, calcaneovalgus, and pes cavus. Pain is uncommon.

CMT type II inheritance is more heterogeneous than in type I, with wider phenotypic variation. The resulting disability ranges from very mild to severe. Type II disease is characterized by less hypertrophic change in myelin and more neuronal or axonal involvement than is seen in CMT type I (19,20,28–31).

HMSN type III, Dejerine-Sottas disease, is another inherited hypertrophic peripheral neuropathy with prominent sensory loss and remyelination. Neurapraxia is typical of this disease (29). Patients present with delay in motor

### TABLE 29.5 Hereditary Sensory Motor Neuropathies

<table>
<thead>
<tr>
<th>Type</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Charcot-Marie-Tooth (CMT) disease</td>
</tr>
<tr>
<td>II</td>
<td>CMT</td>
</tr>
<tr>
<td>III</td>
<td>Dejerine-Sottas disease</td>
</tr>
<tr>
<td>IV</td>
<td>Refsum's disease</td>
</tr>
<tr>
<td>V</td>
<td>With spastic paraplegia</td>
</tr>
<tr>
<td>VI</td>
<td>With optic atrophy</td>
</tr>
<tr>
<td>VII</td>
<td>Friedreich's ataxia, pressure-sensitive hereditary neuropathy, porphyria, Roussy-Levy syndrome, Riley-Day syndrome, Fabry disease, Pelizaeus-Merzbacher disease</td>
</tr>
<tr>
<td>Other inherited diseases associated with peripheral neuropathy</td>
<td></td>
</tr>
</tbody>
</table>
development, difficulty running and jumping, and weakness affecting the arms as well as legs.

Refsum’s disease (HSMN type IV) is characterized by altered mitochondria within the Schwann cell, and a similar abnormality is likely in other HSMN types (24). HSMN type V is associated with prominent spinocerebellar degeneration, type VI with optic atrophy, and type VII occurs with retinitis pigmentosa.

In the HSMNs, ambulation is frequently impaired and falls are common. Rehabilitation interventions focus on maintaining a safe and effective gait. Bracing, particularly ankle-foot orthoses (AFOs), usually provide adequate support. If contractures require surgical release, postoperative bracing or splinting is essential. Attention to footwear is important, particularly as equinus/cavus deformities typically occur. A comfortable, protective shoe with adequate depth and reinforced medial counter helps avoid pain, skin breakdown, and progressive deformity. Exercise is most effective for strengthening the proximal muscles of the lower extremities (32).

**IMMUNE-MEDIATED PERIPHERAL NEUROPATHIES**

In the peripheral nerve, immune-mediated demyelinating protein attacks myelin and produces vasculitis and ischemia (33). The inflammatory mechanism is both cellular and humoral (34). TNF-α modulates the immune response, particularly in T-cell–mediated tissue injury.

The acute demyelination characteristic of AIDP results from postinfection antibodies that recognize glycolipids and gangliosides GM1, GD12, and GD16. Lymphocytic infiltration of the spinal roots and peripheral nerves aids macrophages in myelin stripping.

Acute inflammatory demyelinating polyneuropathy (AIDP; Guillain–Barre syndrome) is most typically a postinfection demyelination of the peripheral nerve with both perineurial and axonal damage. There is a breakdown of the blood-nerve barrier and segmental, macrophage-mediated damage to the myelin sheath. Inflammation and demyelination result in varying degrees of axonal degeneration, and neuroparalytic. Sixty-seven percent of patients with AIDP have a history of preceding viral infection, immunization, surgery, or a disease affecting the immune system. AIDP presents with acute onset of weakness, hypotonia, and areflexia (35). The weakness is progressive and involves the extremities. Bulbar and facial muscles can be affected. Autonomic dysfunction and sensory symptoms are usually mild (34). Respiratory failure occurs in up to 30% of cases within 1 to 2 weeks after disease onset (36). Recovery generally takes 3 to 18 months. Residual weakness is common and usually mild.

The Miller Fisher syndrome is a relatively benign variant, occurring in about 5% of AIDP cases. It is characterized by ophthalmoplegia, areflexia, and ataxia (34). Antibodies to GQ1b are common.

Acute motor axonal neuropathy (AMAN) is an axonal variant that follows *C. jejuni* infection. Wallerian degeneration occurs. The clinical presentation is similar to that of AIDP but myelin is not affected. The antibody mediators include GM1, GD1a, and GD16 (34).

Acute motor sensory axonal neuropathy (AMSAN) is another clinical syndrome characterized by axonopathy. There is also a sensory variant of AIDP and, most rare, an acute panautonomia.

The medical management of AIDP includes the administration of high-dose immunoglobulins, plasmapheresis, or plasma exchange (37–39). Treatment reduces the duration of paralysis and intubation, particularly in the most severe cases (39). Cerebrospinal fluid filtration of antibody complexes can be useful. Corticosteroid therapy has no proven efficacy.

The rehabilitation management of AIDP focuses on the prevention of contractures, skin breakdown, pneumonia, and depression. During the acute phase, communication devices, a trapeze, pressure relief bed surfaces, and bed rails are helpful. Because AIDP presents with evolving weakness, strengthening, bracing, adaptive equipment, and vocational retraining are not appropriate until clinical findings have stabilized. Retraining for activities of daily living (ADLs), wheelchair and ambulation training, and bracing may be necessary if there is significant residual impairment and disability.

Chronic inflammatory demyelinating polyneuropathy (CIDP) is a T-cell–mediated autoimmune peripheral neuropathy. It involves motor and sensory fibers. Disability results from weakness of both proximal and distal muscles (40). Cramps and fasciculations are common in the upper extremities (41).

The differential diagnosis of CIDP includes HSMN and amyotrophic lateral sclerosis. Histologic changes characteristic of CIDP include mononuclear cell infiltrates, prominent endoneurial edema, and wide interfascicle variability. CIDP can be associated with malignancy, particularly melanoma, due to shared immunoreactivity with common surface antigens present in both the myelin and the tumor (42).

The medical management of CIDP includes high-dose intravenous immunoglobulins, immunosuppressive drugs, or immune adsorption (43). Treatment with steroids is probably not effective (44). There is evidence that CIDP responds well to stem cell therapy (45).

Other immune neuropathies include multifocal acquired demyelinating sensory and motor neuropathy (Lewis Sumner syndrome, MADSAM), distal acquired demyelinating symmetric neuropathy, and multifocal motor neuropathy (Table 29-6) (46–48). Other forms of the disease are the subacute inflammatory demyelinating polyneuropathy (SIDP) and mononclonal gammopathy of undetermined significance (MGUS) (49,50).

Critical illness polyneuropathy (CIP) can present as failure to wean from mechanical ventilation (50,51). In septic multisystem organ failure, the systemic inflammatory response syndrome or high fever leads to a polyneuropathy of mixed or
motor nerves (50,52–57). High-dose intravenous steroids have also been implicated in the production of CIP. Electrodiagnostic studies are key in the diagnosis and monitoring of CIP since the clinical findings of CIP can be obscured by the primary medical condition, steroids, neuromuscular blocking agents, or concomitant compression neuropathies. The differential diagnosis includes critical illness-associated transient neuro-muscular blockage, thick-filament myopathy, and necrotizing myopathy (52).

CIP may result from cytokine and free radical release during sepsis which impairs peripheral nerve microcirculation (58). Noninflammatory axonal degeneration and resulting neurogenic muscle atrophy occur in distal motor fibers, although there is often striking weakness of the proximal muscles and diaphragm (59). Sensory fibers are minimally affected (56,59).

CIP produces moderate to severe weakness in all limb, facial, and paraspinal muscles, and atrophy is prominent. Peroneal nerve residual weakness is most common. Loss of deep tendon reflexes is inconsistent. Serial serum creatine kinase levels and serial electrodiagnostic testing are helpful in monitoring the disease course.

Recovery time from CIP ranges from 3 weeks to 6 months (59). The prognosis is good if the patient survives the precipitating critical illness.

Rehabilitation efforts for CIP patients focus on preventing decubitus ulceration, contractures, and compression neuropathies. Strengthening exercise, mobility and ADL retraining, as well as appropriate orthotics and adaptive equipment, should be provided at appropriate stages of recovery.

Other diseases that alter peripheral nerve function include benign monoclonal gammopathy (IgG, IgA, or IgM), chronic liver or pulmonary disease, cryoglobulinemia, giant cell arteritis, gout, and necrotizing angiopathy (60). Paraneoplastic or paraproteinemic peripheral neuropathies occur in association with various malignancies such as lymphoma; multiple myeloma; bronchogenic carcinoma; ovarian, testicular, penile, gastric, oral cavity, and meningal cancers; oat cell carcinoma; and osteosclerotic myeloma.

**TABLE 29.6** Immune-Mediated Peripheral Neuropathies

|---|---|---|---|---|---|---|---|---|

**TABLE 29.7** Toxic, Nutritional, Metabolic and Endocrine Peripheral Neuropathies

<table>
<thead>
<tr>
<th>Etiology</th>
<th>Disease or Agent</th>
<th>Cause</th>
<th>Presentation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Toxic</td>
<td>Organic</td>
<td>Distal axon of longer peripheral nerves</td>
<td>Stocking-glove sensory loss followed by distal motor weakness</td>
</tr>
<tr>
<td>Nutritional</td>
<td>B vitamins deficiency or toxicity</td>
<td>Axon</td>
<td>Pain</td>
</tr>
<tr>
<td>Metabolic</td>
<td>Renal failure</td>
<td>Axon</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Porphyria</td>
<td>Axon</td>
<td></td>
</tr>
<tr>
<td>Endocrine</td>
<td>Diabetes</td>
<td>Axon</td>
<td>Variable</td>
</tr>
<tr>
<td></td>
<td>Thyroid</td>
<td>Myelin</td>
<td></td>
</tr>
</tbody>
</table>

**TOXIC, NUTRITIONAL, METABOLIC, AND ENDOCRINE PERIPHERAL NEUROPATHIES**

Toxic peripheral neuropathy results from exposure to a variety of organic and inorganic toxins, medications, and heavy metals (Table 29-7). Many toxic peripheral neuropathies resolve with appropriate treatment. Many commonly prescribed therapeutic drugs, environmental pollutants, industrial solvents, and other workplace chemicals can be neurotoxic. Peripheral neuropathy is one of the most common responses of the peripheral nerve to chemical attack (61). Most toxins produce distal axonal degeneration in the longer peripheral nerves. There are several toxic agents that damage the neuron directly or induce primary demyelination (61). Therapeutic drugs that can cause neuropathy include α interferon, amiodarone, amitriptyline, chloramphenicol, chloroquine, cimetidine, colchicine, corticosteroids, cyosplatin, dapsone, didanosine, diphenylhydantoin, disulfiram, ethambutol, hydralazine, isoniazid, lithium, metronidazole, nitrofurantoin, nitrous oxide, paclitaxel, phe-nytoin, pyridoxine, sodium cyanate, suramin, tetracycline, clindamycin, and vincristine (14). The pathogenesis differs between therapeutic agents. Cisplatin and suramin, for example, result in apoptosis of neurons in the dorsal root ganglion.
Nucleoside reverse transcriptase inhibitors (62), deplete peripheral nerve mitochondrial DNA, and DNA polymerase activity (63). They may respond to treatment with acetyl-L-carnitine for the treatment of pain and decreased sensory fiber function (62).

Neurotoxic organic compounds include acrylamide, carbon disulfide, dichlorophenoxyacetic acid, ethyl alcohol, ethylene oxide, methylbutyl ketone, and triorthocresyl phosphate. Heavy metals that may be neurotoxic include antimony, arsenic, gold, lead, mercury, and thallium.

Toxic peripheral neuropathies typically present with a glove and stocking distribution of sensory loss (64). This is followed by weakness in the same distribution. The presentation may be protracted and insidious as is common with nutritional or endocrine etiologies. Recovery can take several months to years following appropriate treatment (61).

Nutritional deficiencies (B1, thiamine [beriberi or pellagra], riboflavin [B2], pyridoxine [B6], B12 [pernicious anemia], and protein or calorie deficiency) and endocrine disease (diabetes, thyroid, or parathyroid) exposure can also damage the peripheral nerve axon.

### DIABETIC PERIPHERAL NEUROPATHY

Diabetic neuropathy affects between 5% and 50% of diabetics in the United States (65). The incidence of peripheral neuropathy increases with age, duration of diabetes and mean serum glucose (65,66), smoking, HTN, height, and hyperlipidemia. Average annual incidence is 2% of diabetics and 0.56% have severe symptoms and deficits. The prevalence among people with diabetes for longer than 25 years is 50%. Diabetic peripheral neuropathy increases with small neurotrophic nonneural peptides (“dual-action peptides”) (79).

In diabetic neuropathy, there is endoneurial and epineurial lymphocyte infiltration that disrupts signal conduction (68). The microvascular ischemia affecting the peripheral nerve is due to impaired endogenous fibrinolysis. The vascular changes are mediated by the polyol pathway—sorbitol and aldose reductase deposit glucose in tissues that are insulin independent (retina, kidney, nerve). Accumulation of sorbitol produces osmotic stress on the cell membrane while at the same time decreasing nitric oxide and vasodilation. The net result is microvascular disease due to vasoconstriction, thickening of the capillary basement membrane, endothelial hyperplasia, and a resulting decrease in oxygen tension and hypoxia.

There is an increasing refractory period before overt slowing of nerve conduction velocity is seen (69). Laser Doppler flowmetry can demonstrate sympathetic involvement when testing reveals an abnormal vascular response (70).

Strict control of serum glucose will retard the development or progression of diabetic neuropathy (71,72). There is also increasing evidence that antioxidants, aldose reductase inhibitors, α-lipoic acid, and γ-linolenic acid may play a role in preventing, halting, or improving diabetic neuropathy (73,74). Brain-derived nerve growth factor (NGF), recombinant NGF, neurotropin 3, granulocyte stimulating factor, and other peptides have also been used to treat diabetic neuropathy with limited success (75–77). In animal models of diabetic peripheral neuropathy, an increase in nerve conduction velocity was seen with implantation of hematopoietic mononuclear cells from peripheral blood or marrow (78). Similar results were obtained with small neurotrophic nonneural peptides (“dual-action peptides”) (79).

The most common complaint is of tingling or burning paresthesia in the feet, ankles, and calves. Due to distal weakness, postural control shifts from the ankle to hip, increasing the work of remaining upright. Autonomic neuropathy causes arteriovenous shunting and tissue hypoxia in the feet as well as in other organ systems. Pain, skin breakdown, and amputation are other complications of diabetic peripheral neuropathy.

It is essential to thoroughly educate the patient regarding footwear, foot care, and the need to maintain adequate glycemic control. Shoes should be wide, deep, and have an open throat design (Blucher last is preferable over Balmoral last). There should be adequate medial counter and arch support. Shoe orthotics can improve pain and ambulation when foot deformity or dynamic instability results from intrinsic muscle weakness. Pain can be managed with analgesics, tricyclic antidepressants, tizanidine, antihistamines, and anticonvulsants (including pregabalin and gabapentin). Other agents with some efficacy include topical capsaicin, counterirritants, topical lidocaine, and transcutaneous nerve stimulation (80).

### INFECTIONAL NEUROPATHY

#### Viral Infection

The peripheral nerve can be damaged by infection as well as by the response to infection. The most common viral diseases that affect the peripheral nerve are human immunodeficiency virus (HIV), cytomegalovirus, Epstein-Barr virus, herpes zoster,
Hepatitis C, poliomyelitis, parvovirus, and rabies. West Nile viral infection can produce lesions in the anterior horn cell and motor axon similar to acute poliomyelitis (81). Leprosy remains the most common peripheral neuropathy worldwide.

As many as 50% of patients with chronic HIV have some form of peripheral neuropathy, commonly undiagnosed (82). HIV-related peripheral neuropathies include AIDP, CIDP, mononeuritis multiplex, lumbar polyradiculopathy, and lymphomatous neuropathy (83,84). The degree of neurological impairment parallels the plasma HIV1 RNA level (84).

**Bacterial Infection**

Lyme disease, leprosy, and diphtheria can also produce a peripheral neuropathy.

Infection with *Borrelia* (Lyme disease) can produce several patterns of peripheral neuropathy. The most common pattern occurs late in the disease course. It is a large fiber axonopathy which causes a sensory polyneuropathy (85). Compromise of the endoneurial vascular supply can produce a mononeuritis multiplex (86).

Treatment of the infectious neuropathies addresses the concurrent infection. Appropriate antibiotic or antiviral agents should be prescribed. Hyperbaric O\textsubscript{2} can salvage peripheral nerve fibers from infection as well as ischemic, toxic, or antibody-mediated degeneration (87). Hyperbaric O\textsubscript{2} can improve symptoms of neuropathy in HIV patients (88). In these patients, recombinant NGF also appears to be useful (89).

Rehabilitation management goals include increasing strength or compensating for weakness with bracing or adaptive equipment in order to restore or maintain independence. Exercise may focus on stretching, strengthening, or improving endurance. Splinting, bracing, and/or adaptive equipment are usual components of the rehabilitation prescription.

### NERVE ENTRAPMENT/COMPRESSION SYNDROMES

Nerve entrapment/compression syndromes are common and may involve almost any peripheral nerve (Table 29-9). Long peripheral nerves such as the median, ulnar, radial, peroneal, and tibial nerves are most vulnerable, usually in their more distal segments. Compression can be acute or chronic and result from external compression, swelling of the nerve or compression of vascular supply, and ischemia. Pressure and rigidity in fascial compartments can result in neurapraxia due to segmental demyelination. There are changes in the intraneurial microcirculation, impairment of axonal transport and changes in vascular permeability, all of which contribute to edema formation and block conduction (90). Risk factors include a rise of pressure within a contained site (e.g., fluid retention before menses or tendon thickening/edema from overuse), rigid containment (carpal tunnel borders), a pathologic increase in nerve caliber (edema or hypertrophic remyelination), stretching or tethering the nerve with a related increase in resting tension, or the presence of anomalous muscle or bone in a common confined space. The best diagnostic tool to confirm and stage an entrapment neuropathy is the electrodiagnostic evaluation.

Conservative treatment options for entrapment/compression syndromes relieve pressure and nerve compression. Strategies include splinting to restrict motion of a joint and reduce the intermittent increase in pressure on the nerve. Treatment should be prescribed for any underlying condition that leads to fluid retention, or for any abnormality that produces thickening of tissue such as in myxedema, gout, or acromegaly. Medications, including local injection of steroids, can reduce inflammation, edema, and pain. Other therapeutic modalities utilized include massage and tendon mobilization. Splinting can improve hand position function as well.

### TABLE 29.9 Compression/Entrapment Peripheral Neuropathy

<table>
<thead>
<tr>
<th>Nerve</th>
<th>Sites of Entrapment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cervical/thoracic neurovascular bundle</td>
<td>Scalene muscles</td>
</tr>
<tr>
<td>Thoracic outlet syndrome</td>
<td>Accessory rib, Pancoast tumor</td>
</tr>
<tr>
<td>Median nerve</td>
<td>Flexor retinaculum (carpal tunnel)</td>
</tr>
<tr>
<td>Ulnar nerve</td>
<td>Elbow (ulnar sulcus, heads of the flexor carpi ulnaris, cubital tunnel)</td>
</tr>
<tr>
<td>Radial nerve</td>
<td>Guyon’s canal, Hypothenar eminence</td>
</tr>
<tr>
<td>Radial nerve</td>
<td>Circumflex portion at midhumerus</td>
</tr>
<tr>
<td>Radial nerve</td>
<td>Proximal to supinator margin</td>
</tr>
<tr>
<td>Radial nerve</td>
<td>Arcade of Frohse</td>
</tr>
<tr>
<td>Lateral femoral cutaneous nerve</td>
<td>Along pelvis</td>
</tr>
<tr>
<td>Peroneal nerve</td>
<td>Exiting pelvis</td>
</tr>
<tr>
<td>Tibial nerve</td>
<td>Iliac crest and anterior superior iliac spine</td>
</tr>
<tr>
<td>Tibial nerve</td>
<td>Fibular head</td>
</tr>
<tr>
<td>Tibial nerve</td>
<td>Laciniate ligament (tarsal tunnel)</td>
</tr>
</tbody>
</table>
limit motion that may be contributing to a stretch injury of the nerve. Surgical release is indicated when nonoperative management fails, when acute compression requires immediate relief of pressure to preserve viability of the nerve, or in the presence of an anomaly or foreign object causing nerve compression.

**Thoracic Outlet Syndrome**

The neurologic (versus vascular) thoracic outlet syndrome (TOS) produces symptoms in median and/or ulnar nerve distributions. Examination findings suggest lower trunk involvement, specifically intrinsic hand muscle weakness. Sensory impairment can be seen in a C8 to T1 distribution.

Etiologies most commonly involve compression between the scalene muscles heads, an accessory rib, or compression by a Pancoast tumor. The rehabilitation prescription emphasizes shoulder and neck range of motion, posture, and strengthening of the shoulder girdle. If conservative measures fail, surgical intervention focuses on correction of the anatomic cause of the syndrome.

**Median Nerve**

The most common entrapment syndrome is carpal tunnel syndrome (91,92). The usual presenting complaint is wakening with hand or finger pain and paresthesias. Another common complaint is clumsiness of the hand and dropping objects. In approximately 50% of patients there is no clear etiology. More than 50% of patients eventually have bilateral complaints (93).

Carpal tunnel etiologies include repetitive trauma, pregnancy, rheumatoid arthritis, anomalies of muscle or tendon, gout, myxedema, amyloid deposition, trauma with fracture, and scleroderma.

On examination, there is diminished touch, two-point discrimination, and weakness in a distal median nerve distribution. There is sensory sparing of the thenar eminence (median palmar cutaneous nerve) and of the skin in the ulnar and superficial radial nerve distributions.

Examination of the median innervated thenar muscles can reveal atrophy. Abductor pollicis brevis muscle weakness can be present, but there is no weakness of the median-innervated forearm muscles.

Provocative tests can produce or worsen paresthesias in the median digital nerve distributions. The Tinel's, or nerve percussion, sign (tapping over the median nerve at the flexor retinaculum) and eliciting distal paresthesias indicate demyelination/remyelination. Phalen's maneuver (the wrist in complete flexion) or reverse Phalen's maneuver (maximum dorsiflexion) supports the diagnosis of carpal tunnel syndrome. Direct compression over the distal edge of the carpal tunnel can also provoke paresthesias.

The treatment of carpal tunnel syndrome includes use of a volar wrist splint to limit wrist motion and place the wrist in a neutral position. The more time the splint is used daily the greater the reduction in symptoms (94). Local steroid injection into the carpal tunnel, or medications for neuropathic pain, can be helpful. Modalities, massage, range-of-motion exercises, stretch of the transverse carpal ligament, and tendon-gliding exercises are useful. Ergonomic modifications place the wrist in a neutral position and reduce carpal tunnel symptoms related to work.

When conservative treatment fails, surgery is considered. A good outcome requires complete section of the transverse carpal ligament. The approach may be either open or endoscopic. The open surgery is technically less demanding and has a lower risk of complications (95). The endoscopic method results in earlier return to work and activities of daily living. There is a higher incidence of recurrence after endoscopic carpal tunnel release (95).

Acute carpal tunnel syndrome due to wrist fractures or dislocation with injury to the median nerve is another indication for surgical release.

**Impact on Vocation**

Entrapment neuropathies and compression neuropathies can have a profound impact on vocation. Numbness, tingling, and weakness can lead to a deterioration in the quality and productivity of work (96).

Carpal tunnel syndrome is the most frequent entrapment/compression neuropathy that affects work. Several factors are associated with a poor prognosis for return to work, such as strenuous work activity, an occupational history of repetitive hand movements, involvement in a worker's compensation case, and use of vibrating tools. The etiology of compression is generally flexor tendon inflammatory swelling due to repetitive strain. Treatment includes control of inflammation, ergonomic assessment, and modification of the workstation and tools used.

**Ulnar Nerve**

Ulnar nerve entrapment or compression can occur at several points along the nerve course. Most common are the elbow (within the ulnar sulcus or more distally between the heads of the flexor carpi ulnaris) in the cubital tunnel and at the wrist (Guyon's tunnel, between the hook of the hamate and pisiform). Compression can also occur most distally within the hypothenar eminence (generally due to repetitive strain injury). The ulnar nerve can also be compressed between the elbow and wrist when there is anomalous exit of the nerve through the flexor carpi ulnaris. Examination exposes weakness of the ulnar forearm muscles, flexor digitorum profundus (ulnar), and the flexor carpi ulnaris, depending upon the level of compression. Weakness may be limited to the ulnar intrinsic hand muscles. Sensory deficits in the ulnar branch at the hand dorsum help to differentiate between proximal compression and compression within Guyon's tunnel.

**Radial Nerve**

Radial nerve compression can occur at several sites along the course of the nerve. Exam findings correlate well with the level of compression. Weakness can occur in the elbow, wrist, or fingers. Sensory loss can be noted in the posterior arm, forearm, or hand.
Weakness of the triceps is compatible with entrapment proximal to the circumflex portion of the nerve at the mid-humeral level. Brachioradialis weakness with normal triceps strength suggests a lesion distal to the circumflex portion of the nerve. Sparing of the extensor carpi radialis is noted in entrapment proximal to the supinator margin, at the arcade of Frohse. A posterior interosseous nerve compression involves the flexor carpi ulnaris and long finger extensors without sensory involvement. A superficial radial nerve lesion produces sensory findings without weakness.

Splinting may be helpful if recovery is anticipated. Dynamic orthoses can provide wrist and finger extension if necessary, although power grasp remains compromised. Surgical release is most commonly needed when the nerve is compressed at the arcade of Frohse.

**Lateral Femoral Cutaneous Nerve**

**Meralgia Paresthetica (Bernhardt-Roth Syndrome)**

This is most commonly seen in middle age. There is a gender predisposition for males. Meralgia paresthetica presents as pain or irritation over the anterior or anterolateral thigh and results from injury, compression, or disease of the lateral femoral cutaneous nerve (LFCN). Entrapment can occur near the spinal column, within the abdominal cavity as the nerve courses along the pelvis and as it exits the pelvis (the most common site). It can also be caused by superficial compression near the iliac crest and Anterior superior iliac spine. Tight clothing, trauma (such as acute compression by a seatbelt during rapid deceleration), gynecologic etiologies (related to menstrual cycle, endometriosis, or fetal compression during the second and third trimesters), and obesity are not uncommon causes.

Less likely but possible etiologies include pelvic fracture, pelvic osteotomy, hysterectomy, tumor, hemorrhage, and abscess. On examination, hip extension may increase, and flexion may decrease, angulation and tension of the LFCN and symptoms.

Up to 20% percent of the time, meralgia paresthetica is bilateral. Treatment includes correction of leg length discrepancy to minimize hip hyperextension on the affected side and corticosteroid injection at the inguinal level (97).

**Peroneal Nerve**

The most common site of peroneal nerve compression is behind the fibular head. Ankle dorsiflexor, foot evertor and toe extensor weakness, and hypesthesia in the leg and foot, are typical findings. Sensory loss occurs in both the deep and superficial peroneal nerve distributions. The common peroneal nerve sensory territory at the proximal half of the lateral leg is spared, as are muscles innervated proximal to the knee. Treatment may include prescription of an appropriate ankle-foot orthosis, range-of-motion exercises, and strengthening exercises.

**Tibial Nerve**

The most common entrapment of the tibial nerve occurs at the tarsal tunnel. There is pain and/or paresthesia of the medial heel and plantar foot. The medial arch is variably involved depending upon the saphenous nerve sensory territory. There may be a history of foot injury or deformity such as pes planus. A nerve percussion sign may be present at the tarsal tunnel.

Foot orthoses can compensate for weakness and may help to relieve associated pain. Surgical release of the laciniate ligament is an option when nonoperative treatment does not provide symptomatic relief.

**TUMORS OF THE PERIPHERAL NERVE SHEATH**

The nerve sheath tumors that can undergo malignant transformation include schwannomas and neurofibromas. In neurofibromatosis type 1 (NF1 gene mutation), there is loss of the tumor suppressor gene on chromosome 22, decreasing neurofibromin production. Pain is present 30% of the time. Malignant transformation is most common with NF type 1 (10%) and should be suspected when there is rapid enlargement of the nerve tumor (98,99).

**DIAGNOSIS OF PERIPHERAL NEUROPATHY**

The diagnosis of peripheral nerve injury must be accurate in order to select the most appropriate intervention or repair (Table 29-10). It is based on the clinical examination and the electrodiagnostic evaluation.

**Clinical Findings**

Manual muscle testing will reveal weakness in a peripheral nerve distribution. Muscle atrophy can reach 50% to 70% of the muscle bulk after about 2 months. The sensory exam may identify an impaired sense of sharp or light touch and two-point discrimination. Later, a shrinking area of sensory loss with an enlarging area of partial sensation indicates that anastomotic branches with other nerves are forming (12).

Involvement of autonomic fibers is evidenced by a loss of sweating and impaired pilomotor and vasomotor action. If the skin wrinkles on immersion in water, or if sweating is present, the peripheral nerve damage is incomplete.

A nerve percussion sign is indicative of demyelination/remyelination. When the nerve percussion sign progresses from being positive in the proximal nerve to being positive in sequentially more distal nerve segments, healing takes place. This distally advancing Tinel’s sign occurs as the nerve recovers from a Sunderland II or III lesion (13).

The best diagnostic tool to characterize/confirm the presence of peripheral neuropathy remains the electrodiagnostic examination. Nerve conduction studies focus on sensory, motor, and mixed nerve function in the distribution of symptoms or lesion. Evidence of focal demyelination includes prolongation of distal latencies or slow/absent conduction. In axonopathy, amplitudes of the evoked response are low or absent.
The electrodiagnostic exam is valuable for both diagnosis and prognosis. Nerve conduction studies may be normal until Wallerian degeneration takes place. Neurapraxia is unlikely if sensory conduction remains absent after 7 to 10 days. Table 29-11 summarizes electrodiagnostic findings following nerve injury.

Magnetic resonance imaging (MRI) studies are also used to diagnose peripheral nerve injury (9). Visible changes can be detected by MRI at about the fourth postinjury day. MRI has good sensitivity and specificity, especially on the short-tau inversion recovery signals for axonotmetic or neurotmetic lesions. Neurapraxia yields normal findings. Magnetic resonance neurography and even plain films for associated injuries, such as fracture or dislocation, may be useful.

Laboratory screening can be helpful to identify the etiology of peripheral neuropathy. Routine studies include serum glucose, hemoglobin, leukocytes, platelets, erythrocyte sedimentation rate, thyroid studies, creatinine, and serum protein electrophoresis.

### TABLE 29.10 Mononeuropathies: Diagnostic and Treatment Considerations

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Clinical presentation suggestive of peripheral nerve entrapment/compression syndrome: history and physical examination</td>
<td>1. Avoid or modify aggravating activities—ergonomic modifications</td>
</tr>
<tr>
<td>2. Electrodiagnostic confirmation of a nerve segment lesion</td>
<td>2. Treatment of underlying medical conditions</td>
</tr>
<tr>
<td>3. Correlate history, physical examination, and electrodiagnostic findings</td>
<td>3. Splinting—resting splints: to minimize nerve compression secondary to joint position, muscle/tendon overuse, or stretch of the nerve</td>
</tr>
<tr>
<td>4. Electrodiagnostic excludes radiculopathy, plexopathy, or peripheral polyneuropathy</td>
<td>4. Minimize external compression</td>
</tr>
</tbody>
</table>

### TREATMENT STRATEGIES FOR PERIPHERAL NEUROPATHY

Treatment options for peripheral neuropathy include treatment of underlying disease, splinting, medication, therapeutic modalities, stem cell therapies, and surgery (Table 29-12). Multiple agents have been studied for modifying peripheral neuropathy itself through genetic or direct effect. They might enhance nerve cell survival, promote neural anabolism, reduce oxidative stress, or delay apoptosis. Their value is unclear. Human or animal studies have been generally inconclusive to date (100–107). These agents are listed in Table 29-13.

Stem cell therapy is becoming an increasing valuable tool to encourage peripheral nerve healing. There is evidence in both rats and humans that myelinating Schwann cells can be obtained by seeding with marrow stem cells (recovery from pain and sensory nerve function) or mesenchymal stem cell seeding (animal model). Human amniotic mesenchymal stem
cells can improve growth of a peripheral nerve across any gap in continuity (108–111). Treatment of underlying disease also addresses conditions that lead to fluid retention or thickening of tissues surrounding a nerve. There may be a need to correct nutritional deficiencies, eliminate toxic exposure, or suppress an immune response.

Splinting is useful for restricting joint or tendon motion and reducing intermittent pressure on the nerve. Strategies that mobilize tissues and reduce adhesions (massage, ultrasound, exercise) are useful in entrapment neuropathies before considering surgical release.

**Treatment of Pain**

Medications, local anesthetic or corticosteroid injection, and transcutaneous nerve stimulation are the most frequent methods of addressing peripheral nerve pain. Similar strategies are used for the chronic neuropathic pain that can follow an acute painful peripheral neuropathy.

**TABLE 29.12 Rehabilitation Management of Peripheral Neuropathy**

<table>
<thead>
<tr>
<th>Problem</th>
<th>Functional Consequence</th>
<th>Rehabilitation Intervention</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weakness</td>
<td>Proximal Distal</td>
<td>Rising from chair, jumping stairs, Foot drop/slap, poor toe clearance, Ankle sprains, loss of dexterity, and fine motor control</td>
</tr>
<tr>
<td>Generalized weakness and deconditioning</td>
<td>Distal sensory loss</td>
<td>Loss of endurance</td>
</tr>
<tr>
<td>Impaired motor control</td>
<td>Autonomic dysfunction</td>
<td>As above</td>
</tr>
<tr>
<td>Pain</td>
<td>Deformity</td>
<td>Decreased activity</td>
</tr>
</tbody>
</table>

**TABLE 29.13 Agents With Possible Effect on Peripheral Neuropathy**

- Neurotrophic growth factors
  - Ciliary neurotrophic factor (CNTF)
  - Brain-derived neurotrophic factor (BDNF)
  - Glial cell line–derived neurotrophic factor (GDNF)
  - Recombinant human insulin-like growth factor
- Antioxidants
  - Ascorbic acid
  - Other
- Other
  - Stem cells
  - Curcumin
  - Cannabinoids

Medications used to manage pain include tricyclic antidepressants, selective serotonin/serotonin-norepinephrine reuptake inhibitors (SSRI, SNRI), nonsteroidal anti-inflammatory drugs (NSAIDs)—including transdermal preparations, opiate agonists, topical counterirritants, topical capsaicin, stimulatory peptides, and anticonvulsant or antispasticity agents (111).

Anticonvulsants modulate sodium channels and suppress ectopic or ephaptic discharge in neuropathic pain. Gabapentin and pregabalin have demonstrated efficacy in decreasing pain, improving sleep, and enhancing mood and quality of life for patients with postherpetic or diabetic neuropathy.

Antidepressants are helpful in pain-related depression, and both the tricyclic and the SNRI medications show some efficacy in treating pain. Stimulatory peptides (short-chain amino acids derived from cytokine proteins or growth factors) encourage healing, alleviate pain, and may prevent neuron death (112).

Transcutaneous electrical stimulation (TENS), pulsed galvanic stimulation, and heat and cold modalities (113–116) are useful nonpharmacological means of providing analgesia. Mobilization of a painful limb segment and normal use of a limb is important, particularly as chronic regional pain syndromes may complicate a peripheral neuropathy (117).

**PERIPHERAL NERVE REPAIR FOLLOWING TRAUMA**

The ability of the peripheral nerve to recover from trauma, including both injury and subsequent repair, depends upon a variety of intrinsic and extrinsic factors. When the cell body is viable, the peripheral nervous system demonstrates a great capacity for recovery, beyond that of the central nervous system. This capacity is facilitated or inhibited by intrinsic factors.
factors such as age, state of tissue nutrition, time since injury and type of injury, the specific peripheral nerve involved, and the level of lesion. In general, well-nourished younger patients with recent injury that is located at the distal nerve trunk do better. Similarly, extrinsic factors such as medical treatment and surgical and postoperative management may enhance or inhibit healing (9).

Negative prognosticators for successful nerve repair include advanced age, nerve injury due to dislocation, delay of repair beyond 5 months, prior radiation therapy, nerve discontinuity (gap) exceeding 2.5 cm, proximal nerve injury, and a poor condition of the nerve endings (6). Sensation may improve as late as at 2 years following repair.

Initially, surgery may be exploratory in order to establish an accurate diagnosis when the clinical and electrodiagnostic examinations are inconclusive. Exploratory surgery exposes the extent and severity of the lesion. Surgery may also be considered when there is an incomplete loss of function but no improvement over several weeks or no return of function at about 2 months after injury (4 months for the brachial plexus). The purpose of surgical repair is to improve peripheral nerve recovery and eventual function—if the prognosis for this is poor, surgery should not be done.

There are two basic approaches to reestablishing nerve continuity: macrorepair (reconnecting the nerve ends by epineurial suture) and microscopic (this consists of suturing individual nerve bundles) (118). Surgery is only done when there is adequate length of a viable nerve. Repair is done for Sunderland grade 3–5 lesions. Early repair is desirable, as maximal nerve regeneration occurs at about 3 weeks after injury (7).

Immediate or primary repair is performed within 8 to 12 hours on clean lacerating injuries. Delayed or secondary repair, which is more usual following major trauma, blunt transection or wound contamination, or complete neurotmesis without clear nerve endings, is performed after more than 1 month. This delay in repair allows better definition of the lesion and a reduced risk of infection.

Macroneurorrhaphy, reestablishing the continuity of the epineurium, remains the standard surgical repair. Epineurial repair, fascicular or perineurial nerve repair, and interfascicular nerve repair are also done, but there are insufficient data that demonstrate a better outcome from this type of surgery. Fascicular groups must be large enough to anchor sutures, and proper tensioning of the repaired nerve is essential (119).

Placement of a peripheral nerve graft is done when the resected nerve segment produces a gap between nerve endings greater than 2 cm (120). Resorbable nerve conduits seem to work well following neurona suture, but not for nerve repair (Meek, Nicolai, Robinson). Neurolysis may be needed to remove scar tissue. In order to bridge a gap in nerve continuity, surgical alternatives to grafting include nerve mobilization, nerve transposition, tissue expansion (121), joint flexion, tubulization (with or without Schwann cell–seeded collagen matrix in the conduit), and bone shortening (122).

Trophic factors that enhance nerve regeneration through a graft or tube include neurocytokines (NGF), brain-derived neurotrophic factor (BDNF), ciliary neurotrophic factor, and neurotrophin-3, -4, -5, and -6 (123,124). Trophic factors work by preventing fibroblast activity and inhibiting protease activity.

After surgery, passive and active range of motion of related joints is begun and splinting (12t) is provided as necessary. Hyperbaric oxygen therapy decreases endoneurial edema, pressure, and vascular compromise in ischemic lesions (87).

There is some controversy regarding early passive mobilization following surgery. Most commonly, the operated limb is immobilized for 4 to 6 weeks, apart from providing limited range of motion to prevent or limit joint contracture (125). Early tactile stimulation may enhance recovery of sensory function (122).

Conditions that can complicate the postoperative course include chronic pain syndromes, neuroma formation, peripheral nerve entrapment syndromes, complex regional pain syndromes type 1 and 2, weakness, sensory loss, and atrophic skin changes. If neurosurgery fails, tendon transfers can sometimes offer the patient a partial restitution of function.

Remaining current regarding related research and applications allows the physiatrist to guide patients with pain, weakness, and/or disability due to peripheral neuropathy to optimal rehabilitation outcomes.

REFERENCES


34. Davids H. Guillain-Barre syndrome June 7, 2006 eMedicine


55. Hahn AF. Hereditary motor and sensory neuropathy: HMSN type II (neuronal type) and X-linked HMSN. Brain Pathol. 1993;3(2):147–155.


CHAPTER 29  | PERIPHERAL NEUROPATHY 755


Myopathy

Representing a diverse group of disorders primarily affecting skeletal muscle, myopathies are an important cause of disability affecting patient mobility, self-care, and independence. In addition to weakness, many myopathies have associated dysfunction in other organ systems, such as the cardiac and pulmonary systems. The disability associated with muscle disease depends on the specific type, extent of clinical involvement, and rate of progression.

The number and type of different muscle disorders under the umbrella term myopathy are vast and expanding. With our increasing knowledge about genetic and molecular basis for these disorders, classification and nomenclature regarding myopathies are constantly being reevaluated and modified. A detailed discussion of all the different myopathies is beyond the scope of this chapter; however, this chapter is intended to provide the physical medicine and rehabilitation specialist with an overview of diagnostic approach, clinical characteristics, and care and management of patients with myopathies, with emphasis on few of the most common diagnoses that a physiatrist may encounter.

Although most myopathies remain largely incurable, as is the case for most neuromuscular diseases at this time, they are not untreatable. Rehabilitation specialists have an important role in the care of patients with myopathies to maximize their functional capacities, prolong or maintain independent locomotion and function, prevent physical deformity and medical complications, and provide access for integration into the community with quality of life in mind. The comprehensive management of all the varied clinical problems associated with myopathies and other neuromuscular diseases often requires specialists from neurology, cardiology, pulmonology, and orthopedic surgery as well as clinical specialists in physical therapy, occupational therapy, speech therapy, and orthotics. Coordination of this difficult and demanding task may be best handled by a neuromuscular and rehabilitation medicine specialist knowledgeable in various myopathies. It is important for the rehabilitation physician to understand these diseases in order to appropriately treat the functional problems caused by muscle weakness as well as provide comprehensive interdisciplinary rehabilitation through awareness of other manifestations of the disease.

**DEFINITION AND CATEGORIES OF MYOPATHY**

In the peripheral nervous system, a primary defect may occur at the level of the anterior horn cell, peripheral nerve, neuromuscular junction, or muscle. A disease process in which the primary abnormality is at the level of the muscle itself is termed myopathy. A brief overview of the various myopathies is presented in this section to help the reader classify numerous myopathies in an orderly fashion. A more detailed discussion of the individual myopathies pertinent to rehabilitation specialists is found in the subsequent sections.

There are three basic categories of myopathy: hereditary, acquired, and myopathies associated with systemic disease (Table 30-1). In the hereditary myopathies, all of the myopathies have their inheritance pattern characterized or gene mutations identified. This category of myopathy is further subdivided into muscular dystrophies, congenital, distal, metabolic, and mitochondrial myopathies. In general, the muscular dystrophies present with significant structural defect of the muscle cell due to mutations in genes crucial for its normal function. Typically, the muscular dystrophies are accompanied by progressive muscle fiber degeneration, atrophy, and inflammation as noted by histopathological evaluation. The specific muscular dystrophies that are more commonly encountered, with its characteristic inheritance pattern, affected genes or mutations are listed (Table 30-2). Congenital myopathies are a group of relatively nonprogressive muscle diseases that present during infancy or early childhood and are classified largely based on clinical features and muscle biopsy morphological findings. Distal myopathies, as the name implies, have more pronounced weakness involving distal limb muscles rather than the typical proximal weakness seen with majority of other myopathies. Metabolic myopathies are caused by gene mutations that result in either abnormal glycolysis or lipid metabolism. Usually, deficiency of an enzyme results from a gene mutation and causes an abnormal accumulation of substrate or a deficiency of the product of the enzymatic pathway. Mitochondrial myopathies are a group of muscle disorders with maternal inheritance pattern associated with abnormal structure and function of mitochondria. Since mitochondria are important in energy production throughout the body, other organ system involvements are also common including the nervous, cardiac, gastrointestinal, pulmonary, and endocrine systems. Diagnosis is often based on a combination of clinical findings and associated biochemical defects, along with histological abnormality noted as “ragged red fibers” on the modified Gomori trichrome stain.

Another group of myopathies called channelopathies include primary muscle disorders caused by inherited abnormalities of various ion channels found on cell membranes. These include myotonia congenita, paralympotonia congenita, and primary hyperkalemic and hypokalemic periodic paralysis.
### TABLE 30.1 Myopathies

<table>
<thead>
<tr>
<th>Hereditary Myopathies</th>
<th>Acquired Myopathies</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Muscular dystrophies</strong></td>
<td>Inflammatory myopathies</td>
</tr>
<tr>
<td>Duchenne muscular dystrophy (DMD)</td>
<td>Polymyositis (PM)</td>
</tr>
<tr>
<td>Becker muscular dystrophy (BMD)</td>
<td>Dermatomyositis (DM)</td>
</tr>
<tr>
<td>Myotonic muscular dystrophy (DM1 and DM2)</td>
<td>Inclusion body myositis (IBM)</td>
</tr>
<tr>
<td>Facioscapulohumeral muscular dystrophy (FSHD)</td>
<td>Toxic myopathies</td>
</tr>
<tr>
<td>Limb-girdle muscular dystrophy (LGMD)</td>
<td>Corticosteroid myopathies</td>
</tr>
<tr>
<td>Congenital muscular dystrophy (CMD)</td>
<td>Lipid-lowering agent myopathies</td>
</tr>
<tr>
<td>Oculopharyngeal muscular dystrophy (OPMD)</td>
<td>Alcohol-related myopathies</td>
</tr>
<tr>
<td>Emery-Dreifuss muscular dystrophy (EDMD)</td>
<td>Myopathies related to other medications</td>
</tr>
<tr>
<td><strong>Congenital myopathies</strong></td>
<td>Endocrine myopathies</td>
</tr>
<tr>
<td>Central core, nemaline, centronuclear, multicore</td>
<td>Myopathies with glucocorticoid abnormalities</td>
</tr>
<tr>
<td>Fiber type disproportion, reducing body</td>
<td>Myopathies with thyroid disease</td>
</tr>
<tr>
<td>Fingerprint, cytoplasmic body, myofibrillar</td>
<td>Myopathies with parathyroid disease</td>
</tr>
<tr>
<td><strong>Metabolic myopathies</strong></td>
<td>Myopathies associated with pituitary dysfunction</td>
</tr>
<tr>
<td>Disorders of glycogenoses</td>
<td>Myopathies related to electrolyte disturbance</td>
</tr>
<tr>
<td>Disorders of lipid metabolism</td>
<td><strong>Infectious and granulomatous myopathies</strong></td>
</tr>
<tr>
<td>Respiratory chain defects</td>
<td>Viral, bacterial, fungal, tuberculous, parasitic</td>
</tr>
<tr>
<td><strong>Distal myopathies</strong></td>
<td>Sarcoid myopathy</td>
</tr>
<tr>
<td>Welander, Markesbery-Griggs-Udd, Nonaka, Miyoshi, Laing</td>
<td><strong>Myopathies associated with systemic disease</strong></td>
</tr>
<tr>
<td><strong>Mitochondrial myopathies</strong></td>
<td>Critical illness myopathy</td>
</tr>
<tr>
<td>Kearns-Sayre’s syndrome, Progressive External Ophthalmoplegia (PEO)</td>
<td>Electrolyte disturbances</td>
</tr>
<tr>
<td>Mitochondrial Encephalomyopathy Lactic Acidosis Stroke (MELAS)</td>
<td><strong>Paraneoplastic</strong></td>
</tr>
<tr>
<td>Myoclonic Epilepsy Ragged Red Fibers (MERRF)</td>
<td><strong>Channelopathies</strong></td>
</tr>
<tr>
<td>Neuropathy Ataxia Retinitis Pigmentosa (NARP), Myopathy and external ophthalmoplegia Neurorophy Gastro-Intestinal Encephalopathy (MNGIE)</td>
<td></td>
</tr>
<tr>
<td>Leber’s Hereditary Optic Neuropathy (LHON), Leigh’s syndrome</td>
<td></td>
</tr>
<tr>
<td><strong>Channelopathies</strong></td>
<td></td>
</tr>
<tr>
<td>Myotonia congenita</td>
<td></td>
</tr>
<tr>
<td>Paramyotonia congenita</td>
<td></td>
</tr>
<tr>
<td>Primary hyperkalemic and hypokalemic periodic paralysis</td>
<td></td>
</tr>
</tbody>
</table>

### TABLE 30.2 Inheritance Pattern and Gene Mutations of Muscular Dystrophies

<table>
<thead>
<tr>
<th>Muscular Dystrophies</th>
<th>Inheritance Pattern</th>
<th>Gene Loci</th>
<th>Gene (Mutations)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DMD</td>
<td>X-linked</td>
<td>Xp21</td>
<td>Dystrophin</td>
</tr>
<tr>
<td>BMD</td>
<td>X-linked</td>
<td>Xp21</td>
<td>Dystrophin</td>
</tr>
<tr>
<td>DM1</td>
<td>AD</td>
<td>19q13.3</td>
<td>DMPK (expansion of CTG repeat)</td>
</tr>
<tr>
<td>DM2</td>
<td>AD</td>
<td>3q21</td>
<td>ZNF9 (expansion of CCTG repeat)</td>
</tr>
<tr>
<td>FSHD</td>
<td>AD</td>
<td>4q35</td>
<td>DUX4 (deletions of D4Z4 repeats)</td>
</tr>
<tr>
<td>LGMD (type 1s)</td>
<td>AD</td>
<td>Multiple loci</td>
<td>Myotilin, laminin A/C, caveolin-3</td>
</tr>
<tr>
<td>LGMD (type 2s)</td>
<td>AR</td>
<td>Multiple loci</td>
<td>Calpain-3, dysferlin, sarcoglycans, Fukutin-related protein (FKRP)</td>
</tr>
<tr>
<td>OPMD</td>
<td>AD or AR</td>
<td>14q11.2-q13</td>
<td>PABPN1 and PABP2 (Polyadenylate-binding protein, Nuclear) (expansion of GCG repeat)</td>
</tr>
<tr>
<td>EDMD 1</td>
<td>X-linked</td>
<td>Xq28</td>
<td>Emerin</td>
</tr>
<tr>
<td>EDMD 2</td>
<td>AD</td>
<td>1q21.2</td>
<td>Laminin A/C</td>
</tr>
<tr>
<td>CMD</td>
<td>AR</td>
<td>6q22</td>
<td>Merosin</td>
</tr>
</tbody>
</table>

AD, autosomal dominant; AR, autosomal recessive.
The second category, acquired myopathies, consists of inflammatory, endocrine, toxic, granulomatous, and infectious myopathies. Under inflammatory myopathies are polymyositis (PM), dermatomyositis (DM), and inclusion body myositis (IBM). Muscle disorders associated with various endocrinopathies are now well recognized. These include myopathies associated with thyroid dysfunction (hyperthyroidism or hypothyroidism), adrenal disease, pituitary dysfunction, and parathyroid dysfunction (hyperparathyroidism or hypoparathyroidism). Myopathies can also result from electrolyte disturbances, including abnormalities of serum potassium, sodium, calcium, magnesium, and phosphorus. Under the toxic myopathy category, the most common agents associated with myopathy include HMG-CoA reductase inhibitors (cholesterol-lowering agents), corticosteroids, fibrac acid derivatives (lipid- and cholesterol-lowering agents), chloroquine and amiodarone (amphiphilic drugs), colchicine and vincristine (antimicrotubular agents), zidovudine (HIV medication), and alcohol. Of the toxic myopathies, alcohol-related myopathy is thought to be the most common with both acute and chronic manifestations, often associated with heavy and prolonged alcohol use. Although typically asymptomatic in terms of muscle manifestation, sarcoidosis can present in the form of a granulomatous myopathy. Lastly, infectious myopathies are associated with essentially all types of infectious agents including viral (coxsackievirus, HIV, HTLV-1), bacterial, fungal, tuberculous, as well as parasites.

The last category is myopathies associated with systemic disease. Under this category are myopathies that have significant systemic processes that result in derangement of muscle function and health. The most common etiologies are severe multiorgan failure with sepsis, electrolyte disturbances associated with systemic disease, and underlying neoplasms.

Clinical features and progression vary within and between these categories as pathophysiology of each muscle disorder is different. Some of the myopathies, partly due to their time course of progression, involvement of other organ systems, prevalence in the population, and the availability of rehabilitative treatment options, may be more or less pertinent to rehabilitation specialists. However, in order to devise an appropriate rehabilitation plan, the rehabilitation physician should understand the expected disabilities and prognosis associated with the specific cause of myopathy.

**EVALUATION OF THE PATIENT WITH SUSPECTED MYOPATHY**

**History**

The primary symptom of a patient with suspected myopathy is weakness, defined as a reduction in maximal force generated by a muscle or muscle group (Table 30-3). This weakness may be fairly acute or insidious. Because the weakness is typically in the proximal musculature, certain functional problems should alert the clinician to the possibility of myopathy: difficulty getting up from a chair or toilet seat, trouble descending and climbing stairs, or difficulty with overhead activities, such as dressing, grooming, or reaching cabinets (Table 30-4). Symptoms suggesting distal weakness, such as problems opening jars, may be prominent with certain myopathies. The symptom of muscle fatigue, defined as the inability to sustain a given level of force for a certain period, is often difficult to assess. Although it may be associated with myopathy, when fatigue is the predominant symptom, other pathologic processes, such as neuromuscular junction disease and upper motor neuron disease, are more likely.

Muscle pain, or myalgia, is a common presenting symptom, particularly in the inflammatory myopathies. However, the absence of pain should not distract the clinician from strongly considering the diagnosis of myopathy. When myalgias are the predominant symptom without demonstrated weakness, other disorders are more likely. History of myoglobinuria associated with weakness or fatigue symptoms should be sought and can help in the workup of muscle disorders, especially the metabolic myopathies. The presence of paresthesias or dysesthesias on history is certainly helpful, because they make the presence of myopathy very unlikely. A rare patient might interpret myalgias with descriptors sounding like sensory symptoms, distracting the clinician.

One of the most critical pieces of information is the family history. Whenever a myopathy is suspected that may have a genetic cause, a detailed family history and pedigree chart are essential. In an X-linked recessive disorder such as Duchenne
muscular dystrophy (DMD), men on the maternal side of the family are affected about 50% of the time and women are carriers in an equal percentage. Autosomal recessive disorders, such as many limb-girdle syndromes, frequently have no family members involved, making the diagnosis of a familial disorder more difficult. In an autosomal dominant disorder such as myotonic muscular dystrophy or facioscapulohumeral dystrophy (FSHD), typically 50% of offspring within a pedigree are affected. Sporadic cases resulting from new genetic defects occur with most autosomal dominant and sex-linked dystrophies, making a dystrophy possible even in the absence of a suspicious family history.

In a child presenting with weakness, a developmental history should include milestones of age for head control, independent sitting, standing, and walking. Additional factors related to ambulation include toe walking, excessive lordosis, falls, and running ability.

**Physical Examination**

Examination of the patient with suspected myopathy begins with observation (Table 30-5). In myopathies, muscle atrophy may not be obvious until late in the disease because of a wide normal range of variation in the population and the typical symmetry of muscular involvement. Calf enlargement may be noted in dystrophic myopathies, particularly in DMD and Becker muscular dystrophy (BMD). This “pseudohypertrophy” is caused by increased fat and connective tissue volume, rather than muscle fiber hypertrophy (1) (Fig. 30-1). Observation of facial features, such as a long thin face with temporal and maseter wasting with frontal balding, is typical for myotonic muscular dystrophy.

Other physical examination findings that may be particularly helpful in the evaluation of myopathies are the presence and distribution of rash, contractures, and ligamentous laxity. These may be useful when considering diagnoses such as DM, Emery-Dreifuss muscular dystrophy (EDMD), and muscle diseases associated with collagen dysfunctions. Cardiac examination is also important as some myopathies have an associated conduction abnormality or a cardiomyopathy. Examination of the pulmonary system can provide clues to an accompanying restrictive lung disease process or an aspiration pneumonia secondary to swallowing difficulties.

Because weakness is the predominant symptom, determination of muscle strength is critical. Unfortunately, the manual muscle test typically used by clinicians is only a very rough measure of strength. It is well known that up to 50% strength loss may occur before a muscle is graded as 4/5 using the Medical Research Council (MRC) scale (2). The more powerful pelvic proximal muscles are particularly difficult to measure, because the patient should be able to overcome the examiner’s resistance. The handheld dynamometer is a quantitative device to measure strength, but it shares the same limitation when strong muscles are being tested. It has been shown to provide reliable data in persons with neuropathic weakness (3). Because of a wide range of normality, the handheld dynamometer is probably better suited to measure serial strength than to quantify a specific muscle group as “normal” or “abnormal.”

Probably the most reasonable method to test strength in the clinic is to observe repetitive maneuvers, such as rising from a squat, repeatedly standing on the toes, or raising the arms overhead with resistance. The clinician should observe for Gowers’ sign: The patient rises from a low surface by pushing against the knees and moving the hands up the thighs to substitute for knee and hip extensor weakness.

Facial and neck muscle weakness predominates in several myopathies such as FSHD. The ability to “bury” the eyelashes or the ability of the examiner to easily overcome forced eye closure (because of orbicularis oculi weakness) and difficulty whistling (because of orbicularis oris weakness) are reasonable screening tests. The presence of ptosis or ophthalmoplegia

---

**TABLE 30.5** Key Physical Examination Points for Suspected Myopathy

| Proximal > distal weakness, including neck and facial muscles |
| Observation of facial features |
| Sensation—should be normal |
| Muscle tendon reflexes preserved or mildly decreased |
| Presence of clinical myotonia |
| Waddling gait with Gowers’ sign on standing |
| Positioning of the shoulder girdle |
Laboratory Evaluation
The most important blood study for suspected myopathy is measurement of serum creatine kinase (CK). With muscle fiber injury, this enzyme leaks into the serum. Particularly high elevations of CK (50 to 100 times normal) may be found in acute inflammatory myopathies and the early stages of DMD and BMD. The more slowly progressive dystrophies may have milder to moderate elevations in CK. However, CK is not the ideal screening test for all myopathies because the congenital myopathies, slowly progressive dystrophies, chronic inflammatory myopathies, and myopathies of systemic disease may have normal values. The clinician should be cautious not to overinterpret one mildly elevated CK level, because it may be elevated in healthy persons for several days after vigorous exercise. Conversely, once there is significant muscular atrophy, CK values may be low or normal based on the paucity of remaining muscle tissue to release the enzyme. Other serum transaminases, aldolase, and lactate dehydrogenase are often elevated in myopathy but are much less specific because they are found in liver in equally high amounts. In the metabolic myopathies, measurement of blood lactate and pyruvate may be helpful, particularly atactic lactate levels during ischemic or exercise stress. With abnormalities of glycogen metabolism, there will be no rise in lactate because patients cannot catabolize glycogen.

Electrodiagnosis
Electrodiagnostic studies (electromyography [EMG]) can be extremely important in the evaluation of the patient with myopathy to localize the pathology to the muscle rather than nerve or anterior horn cell. The pattern of EMG findings may indicate the best muscle for biopsy, and certain abnormalities on the EMG occasionally suggest a specific myopathic disease. However, electrodiagnostic studies in myopathy may be normal as well, so a myopathic disorder is not ruled out by normal EMG studies.

Nerve conduction studies should be normal in myopathic disorders, with the exception of a low compound motor action potential obtained when recording over muscles with severe atrophy. With needle EMG, the presence of abnormal spontaneous activity (positive sharp wave/fibrillation potentials) is dependent on whether the myopathy is causing active muscle fiber degeneration. For example, the inflammatory myopathies and rapidly progressive dystrophies frequently demonstrate abnormal spontaneous activity, whereas it is not often encountered in the slowly progressive dystrophies or myopathies associated with systemic disease.

The hallmark needle EMG finding suggesting myopathy is the presence of low-amplitude, often polyphasic, brief-duration potentials with voluntary contraction. Because recruitment of each additional motor unit only slightly augments strength, the electromyographer often notes an excessive number of motor units for a given strength of contraction. These findings may be subtle or absent, particularly in slowly progressive disorders. Particularly important muscles to evaluate with possible myopathy include the paraspinal, supraspinatus, infraspinatus, glutei, and iliopsoas muscles.

Muscle Biopsy
The ideal muscle for biopsy is weak, but not profoundly atrophic. Electrodiagnostic abnormalities increase the likelihood that the muscle will demonstrate useful findings, although one should not biopsy a muscle that has recently been evaluated with a needle electrode because of possible needle-induced fiber damage. The most accessible muscles include the vastus lateralis in the lower limb and the deltoid or biceps brachii in the upper limb. Histologic findings suggestive of myopathy include fiber necrosis, central nuclei indicative of regeneration, atrophied fibers, inflammatory infiltrates, and proliferation of connective
tissue and fibrosis. Certain congenital myopathies, including centronuclear or myotubular, central core, and nemaline rod, have distinctive histologic and electron microscopy findings. In addition to histologic studies, immunohistochemical techniques can provide information about the amounts of dystrophin and other structural membrane proteins.

**Molecular Genetic Studies**
Recent advances in molecular genetic techniques have resulted in remarkable increases in the knowledge of various myopathies. The chromosomal location, causal gene, and mutations have been identified in many neuromuscular disorders and are frequently helpful in diagnostic evaluation. An example of the impact of molecular genetic studies is the evaluation for possible DMD or BMD. Both disorders are caused by mutations in an extremely large gene located on the X chromosome. The protein product of the gene, known as dystrophin, was determined to be an important component of the muscle membrane cytoskeleton, contributing to the stability of the muscle fiber (5). For diagnosis, a clinically available gene deletion study from a blood sample is diagnostic of a dystrophinopathy, but it is able to detect mutations present in only about 65% of DMD patients and 80% of BMD patients. Additional DNA analysis to detect smaller mutations in the dystrophin gene increases the detection rate to approximately 90% of patients with DMD (6). However, a positive test does not clearly distinguish between DMD and BMD. A muscle biopsy for immunohistochemical analysis of the dystrophin protein is necessary in patients testing negative for the mutation or to differentiate between a patient with a particularly severe form of BMD versus a patient with a milder form of DMD. Absent dystrophin or levels less than 3% is consistent with DMD whereas in BMD, the dystrophin may have an abnormal molecular weight or decreased in quantity.

The number of commercially available genetic tests has grown tremendously over the past several years and continues to expand. In addition, there are numerous research laboratories that specialize in specific myopathies and can even offer genetic testing for research purposes, when commercial tests are not available. A list of clinical and research laboratories offering genetic tests for various myopathies or neuromuscular disorders can be found at: www.genetests.org. Although genetic tests occupy an important place among diagnostic tools now available to a clinician, it should not replace a careful history, thorough physical examination, and clinical common sense in the evaluation of a patient with myopathy.

**Clinical Features of Specific Myopathic Disorders**

**Dystrophic Hereditary Myopathies**

**Duchenne Muscular Dystrophy**

DMD is an X-linked disorder with the chromosomal abnormality at the Xp21 gene locus (7). As noted above, the gene codes for the protein dystrophin, an important cytoskeletal component of the muscle cell membrane. It appears that absence of dystrophin makes the muscle cell highly susceptible to mechanical stress, with eventual muscle fiber loss and replacement with fibrotic tissue (5,8).

DMD is the most common form of childhood muscular dystrophy, with an incidence of approximately 1:3,500 male births (9). Although a male inheritance pattern is typical, as many as one third of cases may be due to new mutations, without any previous family history. Typical initial symptoms include abnormal gait, frequent falls, and difficulty climbing steps. Hypotonia and delayed motor milestones occur in earlier onset cases, but in 75% to 80% of cases, onset is noted before age 4 (4). The abnormal gait is often noted by toe walking, which is a compensatory adaptation to knee extensor weakness, or increased lumbar lordosis as a compensation for hip extensor weakness. Another indication of pelvic girdle weakness is Gowers’ sign, demonstrated as the child rises from the floor. The patient generally begins by assuming a four-point stance, then brings the knees into extension while leaning the upper limbs forward, and sequentially moves the hands up the thighs until upright stance is achieved (Fig. 30-2A–D).

On examination, the earliest weakness is seen in the neck flexors, typically during the preschool years. Weakness of the proximal musculature of the shoulder and pelvic girdle is next, with steady progression, although the patient and family may feel that functional loss does not occur gradually but rather quite suddenly. This may relate to a critical point in weakness or range of motion when compensatory actions can no longer suffice to perform a task. Quantitative strength testing shows greater than 40% to 50% loss of strength by age 6 years (4), with fairly linear progression from ages 5 to 13 measured by manual muscle testing. Weakness appears to plateau after age 14 to 15, but this is probably a function of a floor effect and lack of sensitivity of the manual muscle testing scale (10,11).

Rehabilitation concerns are summarized in Table 30-6. In patients not aggressively treated, the average age to wheelchair use is 10, with a range of 7 to 13 years of age. Prediction of transition to wheelchair use may be helped by using timed motor performance tests. In one natural history study, all DMD subjects who took more than 12 seconds to ambulate 30 ft lost the ability to ambulate within 1 year (4). Immobilization, even for an acute illness, may lead to permanent loss of ambulatory ability during this phase of the disease.

<table>
<thead>
<tr>
<th>TABLE 30.6 Rehabilitation Concerns in DMD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maintaining mobility, range of motion, and strength during childhood</td>
</tr>
<tr>
<td>Progressive scoliosis</td>
</tr>
<tr>
<td>Progressive restrictive lung disease</td>
</tr>
<tr>
<td>Cardiac dysrhythmias and cardiomyopathy</td>
</tr>
<tr>
<td>Obesity (early adolescence) and cachexia (late adolescence)</td>
</tr>
<tr>
<td>Psychosocial adjustment and social interaction</td>
</tr>
</tbody>
</table>
FIGURE 30-2.  A–D: Gowers’ sign in an 8-year-old boy with DMD that is due to pelvic girdle weakness.
Unlike many myopathic disorders, joint contractures are a major concern in DMD. Nearly all affected boys older than 13 years have contractures (4,12,13), and these contractures most commonly occur first in the ankle plantar flexors, iliotibial bands, and hip flexors, with subsequent involvement of the knee flexors and elbow and wrist flexors. There does not appear to be a strong correlation between less than anti-gravity strength for a muscle group and the severity of joint contracture, nor for strength imbalance between antagonists across a joint (4). Clearly, lower extremity contractures become a problem after transition to a wheelchair for a significant part of the day. Natural history data suggest that progressive weakness, rather than heel-cord contractures, is associated with loss of ambulation as plantar-flexion contractures greater than 15 degrees are uncommon until after wheelchair reliance (4) (Fig. 30-3).

Scoliosis is a major clinical concern in DMD, and its prevalence is strongly related to age. Although significant curves often coincide with transition into wheelchair mobility, there does not appear to be a cause-and-effect relationship between scoliosis and wheelchair use (4,14). Rather, factors such as the adolescent growth spurt and progressive involvement of the trunk musculature may be responsible for progression of scoliosis during the adolescent years. There is some evidence that severity of scoliosis may be predicted by the type of curve and early pulmonary function measurements (15). When the curves do not involve significant kyphosis or hyperlordosis and peak forced vital capacity (FVC) is greater than 2 L, severe progressive scoliosis appears less likely.

It is now clear that bracing does not slow the progression of spinal deformity (12,16,17). Decision making for surgical management of scoliosis is closely related to pulmonary function. Although FVC volumes increase during the first decade of life close to 100% predicted with DMD, maximal static airway pressure (both maximal inspiratory and expiratory pressures) are impaired by 5 to 10 years of age. After a plateau in the early part of the second decade, there is progressive, fairly linear decline of FVC during adolescence (4). A higher peak FVC obtained at age 10 to 12 may predict less severe restrictive lung disease and spinal deformity developing over the next few years (4). An FVC below 40% predicted may contraindicate spinal instrumentation for scoliosis because of increased perioperative mortality; however, with current improved pulmonary care this is not an absolute contraindication (18). Symptomatic respiratory failure in DMD typically manifests in later adolescence. Management of this complication is covered more in detail at a later section.

It is not surprising that cardiac function is affected in DMD, because the dystrophin protein has been shown to be present in both the myocardium and Purkinje fibers (19). Most DMD patients older than age 13 demonstrate electrocardiogram (ECG) abnormalities (4). The first abnormalities noted are Q-waves in the lateral leads, followed by elevated ST segments, poor R-wave progression, increased R/S ratio, and resting tachycardia and conduction defects (4). ECG findings have been used to predict death from cardiomyopathy and include R wave in lead V1 less than 0.6 mV; R wave in lead V5 or V6 less than 1.1 mV; R wave in lead V5 less than 1.0 mV; abnormal T waves in leads II, III, aVF, V5, and V6; cardiac conduction defects; premature ventricular contractions; and sinus tachycardia (20). Sudden death from ventricular ectopy, a complication of the cardiomyopathy and left ventricular dysfunction, is well described in DMD (21,22). However, progressive congestive heart failure is a more frequent sequela, and some investigators estimate that 40% to 50% of DMD patients die from this complication (23,24). Cardiomyopathy is usually noted after 10 years of age and occurs in nearly all patients by age 18 (25). Cardiomyopathy is typically followed clinically with echocardiography, and the onset of systolic dysfunction is associated with a poor short-term prognosis (26). Once patients with DMD reach adolescence, regular screening with ECG, echocardiography, and Holter monitoring is prudent.

**FIGURE 30-3.** Brothers, ages 8 and 15, with DMD. In the older brother (left), note the presence of profound muscular wasting, scoliosis, and multiple joint contractures. The younger brother (right) demonstrates scapular retraction, increased lumbar lordosis, and stance phase plantar flexion (toe walking) to maintain a weight line posterior to the hip and anterior to the extended knee.
Considering the presence of a dystrophin isoform in brain tissue (27), it is not surprising that DMD patients show mildly decreased IQ scores compared with their peers and normative data (4). There may be a specific deficit with tasks requiring attention to complex verbal information, regardless of IQ (28). Mild impairments are noted on neuropsychological testing as well (4), without a specific area of strength or weakness.

Obesity from reduced physical activity is a major concern in DMD, particularly at the onset of wheelchair dependence (29,30). Since many patients are now placed on corticosteroid treatment, weight gain is the most frequently reported side effect. At later stages of the disease (ages 17 to 21), significant weight loss becomes the predominant nutritional concern (30,31). This probably results from nutritional compromise along with increased protein and calorie requirements during the later stages of DMD (32,33), partially as a result of the increased work of breathing from restrictive lung disease.

At this time, there is no curative treatment available for DMD. Oral corticosteroids have been shown to increase muscle mass, increase strength, and slow muscle deterioration. However, the mechanism of its action is still unclear. Recent studies demonstrate additional potential benefits of corticosteroids including amelioration of cardiac, pulmonary, and scoliosis complications in DMD (34–36). Research involving other pharmacogens that can increase muscle bulk and strength as well as research into the stem cell and gene therapy are ongoing.

**Becker Muscular Dystrophy**

BMD is similar to DMD as an X-linked recessive disorder. It has a similar pattern of muscle weakness, but generally presents with a later onset and a slower rate of progression (Table 30-7). Like DMD, the disorder has an abnormality in the gene location (Xp21) coding for the protein dystrophin. However, in this case, dystrophin levels are usually 20% to 80% of normal, or have the presence of the protein with an abnormal molecular weight. Mutation analysis of BMD has shown that majority are “in-frame” deletions, while DMD results from “frameshift” mutations. BMD is less common than DMD, with an overall prevalence recently estimated as 24 per million (37).

Without dystrophin analysis, it may be difficult to clinically discriminate between DMD and BMD. Although age of onset typically occurs later in BMD, there is significant overlap with DMD (38). The degree of CK elevation does not discriminate between the two diseases. The most useful clinical diagnostic discriminator is the ability to ambulate into adolescence. It is unusual for a patient with BMD to be wheelchair dependent before late adolescence, whereas even DMD outliers are dependent on the wheelchair for mobility by age 16. In fact, some BMD patients may be ambulatory well into middle age and beyond. There may be two distinct patterns of progression in BMD. In the first type, age of onset averages 7.7, and most patients have difficulty climbing stairs by age 20. In the more common milder form, age of onset averages 12, and there is no problem climbing stairs at age 20. The former group also seems to have a much higher rate of ECG abnormalities (39).

The percentage of normal dystrophin cannot be used to predict clinical course with any certainty in BMD (40).

Findings on examination of the BMD patient mirror DMD, although milder. The neck flexors and proximal lower limb muscles are affected early, particularly the hip and knee extendors (38). Subsequently, there is gradual involvement of the proximal upper-limb muscles (Fig. 30-4). Extensors are

<table>
<thead>
<tr>
<th>TABLE 30.7 Comparison of Clinical Features Between DMD and BMD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Age of onset</td>
</tr>
<tr>
<td>Pattern of weakness</td>
</tr>
<tr>
<td>Wheelchair dependence</td>
</tr>
<tr>
<td>Scoliosis</td>
</tr>
<tr>
<td>Cardiac involvement</td>
</tr>
<tr>
<td>Pulmonary dysfunction</td>
</tr>
<tr>
<td>Cognitive involvement</td>
</tr>
</tbody>
</table>

**FIGURE 30-4.** A 36-year-old man with BMD with pseudohypertrophy of the posterior deltoid and infraspinatus resulting in a posterior axillary depression sign.
generally weaker than flexors (38). Calf enlargement occurs, and presence of Gowers’ sign is indicative of the proximal muscle weakness. On standing, there is increased lumbar lordosis, and hip abductor weakness results in a waddling gait with trunk lean over the weight-bearing limb.

Contractures are not a significant early functional problem in BMD (38,39), becoming problematic only after wheelchair dependence. The joint locations for contractures are typical for one assuming the sitting posture, occurring in the hip flexors, knee flexors, and ankle plantar flexors. Significant scoliosis is much less common than DMD, and BMD patients rarely require spinal instrumentation (38,39).

One particular clinical concern in BMD is the potential for significant cardiac disease out of proportion to other manifestations of the myopathy (39,41–45). ECG abnormalities can be detected in about 75% of BMD patients (38,46). Most common abnormalities include abnormal Q-waves, right or left ventricular hypertrophy, right bundle branch block, and nonspecific T-wave changes. Echocardiography demonstrates left ventricular dilatation in 37% of BMD patients, and 63% have subnormal systolic function that is due to global cardiac hypokinesia (46). Cardiac transplantation may even be necessary in some patients (47,48). The degree of cardiac compromise may not be reflected by clinical symptoms, and these patients should be screened at regular intervals with ECG and echocardiographic studies.

Unlike DMD, significant pulmonary dysfunction is not a hallmark of BMD. FVC does not fall below the predicted level until the third to fifth decade of life. Because of relatively greater involvement of the intercostals and abdominal musculature compared with the diaphragm, maximum expiratory pressure is compromised at an earlier age than maximal inspiratory pressure (MIP), similar to DMD (38).

There are no consistent abnormalities on cognitive and neuropsychological testing in BMD other than a mild reduction in some patients (38).

Myotonic Muscular Dystrophy (DM1 and DM2)
There are two subtypes of myotonic muscular dystrophy, DM1 and DM2 (dystrophia myotonica type 1 and type 2). Both are muscular dystrophies that share similar clinical features of myotonia and distinctive effects on other organ systems. However, DM1 and DM2 are genetically separate entities with different genes involved. DM1 is caused by abnormal expansion of the CTG trinucleotide repeats in the dystrophia myotonica protein kinase (DMPK) gene on chromosome 19q13.3, while DM2 is caused by an abnormal expansion of the CCTG repeats in the zinc finger protein 9 (ZNF9) gene on chromosome 3q21 (49–51).

Myotonic muscular dystrophy type 1 (DM1) is the most common slowly progressive dystrophy in adults, with an incidence of 1/8,000 (9), while DM2 is much less common and thought to account for only about 2% of myotonic muscular dystrophy patients. Both are multisystem disorders affecting skeletal muscle, smooth muscle, myocardium, brain, and ocular structures (Table 30-8). This may manifest clinically with cataracts, cardiac conduction defects, endocrine abnormalities, swallowing dysfunction, and skeletal muscle weakness and myotonia.

Inheritance pattern for both DM1 and DM2 is autosomal dominant. As noted above, phenotype of DM1 results from abnormal expansion of CTG trinucleotide repeats within the DMPK gene (49,50). Normal individuals have fewer than 37 CTG repeats, whereas myotonic muscular dystrophy patients may have from 50 to several thousand repeats. The age of onset is inversely correlated with the number of repeats (49–51). With mild, late-onset myotonic muscular dystrophy, there may only be 50 to 150 repeats, whereas congenital myotonic muscular dystrophy may have greater than 1,000 repeats. The congenital form of myotonic muscular dystrophy is typically found only with DM1. The expanded region seems to further expand in subsequent generations, with increased severity of phenotype, in genetic anticipation. Currently, molecular genetic tests are commercially available for both DM1 and DM2 to determine the respective CTG or CCTG repeat expansions.

Patients with myotonic muscular dystrophy often have a characteristic facies. Long-standing myotonic muscular dystrophy is generally associated with a long, thin face, with temporal and masseter wasting. This is sometimes described as “lugubrious facies.” Ptosis can be prominent in these patients and contribute to the characteristic facies. Frontal balding at a young age is common in men.

The pattern of weakness in DM1 may be an unusual exception to the typical proximal greater than distal pattern in myopathies. The distal muscles may be affected to a greater extent, particularly early in the disease. This manifests as footdrop that is due to involvement of the ankle dorsiflexors, ankle invertors, and evertors and grip weakness from affected hand muscles (52). DM2 is different from DM1 in that proximal greater than distal distribution of weakness is noted. However, eventually, the neck, shoulder, and hip girdle muscles all become progressively weak in both DM1 and DM2. Significant contractures are unusual in myotonic muscular dystrophies, and scoliosis is usually a clinical problem only in congenital myotonic muscular dystrophy of DM1 (52).

A distinctive feature of myotonic muscular dystrophy among the dystrophies is the presence of myotonia, a state of delayed relaxation or prolonged contraction of muscle. This may be demonstrated by percussion of the thenar eminence with a reflex hammer, causing sustained flexion and adduction of the thumb. Grip myotonia is provoked by having the

<table>
<thead>
<tr>
<th>TABLE 30.8 Rehabilitation Concerns in Myotonic Muscular Dystrophies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Progressive weakness, often in a distal &gt; proximal distribution</td>
</tr>
<tr>
<td>Clinical myotonia, with difficulty in releasing grip</td>
</tr>
<tr>
<td>Cardiac conduction defects</td>
</tr>
<tr>
<td>Swallowing dysfunction</td>
</tr>
<tr>
<td>Cataracts</td>
</tr>
<tr>
<td>Nocturnal hypoventilation/sleep apnea</td>
</tr>
</tbody>
</table>

Nocturnal hypoventilation/sleep apnea
patient sustain a tight grip, then suddenly attempt to relax. Delayed opening of the fingers will occur. Myotonic “crescendo-decrescendo” discharges are also easily elicited with diagnostic needle EMG, although not specific for myotonic muscular dystrophy. Weakness and myotonia may exist in different proportions in an individual patient, with functional problems often from only one or the other manifestation of myotonic muscular dystrophy.

Cardiac abnormalities are common in both DM1 and DM2. Approximately 70% to 75% of patients demonstrate ECG and echocardiographic abnormalities (53). Cardiac conduction defects are the primary concern, with prolongations of the PR interval, abnormal axis, and infranodal conduction abnormalities leading to potential cardiac morbidity and the possibility of sudden death in a small percentage (54). A cardiac pacemaker may be indicated. Regular cardiac evaluation is recommended, particularly for those patients with dyspnea, palpitations, chest pain, or other cardiac symptoms.

When myotonic muscular dystrophy manifests in infancy as congenital myotonic muscular dystrophy, involvement of the respiratory musculature may cause significant respiratory distress. In noncongenital myotonic muscular dystrophy, restrictive lung disease causes significant morbidity later in life for many patients with DM1 (52). Nocturnal hypventilation and sleep apnea may occur, and clinicians should take a careful history to elicit symptoms common to these disorders: morning headache, frequent nightmares, excessive snoring, difficulty sleeping, and daytime somnolence.

Smooth muscle involvement most frequently manifests as difficulty swallowing and constipation, particularly in congenital myotonic muscular dystrophy. However, videofluoroscopy during swallowing reveals prolonged bolus transit times in adults with myotonic muscular dystrophy, placing them at risk for aspiration (55). Although a higher incidence of frank diabetes mellitus is controversial in myotonic muscular dystrophy, insulin insensitivity is a relatively common finding. Other endocrine abnormalities including primary gonadal dysfunction with infertility and thyroid dysfunction are also seen in patients with myotonic muscular dystrophy and should be evaluated when suspected. Posterior subcapsular cataract is present in almost all DM1 and DM2 patients and requires ophthalmological follow-up with eventual cataract removal if visual acuity worsens.

Cognitive deficits are most profound in the congenital myotonic muscular dystrophy population, with a reduced IQ often in the mentally retarded range (52,56). In noncongenital DM1, there is a wide range of scores with intelligence testing (52), and there appears to be some correlation between cognitive function and the number of CTG repeats at the myotonic muscular dystrophy gene locus.

Facioscapulohumeral Muscular Dystrophy

FSHD was identified as a distinctive muscular dystrophy because of the predilection for slowly progressive muscular weakness in the facial and shoulder girdle musculature. It is the second most common inherited muscular dystrophy in adult population with a prevalence estimate of 10 to 20 per million (9). It is nearly always inherited as an autosomal dominant disorder, with the chromosomal abnormality identified at the 4q35 gene locus (57), with reduced DNA fragment size at the telomere region (58). The reduction in DNA fragment size is due to deletions of a repeat sequence called D4Z4; however, the specific mechanism of disease has not yet been elucidated. Currently, a highly specific and sensitive genetic test is commercially available for the diagnosis of FSHD.

The distinctive clinical feature of FSHD is the presence of facial weakness, primarily involving the orbicularis oris, zygomaticus, and orbicularis oculi. These manifest as difficulty with eye closure (but not ptosis) and expressionless face (Fig. 30-5). The patient will typically have difficulty burying the lashes and pursing the lips, drinking through a straw, or whistling. Onset of symptoms is typically in adolescence or early adulthood. There are no distinctive findings on muscle biopsy, with histology frequently demonstrating mild findings of atrophied fibers.
along with hypertrophied fibers. Serum CK may be normal or only slightly elevated. Molecular genetic testing is available when the diagnosis requires confirmation.

In the shoulder girdle, the scapulae are typically displaced laterally and superiorly, resulting from combined weakness of the serratus anterior, rhomboids, latissimus dorsi, and lower trapezius (Fig. 30-6). Profound winging of the medial border of the scapula is common. Unlike most other muscular dystrophies, asymmetry of muscular involvement is common (59). Some authors feel this may be a manifestation of overwork weakness on the more affected side (60), but this is controversial. Although shoulder abductors and external rotators are typically involved, in some cases deltoid strength may be quite good if the scapula is stabilized. Elbow flexors and extensor involvement is common. In the lower extremities, the proximal musculature of the hip girdle is typically affected, often later in the disease course. However, there also appears to be a predilection for early involvement of the ankle dorsiflexors. About 20% of FSHD patients will eventually require either a power wheelchair or a scooter for mobility (61).

Coats’ syndrome is an early-onset variant of FSHD associated with sensorineural hearing loss. Muscle weakness typically begins in infancy, with progressive weakness and wheelchair dependence by late second or third decade. There is also an associated progressive exudative telangiectasia of the retina, requiring early recognition to prevent permanent visual loss. Interestingly, audiometric studies may demonstrate high-frequency hearing loss in more common forms of FSHD patients as well (62–64).

Significant contractures are uncommon in FSHD (59), and spinal deformity typically presents as hyperlordosis, scoliosis, or a combination. Although hyperlordosis to compensate for hip extensor weakness may be severe in some patients, the scoliosis rarely progresses to the point of requiring surgical instrumentation.

Cardiac complications in FSHD are rare, although some studies report cardiac abnormalities including cardiac fibrosis (65) and cardiac conduction defects (66). Mild restrictive lung disease is present in nearly half of patients with the expiratory muscles affected to a greater extent than inspiratory muscles, similar to other dystrophic disorders (59). In the absence of significant bulbar or respiratory involvement, the life expectancy of FSHD patients is normal.

**Limb-Girdle Muscular Dystrophies**

Before the advent of genetic testing, a group of patients commonly sharing a slowly progressive pattern of proximal greater than distal muscular weakness with either autosomal recessive (type 2) or autosomal dominant (type 1) inheritance were considered to have limb-girdle muscular dystrophies. Recent advances in molecular and genetic analyses have now identified a number of distinct genetic abnormalities with mutations in these patients. Currently, 17 subtypes of LGMD are recognized (67,68). Seven have autosomal dominant inheritance pattern (LGMD type 1, A to E) and ten with autosomal recessive inheritance (LGMD type 2, A to J). A detailed discussion of all subtypes of LGMD is beyond the scope of this chapter; however, the interested reader is directed to a recent review for more information (68). We still use the term limb-girdle muscular dystrophy as a general category of myopathy because of similarities in clinical presentation and progression.

It is important to keep in mind that “limb girdle weakness” is not unique to LGMDs. Other disorders such as BMD, myotonic muscular dystrophy type 2, EDMD, or late-onset spinal muscular atrophy should also be considered in the differential diagnosis. Determining the precise subtype of LGMDs is often difficult and may not be possible in majority of cases; however, a detailed symptom history, family history, mode of inheritance, and presence of associated clinical features may help narrow the differential diagnoses. At this time, genetic testing for few subtypes of LGMD is commercially available (www.genetests.org).

It is now known that some of the autosomal recessive LGMDs are associated with abnormalities in genes coding for components of the sarcoglycan complex, intimately involved
with the dystrophin protein in supporting the cytoskeletal structure of the muscle cell membrane (69). Although muscular weakness may eventually be profound, significant contractures and restrictive lung disease are unusual (70). In the United States, the LGMD 2C to F (sarcoglycanopathies) along with LGMD 2B (dysferlinopathy) and LGMD 2A (calpainopathy) are the most common limb-girdle dystrophies encountered; however, the distribution and prevalence of different subtypes of LGMD may differ around the world (68).

The cardiac involvement is rare and typically less than the dystrophinopathies such as DMD and BMD. However, depending on the LGMD subtypes (LGMD 1A to D, LGMD 2C to F, and LGMD 2L), it can be significant and warrants a close follow-up with cardiology consultation and regular ECG, echocardiography, and Holter monitoring. Respiratory involvement may also be a feature of the above subtypes of LGMD later in the disease course.

One subgroup of sarcoglycanopathy with a more rapid course is known as severe childhood autosomal recessive muscular dystrophy (SCARMD). These patients may mimic DMD early in the disease, with childhood onset and a similar pattern of weakness. However, it occurs equally in males and females and has a slower progression. Loss of ambulation generally occurs between 10 and 20 years.

**Nondystrophic Hereditary Myopathies**

**Congenital Myopathies**

The term congenital myopathy is used for a group of heterogeneous autosomal recessive disorders that typically present as hypotonia in infancy, can clearly be localized to muscle, and have weakness that is nonprogressive or slowly progressive. Frequently, there are delayed motor milestones in early childhood. Mental retardation may be profound in some syndromes. The weakness is usually in a proximal hip and shoulder girdle pattern. Because there is no progressive loss of muscle fibers, these disorders are not generally classified as dystrophic. The diagnosis for each specific disorder is based on histologic and electron microscopic muscle biopsy findings. Examples include central core myopathy, nemaline myopathy, and myotubular (centronuclear) myopathy.

**Metabolic Myopathies**

In patients with defects in glycogen, lipid, or purine metabolism, a common primary symptom is exercise intolerance or exertional muscle pain. There may be no fixed weakness on examination. Symptoms with activity include fatigue, myalgia, and muscle stiffness. When exercise is intense, patients may note a brownish-red color of the urine from myoglobinuria that is due to frank rhabdomyolysis. In some metabolic myopathies, progressive muscle weakness may predominate over dynamic symptoms and mimic a muscular dystrophy. Examples of metabolic myopathies include acid maltase deficiency (Pompe’s disease), myophosphorylase deficiency (McArdle’s disease), and carnitine deficiency.

Acid maltase deficiency (glycogen storage disease type 2) is a disorder of glycogen metabolism, with defects in the lysosomal acid maltase pathway. There are three autosomal recessive forms of acid maltase deficiency. The first presenting shortly after birth as hypotonia, with death by 2 years. The childhood-onset form is associated with delayed motor milestones and proximal weakness, with death by about age 20. In the adult variety of acid maltase deficiency, there is slow progression of proximal weakness beginning in the third or fourth decade. The clinical presentation may mimic limb-girdle dystrophy or PM. A potential treatment in the form of enzyme replacement therapy (ERT) is currently being evaluated in clinical trials. A guideline for diagnostic evaluation and management is available (71). With a potential treatment now on the horizon, it is important to consider the Pompe’s disease in the differential diagnosis of a patient with limb-girdle weakness, respiratory insufficiency without significant cardiac involvement, and findings suggestive of metabolic myopathy. Pompe’s disease represents one of only a handful of potentially treatable neuromuscular disorders at this time, and there is a commercially available diagnostic test.

McArdle’s disease is usually inherited in an autosomal recessive pattern, more commonly in males. Because of the inability to metabolize glycogen, there is exercise intolerance, easy fatigability, and stiffness of the exercised muscles. Physical findings may be normal between episodes, although some patients develop progressive proximal weakness. There are now protocols using nonischemic forearm exercise test in the evaluation of patients for McArdle’s disease (72).

Carnitine deficiency is the most common disorder of lipid metabolism. There are primary and secondary forms: The primary form has severely decreased plasma and tissue carnitine levels, and the secondary form has mild deficiency. Lack of carnitine impairs long-chain fatty acid metabolism in the mitochondria, thereby reducing energy production by the muscle cell. Oral L-carnitine supplementation may be helpful in some cases.

**ACQUIRED MYOPATHIES**

**Inflammatory Myopathies**

The hallmark of an inflammatory myopathy is the predominance of inflammatory cells on muscle biopsy. There are three primary types: DM, PM, and IBM. Although each is distinct, this group of myopathies is thought to involve immune-mediated processes possibly triggered by environmental factors in genetically susceptible individuals. DM and PM may be associated with disorders of the heart and lung as well as neoplasms. In addition, an inflammatory myopathy may be present as part of a multisystem disorder in other connective tissue diseases, most commonly scleroderma, systemic lupus erythematosus, mixed connective tissue disease, and Sjogren’s syndrome. Overall, the age of onset for idiopathic inflammatory myopathies is bimodal, with peaks between 10 and 15 years of age in children and between 45 and 60 years of age in adults. Women are affected twice as often, with the exception of IBM, which is twice as common in men. It is important to diagnose
accurately and in timely fashion for both DM and PM, since treatment is available and prognosis depends on early initiation of immunotherapy.

**Dermatomyositis**

Characteristic features of DM include muscle weakness that may present acutely, subacutely, or insidiously along with a characteristic rash. This violaceous, scaling rash typically involves the eyelids and occurs with periorbital edema, termed a *heliotrope rash*. Other common locations for the rash are the dorsum of the hands, extensor surfaces of the knees and elbows, and ankles. Myalgias may or may not be present. The weakness initially involves the proximal musculature and may progress to the distal muscles. Pharyngeal muscle involvement is evident from the frequent finding of dysphagia or dysphonia. Other manifestations include cardiac dysrhythmias and cardiomyopathy, joint arthralgias, and interstitial lung disease. In adults, there appears to be an association between DM and occult carcinoma, and a judicious workup for carcinoma is advisable in newly diagnosed patients.

Childhood DM differs somewhat from the adult version because of the higher incidence of vasculitis, ectopic calcification in the subcutaneous tissues or muscle, and lipodystrophy. Corticosteroids alone are often highly effective in both inducing a remission and preventing a recurrence and can usually be gradually withdrawn. Adults with DM do not respond to corticosteroids quite so predictably, and other immunosuppressive agents are often required. It may be difficult to fully discontinue pharmacologic treatment.

**Polymyositis**

The diagnosis of PM is often more difficult than DM, because no distinctive rash is present. It rarely occurs before age 20. Proximal limb and neck flexor muscle weakness presenting subacutely or insidiously should raise suspicion for PM. Myalgias are present in as many as one third of patients but are not generally the predominant symptom. CK elevation usually occurs at some point in the disease but may be normal in advanced cases with significant muscle atrophy. In general, serum CK level is a reasonable indicator of disease severity. Potential cardiac and pulmonary manifestations are similar to DM. Underlying carcinoma may less commonly occur than with DM. Treatment with corticosteroids supplemented by other immunosuppressive medications is the primary means of pharmacologic management.

**Inclusion Body Myositis**

A third type of inflammatory myopathy with a different pattern of involvement and course was identified in the 1970s. It was termed *inclusion body myositis* because of the presence of both inflammatory cells and vacuolated muscle fibers with nuclear and cytoplasmic fibrillary inclusions. IBM is now recognized as the most common myopathy in patients aged more than 50 years (73). Males are affected more than females. IBM has distinctive involvement of both proximal and distal musculature. In particular, the wrist and finger flexors are often more affected than the extensors, and the quadriceps may be affected out of proportion to other muscle groups. About one third have dysphagia, and the disease may be mistaken for amyotrophic lateral sclerosis because the age of onset is frequently after 50. IBM is relentlessly progressive in most cases, sometimes to the point of requiring a wheelchair for mobility. Unfortunately, it is not responsive to immunosuppressive medications, and treatment primarily involves appropriate assistive devices.

**Toxic Myopathies**

Medications and toxins can have various effects on muscle and its function. They can either directly injure muscle cells or indirectly via electrolyte disturbances, muscle ischemia, excessive muscle activity, and immune mechanisms. The muscle effects of medications and toxins may be focal or generalized, and it can occur acutely or after prolonged exposure.

**Corticosteroid Atrophy and “Steroid Myopathy”**

Although muscle weakness is an infrequent symptom of patients with endogenous hypercortisolism (Cushing’s syndrome), long-standing administration of exogenous corticosteroids is a common cause of proximal greater than distal muscular weakness and atrophy. Prednisone dosages higher than 30 mg/day increase vulnerability compared with lower dosages or an alternate-day regimen (74). The extent of weakness does not necessarily correlate with the duration of drug treatment. CK levels are generally normal or reduced, and muscle biopsy demonstrates type 2 greater than type 1 fiber atrophy. This is probably due to reduced muscle protein synthesis rather than increased catabolism. There is some evidence that resistance exercise training may help reduce or prevent steroid-induced myopathy (75).

**Medication-Associated Myopathies**

The HMG-CoA reductase inhibitors have been developed as treatment for hypercholesterolemia. With initiation of these cholesterol-lowering agents, often called collectively the “-statin” medications, transient myalgia and mildly elevated CK levels can be observed. However, the symptoms rarely progress to a necrotizing myopathy. When necrotizing myopathy does occur, acute to subacute proximal weakness, prominent muscle tenderness, and significantly elevated serum CK can be seen. In severe cases, myoglobinuria can also occur. Stopping the medication typically leads to resolution of symptoms and return of serum CK to normal levels within several weeks. Other cholesterol- and lipid-lowering agents that are fibric acid derivatives may cause myopathy. Amphiphilic agents (used as antimalarial and antirheumatic medications) and antimicrotubular agents (such as colchicines and vincristine) have also been shown to cause myopathies.

**Alcohol-Related Myopathies**

Alcohol-related myopathies are thought to be the most common among toxic myopathies. Muscle effects of alcohol are diverse and span from acute necrotizing myopathy with myoglobinuria to chronic atrophic myopathy. The common
symptoms of acute alcoholic myopathy are diffuse muscle cramps, myalgia, muscle swelling, and weakness. The exact mechanisms of alcohol-related myopathies are unknown at this time. However, severity of alcoholic myopathies appears to be related to a history of heavy, either acute or prolonged alcohol intake. Treatment is withdrawal of alcohol and correction of any electrolyte disturbances.

**Endocrine Myopathy**

**Thyroid Disorders**

Symptoms of muscle pain, cramps, or spasms are frequent in untreated hypothyroidism. However, only about one third of patients have demonstrable muscle weakness (76). In these patients, subtle proximal weakness with hyporeflexia and delayed relaxation of reflexes is evident. A PM-like syndrome with more profound proximal weakness and elevation of muscle enzymes may also occur (77). CK may be elevated 10- to 100-fold. Muscle histology is nonspecific, and the weakness resolves with appropriate thyroid hormone replacement. Thyrotoxicosis may present with symptoms of muscle weakness and proximal weakness, atrophy, and preserved or even brisk reflexes on examination. CK level is usually normal or low, and muscle biopsy may be normal or show predominant atrophy of type 2 fibers. The etiology is probably from increased catabolism of muscle tissue. Treatment of thyrotoxicosis improves the myopathy.

**Infectious Myopathy**

**HIV Infection**

HIV infection may be associated with an inflammatory myopathy or myositis, generally in patients with acquired immunodeficiency syndrome (AIDS). The clinical presentation is similar to idiopathic PM, with symmetric proximal weakness and elevation of serum CK. This disorder needs to be distinguished from the myotoxicity of zidovudine as well as the generalized weakness that may occur with advancing HIV infection.

**Myopathies Associated with Systemic Disease**

Many major illnesses are associated with muscle compromise, not generally demonstrable by manual muscle testing but resulting in poor muscle endurance and reduced motor performance with functional tasks. This is not generally classified as a “myopathic” disorder. However, certain systemic diseases or conditions do have a particular predilection to involve muscle and may be considered myopathies.

With severe illness, often in the ICU setting, patients can develop critical illness myopathy, critical illness polyneuropathy, or a combination of these syndromes (78,79). Critical illness myopathy may develop in patients who received high doses of corticosteroids with or without neuromuscular blocking agents or who have sepsis with multiorgan system failure. Necrotizing myopathy may be associated with underlying neoplasm, especially adenocarcinoma or the gastrointestinal system or lung cancer.

Electrolyte disturbances associated with systemic illness of various causes can produce myopathy. Hypokalemia is the most common electrolyte disturbance producing muscle weakness. This may be related to the use of diuretics or chronic diarrhea. Weakness is usually proximal, and CK elevation corresponds to the level of weakness. Frank rhabdomyolysis may be present, with muscle fiber necrosis noted on muscle biopsy. Clinical resolution occurs with correction of the hypokalemia. Other electrolyte abnormalities, which may be associated with muscle weakness, include hypocalcemia and hypercalcemia, hypophosphatemia, hyponatremia, and hypermagnesemia.

**REHABILITATION CONCERNS AND STRATEGIES IN MYOPATHIC DISORDERS**

**Weakness and Resistance Training**

The primary clinical manifestation of myopathies is muscle weakness. With the exception of certain inflammatory and metabolic myopathies, there is currently no effective pharmacologic management. Thus, the primary physiatric goal is to maintain strength, function, and independence. By using appropriate bracing, gait aids, and other assistive devices, functional mobility may be prolonged. Because of the proximal predilection of weakness, coming from sit to stand, managing stairs, and completing overhead activities such as dressing are primary functional problems.

An essential tool to maintain strength in most chronic diseases is resistance exercise. This is a controversial issue in myopathies (80,81). There is not clear evidence that persons with myopathy respond to strengthening exercise in a similar fashion as the able-bodied, and significant concern about overworking weak muscles prevents widespread acceptance by clinicians. Unfortunately, probably because of caution and concern by caring providers, this population often adopts a sedentary lifestyle, resulting in a component of disuse weakness (82).

Traditional concern about overwork weakness in myopathy primarily stems from case reports and circumstantial evidence noted in individual patients (60,83,84). However, newer knowledge about the function of dystrophin and related glycoproteins does provide additional concern about overwork. These proteins appear to be essential in maintaining the cytoskeletal framework of the muscle fiber during muscle contraction (69,85). In animal models of dystrophin-deficient dystrophy, there is increased damage to muscle using eccentric contractions, which particularly stress these cytoskeletal elements (86,87). Thus, it is conceivable that intense muscle contractions, particularly when including an eccentric component, may damage myopathic muscle to a greater extent than in the able-bodied. This is a particular concern in those diseases known to involve structural proteins of the muscle cell, such as DMD, BMD, and many of the limb-girdle syndromes. A recent study demonstrated a similar response to an acute bout of eccentric contractions in myopathic subjects and controls (88), but the effect over a longer time period is unknown.
The role of exercise differs between rapidly progressive disorders such as DMD and the more slowly progressive or static myopathies. In DMD, there is a rapid progression of strength loss when measured using qualitative (manual muscle testing) and quantitative (isokinetic dynamometry) methods (31,89). Considering this natural history, maintenance of strength would be the primary rehabilitation goal. Although hampered by methodological limitations, several investigations have demonstrated the ability to maintain or even slightly improve strength in DMD using resistance exercise (89–91). No protocol has demonstrated signs of overwork weakness. In DMD, resistance exercise activities are probably best encouraged by incorporating them into normal play and games of the child rather than a weight-lifting program (92).

In slowly progressive or static myopathies, the goal of resistance exercise is to increase strength, thereby giving the patient increased capacity to perform daily functions. A number of investigations combining patients with slowly progressive myopathies along with other peripheral neuromuscular disorders demonstrate modest benefits of strengthening exercise in slowly progressive disorders (93–97). Whether strength gains occur through direct hypertrophy of diseased muscle fibers or through reducing the effects of disuse weakness is not known. A systematic review of the literature found only two randomized, controlled trials of resistance exercise (in myotonic muscular dystrophy and FSHD) fulfilling inclusion criteria. The only conclusion drawn was that moderate-resistance strength training appeared to do no harm in these disorders, but there was insufficient evidence for benefit (98).

Therefore, unanswered questions include the appropriate regimen and whether increased strength translates into an improved ability to perform daily work tasks. There is some evidence that compared with a moderate-resistance strengthening program, a high-resistance training program offers no additional strength benefits (95,96). The moderate-resistance program included three sets of four to eight repetitions at 30% to 40% maximal strength of the knee extensors, with a similar regimen of 10% to 20% maximal strength training of the elbow flexors. No controlled investigations have demonstrated evidence of overwork weakness in muscular dystrophy.

Traditionally, patients with inflammatory myopathies were discouraged from physical activity due to fear of exacerbating muscle inflammation. Although sample sizes are small, more recent work suggests that moderate-intensity resistance exercise may improve strength and function without signs of increased muscle inflammation (99). Response to exercise may vary depending on disease activity, medications, and degree of disability. Patients with stable, chronic inflammatory myopathy may be able to tolerate more intensive strengthening regimens (ten maximal muscle contractions 3 days/week) without untoward effects (100).

Aerobic Training
Aerobic exercise training in myopathies has received scant research attention. Involvement of the cardiac and pulmonary musculature in both dystrophic and inflammatory myopathies may reduce cardiopulmonary fitness, compounding the effects of deconditioning. Decreased aerobic capacity compared with controls has been demonstrated in both adult patients with inflammatory myositis (101) and juvenile DM (102). Combining multiple neuromuscular diseases, several investigators have demonstrated improved oxygen uptake (103) or reduced heart rate at a submaximal workload after stationary bicycle training (104). Focusing specifically on patients with LGMD type 2L, moderate-intensity cycle endurance training was found to be safe and result in improved work capacity (105). Similarly, aerobic exercise in patients with FSHD and myotonic muscular dystrophy improved maximal oxygen uptake without signs of muscle damage (106,107). Improved oxygen uptake in subjects with inflammatory myositis was reported in two investigations combining strength and aerobic training (stationary cycling and step aerobics) not specifically designed to improve aerobic performance (108,109). There have been no randomized, controlled studies of aerobic exercise meeting Cochrane review criteria for recommendations in muscle disease (98).

Because maximal aerobic capacity is rarely the limiting factor in performing daily work tasks (110), improving muscular strength and endurance through resistance training will more likely enhance the ability to perform physical work tasks. However, an aerobic training program may help modify cardiovascular risk, because epidemiologic studies show that the physically inactive lifestyle so common in myopathic patients is associated with twice the risk of coronary artery disease in an able-bodied population (111). Swimming in warm water is a particularly useful activity to maintain aerobic capacity in myopathic patients.

Management of Cardiac Complications
Symptomatic cardiac complications associated with myopathies are primarily seen in DMD, BMD, myotonic muscular dystrophy (DM1 and DM2), EDMD, certain LGMDs, and mitochondrial myopathies (112,113). Asymptomatic electrocardiographic abnormalities are common in the inflammatory myopathies, although significant supraventricular arrhythmias, cardiomyopathy, and congestive heart failure may occur.

In managing patients with chronic myopathies, yearly ECG screening is warranted as it is generally the first cardiac test to become abnormal (114). The dystrophin protein is normally found in the Purkinje fibers of the heart, but its absence likely contributes to the rhythm abnormalities seen in the DMD and BMD patients. When the ECG is abnormal, echocardiography is indicated, also on a yearly basis, as well as involvement of the cardiology specialist. At times, cardiac conduction abnormalities and arrhythmias associated with myopathies can be severe enough to require a pacemaker, such as in DMD and BMD, myotonic muscular dystrophy, and EDMD. Cardiomyopathy is treated in the standard fashion, using angiotensin-converting enzymes when ejection fraction is less than 35% (115), as well as digitalis and diuretics with symptomatic heart failure (116). However, special caution with kaliuretic diuretics is warranted, because hypokalemia may exacerbate weakness (117,118).
The clinician should not confuse the presence of cor pulmonale that is due to chronic respiratory failure with intrinsic cardiomyopathy. Correction of hypoxemia and respiratory failure should be performed before treating cardiomyopathy. As detailed below, this usually requires mechanical ventilatory assistance rather than provision of supplemental oxygen, which may be inappropriate and perhaps even detrimental in patients with CO₂ retention.

Management of Pulmonary Complications
Breathing disorders are the leading cause of mortality in neuromuscular diseases (119). The causes of respiratory failure include direct respiratory muscle involvement from the skeletal myopathy, alteration in respiratory mechanics, poor secretion management, infections, and occasionally a problem with central control of respiration. However, with improved pulmonary care, patients with progressive neuromuscular diseases and pulmonary system involvement are living longer. In DMD, the average life expectancy has increased from 19 to 25 years, with increasing numbers of patients in their thirties (120).

Measurements of respiratory muscle strength and function have allowed clinicians to better determine the need for ventilation and cough assistance. Serial FVC measurements have demonstrated to be highly predictive of respiratory impairment and survival. An FVC of less than 1 L has shown to be the best negative predictor of survival in DMD, with mean survival of 3.1 years and 5-year survival of approximately 8% if ventilatory support is not provided (121). Other spirometric measurements including MIP, maximal expiratory pressure (MEP), and peak cough flow are also useful in the assessment of respiratory muscle weakness. When these values decline (peak cough flow <160 L/min or MEP <45 cm H₂O), they can indicate poor airway clearance function and hastened respiratory failure (122,123). Manual techniques or mechanical insufflator-exsufflators (cough-assist machines) can help in improving the airway clearance and secretion management. Intrapulmonary percussive devices and ventilators are also available to help mobilize secretions and improve pulmonary hygiene.

Progressive respiratory muscular weakness leads to restrictive lung disease manifesting as hypoventilation, hypercarbia, and ultimately respiratory failure when FVC dwindles to less than 30% predicted (124). The work of breathing is affected by the presence of pulmonary secretions, increased elastic and resistive loads, and kyphoscoliosis. In DMD, the presence of chronic hypercarbia and hypoxemia reduces central respiratory drive to these conditions, which may result in apnea during sleep and further worsening of the hypercapnia (125).

There may be a role for inspiratory muscle training through breathing against resistance in DMD. However, the current research data on the efficacy of respiratory muscle training are conflicting and no recommendation is available at this time. A study demonstrated improvement in MIP and 12-second maximal voluntary ventilation after 24 months of inspiratory muscle training (126). Others have also reported increased ventilatory strength and endurance (127–129).

However, several investigations found no significant improvements in respiratory function (130,131). In addition, it was recently found that nitric oxide release in exercising muscle as a protective mechanism is impaired in DMD (132). It has been postulated that already dystrophic muscle may be further damaged during respiratory muscle training protocols. If instituted, one should not wait until the later stages of the disease, when respiratory muscles are already working to near their fatigue threshold (133).

Other general measures for patients with restrictive lung disease include yearly influenza vaccination and a pneumococcal vaccination unless there are contraindications. Maintenance of an optimal body weight through proper nutrition (avoiding obesity or cachexia) is crucial. A consensus statement regarding respiratory care of patients with DMD is available (134).

Indications for Ventilatory Support
Signs and symptoms suggesting respiratory impairment include those related to sleep-disordered breathing (nightmares, morning headache, and daytime drowsiness) as well as respiratory dysfunction (exertional dyspnea, orthopnea, generalized fatigue, and paradoxical breathing patterns) (134–136) (Table 30-9). A polysomnography with continuous CO₂ monitoring is helpful in determining sleep-related hypoventilation associated with myopathies. However, a nocturnal pulse oximetry in the home environment can serve as a screening tool for sleep-related oxyhemoglobin desaturation and alveolar hypoventilation when polysomnography is unavailable (137).

Several studies have examined the indices useful to predict the need for ventilatory support. Vital capacity and MIP are generally the most useful pulmonary tests. Hypercapnia generally may occur when vital capacity falls to 500 to 700 mL (135), or with a vital capacity less than 55% predicted when the MIP is less than 30% predicted (124). When monitoring a patient with progressive myopathy, general warning signs for the impending need for mechanical ventilation include severe restrictive lung disease (vital capacity <45% predicted),

### TABLE 30.9 Clinical Features Suggesting the Need for Ventilatory Support in Myopathy

<table>
<thead>
<tr>
<th>Symptoms/signs</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Nightmares</td>
<td>Morning headache</td>
<td>Daytime drowsiness</td>
</tr>
<tr>
<td></td>
<td>Generalized fatigue</td>
<td>Dyspnea at rest</td>
</tr>
<tr>
<td></td>
<td>Orthopnea</td>
<td>Paradoxical breathing pattern</td>
</tr>
<tr>
<td>Pulmonary function testing</td>
<td>Vital capacity &lt;45–55% predicted</td>
<td>MIP &lt;30% predicted</td>
</tr>
<tr>
<td></td>
<td>Hypercapnia</td>
<td></td>
</tr>
</tbody>
</table>

[Table continues...]

...
respiratory muscle weakness (MIP <30% predicted), dyspnea at rest, and hypercapnia (138).

Methods of Ventilatory Support in Myopathies
Respiratory failure often presents insidiously in chronic myopathic disorders. Noninvasive methods of managing early respiratory failure are quite effective in maintaining adequate respiratory function and usually first instituted during the nocturnal period, with demonstrated improvements in PaO2 and reduction in PaCO2 (139–143). These benefits may be carrying over into periods off the ventilator and may result in improved daytime alertness and functional ability. Investigators report that ventilator-assisted individuals with DMD have similar life satisfaction scores to the general population (144) and similar levels of health perception and social function to ventilator users with nonprogressive conditions (145).

Noninvasive positive pressure ventilation (NIPPV) may be delivered via mouthpiece with or without a lip seal, via nasal mask, or via full-face mask. In any method, the most important factor is to obtain a good seal. This method is particularly convenient for nighttime use. Nocturnal NIPPV with bilevel positive airway pressure (BiPAP) has demonstrated its efficacy in use with sleep-disordered breathing and nighttime hypoventilation in DMD (146–148). In general, BiPAP mode of ventilation rather than the continuous positive airway pressure (CPAP) is appropriate for majority of restrictive lung volume processes as seen with myopathies. Frequent monitoring for adequate mask fit and appropriate ventilator pressure level settings is necessary.

Negative pressure ventilators apply a negative pressure to the surface of the chest and abdomen, expanding the chest wall and lungs. The cuirass, or chest shell, is the prototypical negative pressure device and is occasionally used but limited by the need for a snug fit to avoid air leaks. Some patients requiring intermittent ventilatory support may be amenable to the pneumobelt, an inflatable rubber bladder strapped around the abdomen. During inflation of the device, abdominal contents are displaced upward, resulting in upward displacement of the diaphragm and passive exhalation. When the device is deflated, descent of the diaphragm results in spontaneous inhalation. One drawback is that the patient must be seated at a greater than 30-degree angle from horizontal, making nocturnal use impractical. However, it may increase speech volume and allow time off other methods of ventilation.

Continuous invasive ventilatory support via tracheostomy should be considered when contraindications and patient aversion to noninvasive ventilation are present or when noninvasive ventilation is not feasible due to severe bulbar weakness or dysfunction. For those patients with myopathy requiring full-time ventilatory support, portable ventilators can be attached to power wheelchairs, markedly improving quality of life in the community.

Limb Contractures and Bracing
Many chronic myopathic disorders are associated with limb contractures in later stages of the disease, usually occurring with wheelchair dependence. Contractures may be myogenic, arthrogenic, or from soft-tissue shortening. Potentially important factors in the development of contractures in myopathies include (149) replacement of normal muscle tissue with collagen and fatty tissue, resulting in shortened muscle length; inability to obtain full range of motion of a joint; imbalance of agonist and antagonist muscle strength across a joint; static positioning in sitting; compensatory postural changes to stabilize joints for standing, most frequently seen in the equinus positioning of the ankles in DMD to keep the ground reaction force in front of the knee; and inability to obtain good muscle stretch in multijoint muscle groups, such as the hamstrings, tensor fasciae latae, and gastrocnemius.

Prevention and Management of Contractures
In some myopathies such as DMD, contractures are inevitable with wheelchair reliance, despite aggressive stretching programs. Range-of-motion programs are most useful in ambulatory patients with mild contractures. In DMD, an early stretching program has been shown to slow the progression of contractures (150–152). A minimum of 2 to 3 hours per day standing and walking may be necessary, along with a passive stretching program to avoid contracture development (153). Proper stretching technique is essential. The stretched position should be held for a slow count of 15, and each exercise should be repeated 10 to 15 times during a session. The gastrocnemius–soleus group, hamstrings, and iliotibial bands are particularly important to stretch in younger DMD patients (154).

Written instruction materials for the family effectively supplements verbal instruction. Physical modalities, such as heat to augment stretching, have not been adequately studied to recommend their use. Probably the most important component of contracture development in the weaker myopathic patient is the static position of the joint throughout each day and night.

The use of nighttime resting splints or ankle-foot orthoses to prevent plantar-flexor contractures is controversial. Some authors support their use (150), whereas others feel they are ineffective (154,155). A randomized trial in ambulatory DMD patients of stretching versus stretching and night resting splints demonstrated a 23% annual delay of heel-cord contracture in the combined group, although dropout rate was high (156). However, there is no evidence that delaying contracture development prolongs ambulation. Once contractures are present, early release of the hip flexors, iliotibial bands, and Achilles

<table>
<thead>
<tr>
<th>TABLE 30.10</th>
<th>Pros and Cons of Lower Extremity Bracing in DMD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pros</td>
<td>Cons</td>
</tr>
<tr>
<td>Reduced contractures during adolescence</td>
<td>Often requires surgical release of contractures before bracing</td>
</tr>
<tr>
<td>Prolonged standing during adolescence</td>
<td>No prolongation of functional ambulation</td>
</tr>
<tr>
<td>Possible psychologic benefits</td>
<td>Risk for falls</td>
</tr>
<tr>
<td></td>
<td>Need for ongoing physical therapy</td>
</tr>
</tbody>
</table>
In DMD, scoliosis affects 75% to 90% of nonambulatory patients (31,162). The prevalence of scoliosis has shown to be closely related to age and usually becomes noticeable between the ages of 10 and 14. Although the severity of the spinal curve increases with the length of time in wheelchair, no causal relationship has been established between scoliosis progression and wheelchair use. It is likely that both loss of ambulation and increasing spinal deformity represent disease severity and progression and are related to factors such as age, adolescent growth spurt, and weak truncal muscle. Routine screening for scoliosis with radiography before the transition to wheelchair is usually not productive. Once transitioned to wheelchair, scoliosis progression can be monitored with serial radiographs every 4 to 6 months. The reported degree of curve progression is variable (11 to 42 degrees per year) but can progress rapidly, so close follow-up is needed.

Spinal bracing in DMD does not change the natural history of the scoliosis, with continued progression despite brace usage (16,17,163). Modification of wheelchair seating system to slow the scoliosis progression has been ineffective (164). Other approaches, such as spinal exercises, manipulation, and electrical stimulation of paraspinal musculature, are also unproven. Recent studies in DMD suggest that the progression of scoliosis may be slowed with treatment with corticosteroids; however, greater side effects related to long-term corticosteroids use were also noted (34,162). At this time, the only effective treatment for progressive spinal deformity in severe myopathy remains surgical spinal instrumentation.

Indication for surgical intervention and optimal timing depends on the degree of scoliosis and cardiopulmonary status of the patient. It is generally agreed that surgical intervention should be sought when the Cobb angle is between 30 and 50 degrees (165–167). Preoperative pulmonary function testing is essential, and surgery should probably not be performed with FVC less than 30% predicted. Although a recent study demonstrated no clinically significant increase in operative and postoperative complications for those patients with FVC less than 30%, a careful preoperative risk assessment and planning as well as routine postoperative aggressive ventilatory support are stressed to minimize morbidity (168). Nocturnal pulse oximetry can provide valuable information about potential postoperative ventilation need. In those patients at risk, preoperative mask fitting and initiation of nocturnal NIPPV can improve postoperative respiratory recovery. Cardiology consultation, including preoperative ECG and echocardiography, is also important to evaluate for conduction abnormalities and cardiomyopathy. Impaired left ventricular function and cardiomyopathy are risk factors for dysrhythmias and sudden death. A potential complication of malignant hyperthermia during anesthesia is not specific to DMD. It can also occur in other myopathies and should be kept in mind for those patients undergoing anesthesia. Postoperative management after scoliosis surgery includes early involvement of therapies, mobilization out of bed when clinically stable, pain control, ventilatory support as needed, and appropriate pulmonary toilet.

Although some improvement in spinal deformity may be expected with surgical instrumentation, it is still unclear whether the procedure leads to improved or sustained pulmonary function. Progressive loss of vital capacity in DMD results from disease of the respiratory musculature, although there is probably a component from the scoliosis as well. Some studies support preserved pulmonary function and survival (17),

**Bracing for Ambulation**

Because of the proximal predilection of weakness in most myopathies, bracing for ambulation becomes much more complex and controversial than distal disorders. The greatest controversy is with DMD (Table 30-10). At the time when upright ambulation becomes difficult in DMD, hip and knee flexion contractures are usually mild (31,150), but the heel cords and iliobial bands often require surgical release if braced ambulation is to be successful (158,159). Bracing then consists of bilateral knee-ankle-foot orthoses with solid ankles set at 90 degrees, drop-lock knee joints, and ischial weight-bearing upper-thigh components to help maintain the upright posture with reduced lumbar lordosis.

Clearly, continued standing and limited ambulating with the braces provide good stretch to the hip, knee, and ankle musculature. Heel-cord and knee flexion contractures are less pronounced at age 16 in patients who continue daily standing (154). Although reported in the literature (160), it is not clearly known if scoliosis becomes less pronounced with the use of lower extremity bracing. However, a slowing in the progression of contractures by itself does not prolong standing and ambulation, because weakness, not contractures, is primarily responsible for the transition to wheelchair reliance (149). In addition, using braces requires significant energy expenditure on the part of the patient, requires ongoing physical therapy, and increases the risk of falls. Clearly, quality of life issues require detailed discussions with the patient and family before proceeding with aggressive surgical and orthotic management in rapidly progressing muscular dystrophy. A systematic review of the effectiveness of knee-ankle-foot orthoses in DMD reports probable prolongation of standing, but not necessarily functional ambulation (161).

**Management of Spinal Deformity**

Spinal deformity in neuromuscular diseases and myopathies can be very severe and progress to result in multiple problems. Severe scoliosis and pelvic obliquity can lead to pain, poor sitting balance and upright seating position, difficulty in attendant care, skin ulcers, and potential exacerbation of restrictive lung disease. Although various different myopathies can be associated with progressive spinal deformity, the most common are those that are in the hereditary myopathy category and those that are relatively more severe in disease phenotype. These include DMD, congenital muscular dystrophy (CMD), congenital myotonic muscular dystrophy, and SCARMID. Scoliosis may also be an occasional concern in facioscapulohumeral muscular dystrophy and congenital myopathies.

In DMD, scoliosis affects 75% to 90% of nonambulatory patients (31,162). The prevalence of scoliosis has shown to be closely related to age and usually becomes noticeable between the ages of 10 and 14. Although the severity of the spinal curve increases with the length of time in wheelchair, no causal relationship has been established between scoliosis progression and the length of time in wheelchair, no causal relationship has been established between scoliosis progression and differential care, skin ulcers, and potential exacerbation of restrictive lung disease. Although various different myopathies can be associated with progressive spinal deformity, the most common are those that are in the hereditary myopathy category and those that are relatively more severe in disease phenotype. These include DMD, congenital muscular dystrophy (CMD), congenital myotonic muscular dystrophy, and SCARMID. Scoliosis may also be an occasional concern in facioscapulohumeral muscular dystrophy and congenital myopathies.

In DMD, scoliosis affects 75% to 90% of nonambulatory patients (31,162). The prevalence of scoliosis has shown to be closely related to age and usually becomes noticeable between the ages of 10 and 14. Although the severity of the spinal curve increases with the length of time in wheelchair, no causal relationship has been established between scoliosis progression and differential care, skin ulcers, and potential exacerbation of restrictive lung disease. Although various different myopathies can be associated with progressive spinal deformity, the most common are those that are in the hereditary myopathy category and those that are relatively more severe in disease phenotype. These include DMD, congenital muscular dystrophy (CMD), congenital myotonic muscular dystrophy, and SCARMID. Scoliosis may also be an occasional concern in facioscapulohumeral muscular dystrophy and congenital myopathies.
whereas others note no change in pulmonary decline and age at death in their DMD patients (31,169). It is important to keep in mind that the primary goals for prevention of scoliosis and corrective surgery are to provide the patients with improved sitting balance, which can help to maintain use of wheelchair, facilitate nursing care, and enhance their quality of life.

Body Composition and Metabolic Issues
Loss of skeletal muscle, gain of excess body fat, and changes in physical activity and energy metabolism are common to both rapidly progressive and slowly progressive myopathies. Excess body weight places greater burden on already weakened skeletal muscle, reduces mobility, and may increase the work of breathing. In one study on DMD, 40% of patients were overweight (weight-for-age >90th percentile) between ages 9 and 17, 40% were underweight (weight-for-age <10th percentile) between ages 13 and 17, whereas 65% were underweight older than age 17 (31). This supports a divergence in weight in early adolescence, but a tendency toward underweight in later stages of the disease. Because of the replacement of muscle tissue by fat, at a given weight-for-age, the DMD individual will have a lower percentage of lean tissue. Although it is unclear why only certain individuals with DMD develop obesity, the weight loss in later stages of DMD may be from higher energy and protein requirements (32) with inadequate caloric intake to meet nutritional needs. Placement of a percutaneous feeding tube may be beneficial to provide rapid provision of calories and fluids, even if the patient can still swallow. Oral intake for enjoyment is still possible. Adequate nutritional intake may be a concern in slowly progressive myopathies as well. In a study of adults with nondystrophin-related muscular dystrophies, there was inadequate intake of various micronutrients and macronutrients, with significant correlations between copper, water-soluble vitamins, and strength measures (170).

The relationships between metabolism, energy expenditure, and chronic disease risk are now being considered in persons with neuromuscular diseases. Several studies have shown a lower basal metabolic rate for patients with slowly progressive myopathies (82,171,172). In a study combining patients with a number of slowly progressive neuromuscular diseases, the neuromuscular disease subjects were shown to have reduced 24-hour energy expenditure as a result of reduced physical activity and a higher percentage of fat mass than controls (82). A similar population was found to have a high risk for developing chronic diseases resulting from obesity and a sedentary lifestyle (173). Using pedometer-derived activity goals combined with dietary intervention, persons with slowly progressive neuromuscular diseases can increase physical activity and reduce caloric intake, although improved health outcomes have not yet been shown (174).

Quality of Life/Psychosocial Issues
A survey on patients with neuromuscular diseases, including chronic myopathies, supports the notion that as a group, the quality of life of individuals with neuromuscular disease in various disease stages is not much different from that of the able-bodied (175). However, analysis of certain components of quality of life is of interest. Although level of impairment and disability did not predict life satisfaction, certain factors such as lack of information about the disease, poor coordination of services, negative attitudes, and decreased expectations of their potential were identified as important issues. The ability to attain independence, either through themselves or through personal care assistants, was an important key to quality of life. Support groups often serve as an excellent resource for psychological support and identifying other resources in the community. A recent study demonstrates that those patients who are single, childless, and with an earlier onset neuromuscular disease may have their quality of life impacted more by their impairments and disabilities than previously thought (176). Psychosocial issues are summarized in Table 30-11. It is important for the rehabilitation specialist to be attuned to these factors in their care of patients with myopathies, so that therapeutic interventions can be timely and effective.

Clinicians may underestimate the presence of pain in their patients with myopathies. The causes of pain are multifactorial and may be treatable, often related to joint degenerative conditions and myofascial origin rather than directly from the myopathy. A survey in slowly progressive neuromuscular diseases found that the frequency and severity of pain reported in these disorders were comparable to persons with osteoarthritis and chronic low back pain (177). Appropriate assessment and treatment of pain may significantly improve quality of life and should be part of overall physiatric management.

In a recent survey of employment among those with neuromuscular diseases including muscular dystrophies, the major barrier to employment was lack of education (178). Level of disability and physical impairment were not critical factors. Unemployment is a particular problem in myotonic muscular dystrophy, possibly related to nonphysical factors. Although much more work needs to be done to determine ways to improve quality of life with chronic diseases such as myopathy, it is incumbent on the rehabilitation specialist to see the person beyond the disease, one in which others may view as “untreatable.” There are always interventions to be made to improve quality of life, even diseases that have no cure.

CONCLUSION
Although great strides in genetics and molecular biology have helped us improve our understanding of many muscle
diseases, the treatment for these and most other neuromuscular diseases remains supportive rather than curative. An important component of this treatment involves appropriate and judicious use of physical activity, assistive devices, and other common physiatric interventions. Knowledge of these diseases also provides the basis of expectant management to prevent secondary complications. With this knowledge, the physician can orchestrate a rehabilitation plan matching the goals of the patient and family. With proper care, many persons with myopathy have the potential to have a reasonable quality of life.

ACKNOWLEDGMENT

Supported in part by Grant H133B03111805 from the National Institute of Disability and Rehabilitation Research (NIDRR). The authors would like to thank Elaine Shirley for manuscript preparations.

REFERENCES


Osteoarthritis

INTRODUCTION

Physiatrists, particularly those in the outpatient setting, will likely increasingly take care of patients with osteoarthritis (OA). This is due to both the growing prevalence of OA within the general population and the recognition by physiatrists, other physicians, and the general public that physiatrists possess the knowledge base and skill set to manage patients with various musculoskeletal conditions. Traditional physiatric training emphasizes function, therapeutic exercise, orthoses and assistive devices, and coordination of care with other health care professionals, especially physical therapists, occupational therapists, and social workers. Training nowadays often also includes instruction in musculoskeletal injection procedures and pain management pharmacotherapy. This combination of traditional physical medicine and rehabilitation along with interventional and pharmacologic pain management techniques places the physiatrist in a position of being able to provide the full spectrum of nonsurgical patient care.

Although the field of PM&R has not yet drafted specific management guidelines analogous to those published by the American College of Rheumatology (ACR) (1), the ACR guidelines center around many basic nonpharmacologic principles that are integral to physiatric training. Therefore, part of the nonpharmacologic treatment section of this chapter is loosely structured based upon these management strategies. The pharmacologic treatment section of this chapter is also based upon the ACR guidelines as medication recommendations were derived in part from this evidence-based medical research. Of note, the ACR guidelines are scheduled to be undergoing revision under the guidance of a subcommittee that now includes a physiatrist. These revised guidelines are scheduled to be completed during 2010.

Although physiatrists are also frequently involved in the postsurgical rehabilitation of patients who have undergone arthroplasty, this topic is covered elsewhere in this textbook.

PATHOLOGY AND PATHOPHYSIOLOGY

OA is really a group of disorders with different etiologies but similar pathologic changes. In primary OA, it is believed that excessive loads cause failure of an otherwise normal joint (2). The changes eventually involve all of the tissues including the articular cartilage, subchondral bone, synovial tissue, joint capsule, ligaments, and muscles that act on the joint. Histologically, small tears known as fibrillations and larger tears known as clefts both develop. These defects begin in the superficial zone of cartilage, extend into the transitional zone, and are also propagated by enzymatic breakdown of cartilage (3,4). Eventually, large areas of cartilage loss occur, thus essentially exposing the underlying subchondral bone. Other changes within the cartilage include an eventual decline in the ability of the chondrocytes to replicate, an initial increase in the water content of cartilage, a significant reduction in proteoglycan content, a reduction in the size of type II collagen fibers, shortening of the glycosaminoglycan chains, a diminution of the keratin sulfate concentration, and an increase in the proportion of chondroitin-4-sulfate consistent with immature cartilage being produced in an attempt to regenerate the lost cartilage (5). Cartilage matrix changes lead to increased matrix permeability and decreased matrix stiffness, both of which may increase tissue vulnerability to additional mechanical damage. Although the chondrocytes within the osteoarthritic joint acquire the ability to replicate and produce new chondrocytes that are very metabolically active, they produce collagen, proteoglycan, and hyaluronan (HA) that do not aggregate well and are not adequately stabilized in the extracellular matrix.

Eventually, this chondrocytic repair response declines, although the exact reasons for this are poorly understood. Theories, however, include failure of the matrix to stabilize and protect the chondrocytes and a down-regulation of the chondrocytic response to anabolic cytokines. The order of appearance of the changes in the cartilage and underlying subchondral bone is debatable. Specifically, one theory is that the subchondral bone undergoes remodeling in response to the external load applied to it due to cartilage breakdown (6). In contrast, Radin hypothesized that subchondral bone stiffening associated with remodeling in response to externally applied loads precedes and subsequently promotes cartilage loss (7). Fibrous, cartilaginous, and osseous prominences known as osteophytes eventually develop around the periphery of the joints (marginal osteophytes) but can also develop along joint capsule insertions (capsular osteophytes) or protrude from the degenerating joint surfaces (central osteophytes). Cystlike bone cavities containing myxoid, fibrous, or cartilaginous tissue eventually form within the bone. Regardless of the exact sequence of events, both cartilage degeneration and subchondral bone remodeling are found by the time the patient is symptomatic (8). Changes also occur within the synovial membrane and the synovial fluid.
Specifically, the synovial membrane may have a mild to moderate inflammatory reaction and may contain fragments of articular cartilage (9). Within the synovial fluid, pathological changes include significant alteration in synovial fluid HA, including a decrease in the concentration of normal molecular weight hyaluronate and the production of abnormal hyaluronate. The resultant decreased hyaluronate concentration is a result of both defective hyaluronate production and increased hyaluronate breakdown. In addition, there is increased water content and an increased concentration of inflammatory mediators (10). These pathological changes result in defective synovial fluid viscosity, elasticity, barrier exclusion and shielding. Exposure of synovial nociceptors perhaps explains in part the pain associated with the osteoarthritic joint.

EPIDEMIOLOGY

More than 21% of US adults (46.4 million persons) have been reported as having “self-reported doctor-diagnosed arthritis” — for example, the patients themselves reported that a physician had previously diagnosed them with arthritis. Based on cognitive and validation studies, “self-reported doctor-diagnosed arthritis” is thought to provide the most credible estimate of overall arthritis prevalence, with acceptable sensitivity and specificity for surveillance purposes (11). As the average age of the US population increases, an estimated 67 million Americans will acquire some form of arthritis by the year 2030 (12). Among various types of arthritis, OA is the most common and affects 27 million people in the United States (13). Approximately 10% to 30% of those affected with OA have significant pain and disability (10). OA is second only to ischemic heart disease as a cause of work disability in men over age 50 (14).

Although incidence and prevalence may vary widely according to the type of epidemiologic studies, as well as by whether clinical or radiological definitions are used, the overall epidemiology of OA has been described. At least 37% (and up to 68% in some studies) of persons 60 years and older have radiographic evidence of knee OA. By contrast, 27.2% of adults 26 years and older have radiographic evidence of hand OA. Clinically, 12.1% of adults aged 60 or older have symptomatic knee OA. The prevalence of symptomatic hand OA was higher in women (9.2%) than in men (3.8%) (10,15,16).

The causes of OA are multifactorial and the risk factors may differ for each joint. Given the anticipated increase in OA prevalence, understanding the risk factors associated with OA can aid in further clarifying the disease process as well as helping to potentially delay OA progression.

RISK FACTORS

Genetic Factors

Multiple genes appear to confer an increased risk for OA, and the contribution of heritability and possibly genetics to the occurrence of OA varies by joint. The genetic predisposition to OA seems to be based on a variety of physiologic vulnerabilities including changes in collagen and enzymes within cartilage, variations in cytokines or growth factor profiles in cartilage, and genes that dictate joint shape and structure (16,17).

Age and Gender

The prevalence and incidence of OA correlates markedly with age; however, age alone does not appear to cause OA. Rather, the increase in OA with age is likely a consequence of biologic alterations that occur with the aging process. These include decreased strength, slower neurologic responses with a decline in proprioception, a decreased response of chondrocytes by growth factors, and age-related accumulation of glycation end products, which is a haphazard process impairing the function of the cartilage (18).

Younger adults also develop OA secondarily from injury and loss of biomechanical integrity. Data show that a greater number of men present with OA before age 50; while after age 50, the condition is more common in women (19). Women experience greater severity of symptoms and report more problems with morning stiffness, joint swelling, and night pain. Women also have a higher incidence of Heberden’s nodes (swelling of the distal interphalangeal [DIP] joint of the hands), a phenomenon with a familial tendency.

Ethnicity

African Americans demonstrate a greater severity rating of knee OA than Caucasians (20). The prevalence of hip OA is higher in European Caucasians than in African-Americans. Hip OA is rare in China and in those of Chinese descent in the United States (21,22). Anatomic abnormalities that are prevalent in the United States are rare in the hips of persons of Chinese descent, which may indicate that genetic predisposition to developmental abnormalities is a factor in this ethnic variation (23). However, it may also be related to other variables including variations in body mass index (BMI), nutritional factors, and the impact of lifestyle differences and health care disparities between populations.

Obesity

Obesity is a risk factor for the development of OA. The reason why increased body weight is associated with OA remains somewhat uncertain but may involve cartilage degeneration via excessive load bearing (24). Another mechanical factor of knee OA in overweight individuals appears to be knee alignment. In patients with varus knees, BMI appears to be related to OA severity, especially medial tibiofemoral OA (25).

In the Framingham cohort study, the BMI measured at entry into the study predicted the presence of radiographic knee OA 36 years later. It also found there was strong association between body weight and knee OA, especially in women; for women in the most overweight, relative risk was 2.07, and for men, it was 1.51 (26).

In contrast to the knee, a more modest association between body weight and bilateral hip OA has been described (27).
Various factors including abnormalities that are unrelated to obesity; adiposity lying below the hip joint, which does not contribute to loading; and distribution of body weight load more broadly may be an explanation of less clear relationship between obesity and hip OA (16).

Some studies have shown that there is an association between obesity and hand OA, with BMI being directly proportional to carpometacarpal (CMC) OA in both genders, suggesting that obesity may predispose to OA, perhaps via an inflammatory or metabolic intermediary that has not yet been identified (28–30).

**Occupation**

Occupations characterized by repetitive joint loading (overuse) or high physical demands may increase the risk of developing OA. For example, farmers are at high risk for hip OA, jackhammer operators for shoulder and elbow OA, miners for knee and spine OA, and dockyard workers for knee and finger OA (16). These associations especially correlate in the setting of injury.

A systematic international review and synthesis (OASIS) states with a high degree of scientific evidence that sports and recreational activities are a risk factor for knee and hip OA and that the risk correlates with intensity and duration of exposure. But the risk is lesser than that associated with a history of trauma and overweight (31). Early diagnosis and treatment of sports-related injuries, with a goal of maintaining joint-surface integrity, should help decrease the subsequent risk of developing OA at the injured joint (32).

**Muscle Weakness**

Poor quadriceps strength may predispose individuals to the development of OA (33). In a longitudinal study, women who had no radiographic evidence of OA initially but who had knee-extensor weakness were more likely to develop OA than women with no initial weakness (34). On the contrary, the progression of knee OA is more likely associated with the factors such as level of hyaluronic acid and generalized OA rather than quadriceps weakness (35,36).

**JOINT BIOMECHANICS**

A deviation from normal joint biomechanics amplifies joint vulnerability leading to OA. Biomechanical alteration may include disruption or incongruity of the articular surface, dysplasia, malalignment, instability, disturbance of innervation of the joint, ligament and muscles, and inadequate muscle strength or endurance.

Knee OA with varus alignment increases the risk of medial joint disease progression, and knee OA with valgus alignment increases the risk of lateral joint space progression (37). In a study of varus-aligned knees, a thrust was associated with a threefold increase in the likelihood of OA progression (38).

A higher incidence of varus-valgus laxity is seen bilaterally in the knees of those with OA (both the involved and the nonarthritic knees), suggesting that laxity may predate disease development and contribute to the disease process (39). Impaired proprioception has been seen in patients with OA compared with age-matched controls, which may also indicate that loss of proprioception precedes disease development (40).

The role of mechanical loading in the development of OA is clearer. Moderate, intermittent cyclic joint loading has been shown to be beneficial and essential to healthy joint function, but continuous compression of the cartilage suppresses metabolic activity including collagen and proteoglycan synthesis and causes tissue damage. Joint immobilization has also been shown to be detrimental, reducing cartilage thickness and proteoglycan content. In addition, intense exercise or a sudden increase in exercise, particularly in older persons, produces catabolic changes in cartilage (41,42).

**SYNOVIAL JOINTS**

**Normal Anatomy/Physiology**

The normal synovial joint consists of articular cartilage, subchondral bone, the synovial membrane, synovial fluid, and the joint capsule (Table 31–1). Some synovial joints also have labral tissue, interosseous ligaments, menisci, and fat pads. Most synovial joints are supported by the periarticular muscles, tendons, and ligaments. These structures are all important in ensuring proper joint function.

Normal cartilage has a surface zone, a middle zone, a deep zone, and then a zone of calcified cartilage, where the cartilage attaches to the underlying subchondral bone. In the surface zone, the collagen content is highest and the collagen fibers are oriented parallel to the joint surface. In the middle zone, collagen fibers are oriented in multiple directions, and proteoglycan content is increased. In the deep zone, collagen fibers are oriented perpendicular to the articular surface. Collagen fibers attach to the subchondral bone after a transition through the zone of calcified cartilage (43).

The articular cartilage has several functions including friction reduction, shock absorption, and the spread and transmission of weight loads to the underlying bone. Articular cartilage is composed of an extracellular matrix and chondrocytes. Under normal circumstances, there is a balance of repair and degradation of the cartilage. In OA, this balance is disrupted resulting in altered matrix and chondrocyte. Cartilage nutrition is maintained by diffusion of nutrients from the synovial fluid and facilitated imbibition (the process by which cartilage absorbs synovial fluid in response to sequential loading and unloading).

The subchondral bone also plays a role in normal joint protection as a shock absorber. The interdigitation of the cartilage and bone between the deepest layer of cartilage and the subchondral bone plate (cortical end plate) serves to transform shear forces into tensile and compressive stresses. Subchondral bone can attenuate about 30% of the loads through the joint, whereas articular cartilage attenuates only 1% to 3% of load forces (32,44). In addition to its shock-absorbing function, the subchondral bone plays a supportive role in maintaining
the joint environment. The subchondral bone contains not only bone marrow and trabecular bone but also end arteries and veins. The subchondral bone has marked porosity, with vessels penetrating the calcified cartilage zone. These help provide nutrients to the cartilage and facilitate the removal of metabolic waste products. The perfusion of these vessels accounts for more than 50% of the glucose, oxygen, and water requirements of the cartilage (44). The underlying bone also undergoes turnover as per Wolff’s law that states that the bone remodels in response to externally applied forces. Thus, loads applied to the joint also help the subchondral bone to remain metabolically active.

The synovial membrane is another structure that provides joint protection. The synovial membrane is normally 1- to 2-cell layers thick. The cells themselves are known as synoviocytes and are responsible for the formation of synovial fluid via plasma ultrafiltration and production of hyaluronate specifically by type B synoviocytes. The synovial fluid undergoes the normal process of turnover, as it is removed from the joint via diffusion into the surrounding lymphatics and capillaries. Synovial fluid provides nutrient support to the periarticular cartilage and acts as either a “joint lubricant” or a “shock absorber” depending upon the load imposed upon the joint. This differential action is known as viscoelasticity and can be understood by the rate of change of an externally applied load. Specifically, under conditions of a slowly changing load, such as slow range of motion (ROM) of the joint, the synovial fluid molecules line up so as to act as a viscous liquid (aka joint lubricant) and dissipate the externally applied energy as heat. By contrast, under high load conditions, such as stepping off of a curb, the molecules within the synovial fluid act as an elastic liquid by absorbing the externally applied energy. Another function of synovial fluid is referred to as barrier exclusion by which the movement of inflammatory cells (e.g., neutrophils) and molecular debris (e.g., fragmented cartilage know as “joint mice”) within the joint is limited due to the overlap of the large molecular weight hyaluronate molecules. Finally, shielding effects refer to the coating of articular nociceptors, thereby helping to prevent pain receptor binding by inflammatory mediators (e.g., prostaglandins and bradykinins), and the coating of synoviocytes, thereby decreasing synovial membrane permeability.

The surrounding joint capsule is composed of dense connective tissue containing collagen bundles and some elastic fibers. The capsule encloses the joint and provides stability and limits ROM. The periarticular structures include the surrounding muscles and their tendinous insertions onto the joint. The muscles and tendons function to move the joint and are also theorized to protect the joint by providing support.
Overall, the joint is protected by several different tissues that serve as shock absorbers including articular cartilage, subchondral bone, synovial fluid, and the periarticular muscles/tendinous insertions. Since articular cartilage is only 1 to 2 mm thick in most regions, it is too thin to serve as the sole shock absorber. A defect, therefore, in any of the adjunctive shock absorbers contributes to degeneration of the joint, as is discussed below.

**INTRODUCTION TO NONPHARMACOLOGIC TREATMENT**

A significant amount of research has been performed on nonpharmacologic treatment interventions for OA. Several of these have been validated by prospective studies and are part of the updated 2000 ACR guidelines for the medical management of hip and knee OA (1). The following 12 key nonpharmacological modalities appear in the guidelines: patient education, self-management, social support, weight loss, aerobic exercise, physical therapy (PT) ROM exercises, muscle-strengthening exercises, assistive devices for ambulation, patellar taping, appropriate footwear, lateral wedged insoles, and occupational therapy (OT) (Table 31-2).

There is relatively good evidence to support the efficacy of nonpharmacologic interventions. Unfortunately, compliance with one of the most effective interventions, therapeutic exercise is often poor as has been demonstrated in the elderly (45). Other patients might not be eligible to be enrolled in a structured therapeutic exercise program due for instance to insurance-related restrictions. In other cases, a patient might not be willing to undergo therapeutic exercise due to a previous unsuccessful attempt at PT. In these circumstances, it becomes important for the physiatrist to distinguish a true nonresponder from a patient who did not actually receive a valid attempt at PT.

The following sections on nonpharmacologic treatment should help convince the reader that the physiatrist is in an excellent position to direct this aspect of management as it incorporates many of the educational, exercise, orthotic, and functional aspects of physiatric training.

**Patient Education and the Arthritis Foundation**

Overview

Despite the link between OA and high disability rates in the United States, the cost-effective interventions such as patient education to reduce the burden of OA are currently underused and often go unnoticed. Thankfully, one national voluntary health agency champions the cause for this population. The Arthritis Foundation is the largest private, not-for-profit contributor to arthritis research in the world, funding more than $380 million in research grants since 1948. The Arthritis Foundation’s overall mission is to improve lives through leadership in the prevention, control, and cure of arthritis and related diseases. In this capacity, it has been particularly active in the area of patient education and in the delivery of community-based exercise programs (46). During its 60-year history, the foundation has grown into over 150 chapters and service points nationwide to provide community-based programs, to fundraise, and to act as the primary vehicle to inform and empower the millions of Americans with arthritis. The foundation emphasizes education for the public, patients, and their caregivers. By focusing on how individuals may

<table>
<thead>
<tr>
<th>General Guideline</th>
<th>Example of a Specific Intervention</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient education</td>
<td>• Self-management programs (e.g., Arthritis Foundation Self-Help course)</td>
</tr>
<tr>
<td>Health professional social support</td>
<td>• Arthritis Foundation publications</td>
</tr>
<tr>
<td>Weight loss (if overweight)</td>
<td>• Education of the patient’s family, friends, or other caregivers</td>
</tr>
<tr>
<td>Physical therapy</td>
<td>• Telephone contact</td>
</tr>
<tr>
<td></td>
<td>• Direct health professional contact</td>
</tr>
<tr>
<td></td>
<td>• Exercise</td>
</tr>
<tr>
<td></td>
<td>• Dietary changes</td>
</tr>
<tr>
<td></td>
<td>• ROM exercises</td>
</tr>
<tr>
<td></td>
<td>• Strengthening (especially quadriceps strengthening for knee OA)</td>
</tr>
<tr>
<td></td>
<td>• Pain-relieving modalities</td>
</tr>
<tr>
<td></td>
<td>• Assistive devices for ambulation</td>
</tr>
<tr>
<td></td>
<td>• Specific interventions for some knee OA patients</td>
</tr>
<tr>
<td>Orthotic intervention</td>
<td>• Patellar taping</td>
</tr>
<tr>
<td></td>
<td>• Lateral-wedged insoles if genu varum deformity</td>
</tr>
<tr>
<td></td>
<td>• Shock-absorbing insoles if no genu varum deformity</td>
</tr>
<tr>
<td></td>
<td>• Knee bracing</td>
</tr>
<tr>
<td>Occupational therapy</td>
<td>• Joint protection</td>
</tr>
<tr>
<td></td>
<td>• Energy conservation</td>
</tr>
<tr>
<td></td>
<td>• ADL training</td>
</tr>
</tbody>
</table>
take control over the debilitating aspects of the disease, this volunteer-driven organization supplemented by dedicated staff to provide professional expertise and support for volunteer efforts helps individuals with arthritis improve quality of life through its many community-based, educational programs and activities. Outreach is provided through telephone and e-mail information services as well as the foundation’s internet site, www.arthritis.org, which serves as an outstanding resource for online information and support. Professional education directives also benefit health care providers to better diagnose and treat patients with arthritis. Publications, for both professional and consumer groups, continue to provide accurate and sound health information to record numbers.

Rationale for Educational Interventions
OA often produces pain, weakness, and decreased physical function, thus negatively influencing quality of life. Psychological and socioeconomic variables influence the pain experience of OA; for example, depression may affect a patient’s belief about his or her ability to cope with OA pain. High levels of anxiety, manifested by persistent attempts to avoid knee pain, may lead to loss of muscle bulk, generalized deconditioning, and subsequently a loss of confidence. With avoidance of movement, a chronic cycle of inactivity, muscle atrophy, and weakness produces further pain and inactivity. Eventually, a vicious cycle develops and perpetuates, often with a concomitant decline in function.

Since psychosocial and physical factors explain some of the pain and functional variability of persons with OA, interventions must address either or both of these domains. In a study of 69 hospital-based outpatients, pain severity, obesity, and helplessness were the most important determinants of disability (47). Patient education is one of the interventions that has demonstrated a reduction in pain and enhanced function despite continued structural pathology.

A tenet of all health care management is that the client/patient is an integral member of the team. By acquiring knowledge about the disease and the options available, an individual can best manage the implications of an arthritis diagnosis. Patient education as a planned, organized set of learning experiences designed to facilitate voluntary adoption of behaviors or beliefs conducive to health empowers the patient with new information, skills, beliefs, and attitudes to best influence health status, quality of life, and possibly health care utilization. Self-management strategies and self-efficacy, the belief that one is capable of successfully executing the behavior required to produce a certain outcome, intermingle into a framework for the adoption of the new knowledge and behaviors and strengthen the individual’s confidence that he or she can perform these new actions.

Educational Interventions
To affect behavior changes, self-help educational programs use five major strategies: contracting, feedback, modeling, reinterpreting physiological symptoms, and persuasion. The Arthritis Self-Help Program developed at the Stanford Arthritis Center incorporates these educational principles to instruct course members about pain, fatigue, and stress management; development of an individualized exercise program; purposes and effective use of medications; identification of coping strategies; and forming a partnership with the health care team. Past participants have experienced such benefits as increased knowledge about their arthritis, increased frequency of exercise and relaxation, increased self-confidence, decreased depression, decreased pain, and decreased physician visits.

Studies have also looked at the effects of patient education combined with strengthening exercises. Persons with knee OA, who participated in an 8-week walking and education program, reported less pain and medications than controls (48). In a similar study, after the 8-week intervention, persons with knee OA had an increase of 15% in distance walked during the timed 6-minute walk test, with a concomitant increase in stride lengths at both normal and fast speeds (49). The Arthritis Foundation’s research-based Walk with Ease program motivates individuals to increase their exercise levels through development of a solo or group walking programs in a safe and effective manner. Participants report such benefits as increased physical activity, increased walking distance and speed, decreased pain, and decreased depression. This program is beneficial for the participant with OA who is apprehensive about performing exercise. The curriculum empowers the individual to take the first step toward a regular physical activity program and the adoption of a healthy behavior.

For those wishing to engage in moderate physical activity, they may join one of the foundation’s three Life Improvement Series Exercise programs. These evidence-based programs include the Arthritis Foundation Aquatic Program (AFAP) (to be discussed later in this chapter), the Arthritis Foundation Exercise Program (AFEP), and the new Arthritis Foundation Tai Chi Program. The latter two programs can be performed either in sitting or in standing and offer adults with arthritis options to increase physical activity while minimizing the stresses to their joints. The AFEP, formally known as people with arthritis can exercise (PACE), has been given at community centers around the country for over 20 years. This program consists of patient education infused with gentle exercise and activities that help to improve flexibility, strength, endurance, and function. Evidence shows that regular participation in these classes helps participants to better control the symptoms of the disease and the friendly and supportive atmosphere that can lead to improved mood and self-esteem. In a randomized control trial with 346 adults with arthritis, Callahan et al. found that if participants attend a majority of AFEP classes of an 8-week series of sessions twice a week, they may expect improvements in symptoms, self-efficacy for arthritis management, and upper and lower extremity function (50). They also noted that sustained improvement may require continued participation in the program.

In addition to flexibility and strength deficits with this disease, many adults with arthritis report difficulties with balance and walking. Tai Chi is an ancient, gentle form of exercise that emphasizes breathing and mental focus during the
performance of slow, continuous movements and postures. These techniques are thought to integrate the mind and the body resulting in a harmonious inner and outer self. Known benefits include increasing flexibility, muscle strength, fitness, balance, and cardiovascular (CV) health. The Sun Style of Tai Chi is especially suited for persons living with arthritis. This style utilizes forward or backward stepping motions, which is believed to improve mobility. When one foot moves forward or backward, the other foot follows. The higher stance places less stress on lower extremity joints. The coordinated breathing tied to motion promotes symptom relief and relaxation. Thus, the target audience for the Arthritis Foundation Tai Chi program is individuals with arthritis or related conditions characterized by diffuse muscle pain, decreased muscle strength, and fatigue. Preliminary outcome data suggest that participation results in improved lower extremity strength, balance, and reach.

**Weight Loss and OA**

**Overview**

Obesity is a growing epidemic in the United States and worldwide. It is estimated that currently 33.2% of women and 27.6% of men in the United States are overweight or obese (51). These statistics in conjunction with the apparent correlation between obesity and OA have prompted a number of studies. Multiple population-based studies have shown that increased body weight is associated with OA, especially of the knee, and that being overweight actually precedes the development of knee OA, rather than vice versa (52,53). In addition, it has been demonstrated that risk ratios for radiographic knee OA increase with weight, even in normal BMI range, and that weight loss protects against knee OA development (54).

A prospective study of 1,180 male medical students followed them over a median duration of 36 years. For each 8 kg (18 lb) heavier that men were during ages 20 to 29 years, the incidence of subsequent knee OA was substantially increased (relative risk =1.7) but without an increased risk of hip OA. In addition, BMI at ages 20 to 29 years was more predictive of future knee OA than BMI at ages 30 to 39 or 40 to 49 years, indicating that duration of obesity may play a cumulative role (55).

Human obesity is associated with an increased risk of knee OA, specifically at the patellofemoral compartment and the medial tibiofemoral compartment, and most notably with simultaneous OA at both of these sites and in the knees bilaterally (52,56). Conversely, knee injury predisposes to knee OA that is usually unilateral, rather than bilateral (57). Body weight has a weaker association with hip OA than with knee OA (58). Interestingly, overweight patients may also be at increased risk for OA of the hand, thus affecting joints other than those intuitively considered to be weight bearing (58).

Women are two to three times more likely to develop knee OA in comparison with men and are twice as likely to suffer from knee OA bilaterally (59). A cross-sectional survey of 6,987 males and 7,689 females found that, without exception, for any given category of body weight, the risk of OA was higher in women than in men within the same BMI category (60). It is believed that gender difference in reference to gait may play a role in this finding. In the McKeen study done on 42 healthy subjects (18 men and 24 women) and 39 individuals (24 men and 15 women) afflicted by OA (59), it was found that OA women generated less torque at the knee and ankle and had decreased ROM at the knee. Meanwhile, OA men had results that were comparable to healthy men.

**Effect of Weight Loss on OA**

In keeping with the theory that increased weight causes a proportionate increased risk in development of OA, studies were done to determine if losing the weight could lessen the risk. The Framingham Study showed that women with weight loss of approximately 11 lb during the 10-year period preceding the study decreased their risk of symptomatic knee OA by over 50%. Conversely, previous weight gain was associated with a slightly increased risk of subsequent knee OA (61). Thus, it appears that weight loss can be effective as a primary means of prevention against the onset of symptomatic knee OA.

Weight reduction seems to not only help prevent onset of OA, but it may also alleviate symptoms when the diagnosis of OA has already been made, especially for OA of weight-bearing joints (62,63). In one study where the participants lost an average of 100 lb over a 12- to 18-month period through gastric bypass surgery, there was a considerable improvement in pain, function, and knee ROM during stance and swing phase of walking (54). Research appears to indicate that even small amounts of weight loss may be beneficial for OA, especially for knee OA. For example, a study done in 2005 on elderly women in Denmark found that reducing BMI greater than two units over a 10-year period was associated with a 50% decreased risk in developing OA (64). Meanwhile, phentermine use as an anorectic agent in obese women with knee OA has been shown to be more effective than placebo at achieving 12 lb of weight loss and improving OA symptoms at the hips and especially at the knees (65).

While an ideal weight loss program would typically start with an appropriate exercise program and dietary modification as a first-line approach, the various health risks of obesity may justify the more aggressive treatment, for example, also using anorectic agents and surgical approaches, under close medical guidance (66). The potential short-term and long-term benefits of weight loss might outweigh the potential side effects (67). Thus, many clinicians and researchers now feel that pharmacotherapy may also be an appropriate and effective supplementary treatment, in addition to diet and exercise (68,69).

It has been demonstrated that the most effective nonpharmacologic weight loss interventions combine fat and caloric restriction, increased physical activity, behavioral reinforcement, and an extended weight maintenance program, with support from the physician and weight-loss support groups (68). These measures require a considerable commitment by the patient, and substantial weight reduction is often difficult to achieve and maintain (70).
The overall goals of weight loss include benefits not only to the patient’s OA but also to the patient’s general health and functional status (71). For example, in patients with knee OA, obesity is an independent risk factor for disability (72). An extensive review of clinical studies has shown that moderate weight loss (5% to 10% of baseline body weight) has clear benefits in terms of overall health (73).

There are still questions to be answered, but it appears as though weight loss can have a resounding beneficial effect only on the likelihood of developing OA, the symptoms of OA, and on the likelihood of progression of OA. This can be accomplished by a carefully monitored, comprehensive weight loss program including dietary modifications, appropriate exercise, and possibly pharmaceutical and surgical interventions.

**Physical Therapy and OA**

**Background**

While medications, injections, and orthotic interventions are valuable treatments discussed elsewhere within this chapter, it is important for rehabilitation specialists to also have specialized knowledge of the role that PT, particularly with respect to therapeutic exercise, can play in patients with OA. Many medical treatments are more successful at decreasing pain than they are at decreasing disability. But exercise often can target factors that lead to disability (74), and long-term benefits may include improving fitness and encouraging participation in physical activity programs that may help to reduce the risk of various comorbid conditions associated with obesity and a sedentary lifestyle (75). Such exercise programs may include stretching, strengthening, aerobic conditioning, aquatic programs, proprioceptive training, and use of joint-protection techniques. Until fairly recently, many physicians recommended against most exercises in OA patients, but the benefits of a well-planned exercise program have become increasingly clear within the medical literature (76–78). In 2005, the Ottawa Panel evaluated the strength of the scientific evidence on the efficacy of therapeutic exercise (79). They found 26 randomized clinical trials and controlled clinical trials for analysis. This resulted in 16 positive recommendations for therapeutic exercise and physical activity for pain management and improvement of physical function in this population. Exercise is now considered to be an integral component of primary prevention (preventing the occurrence of OA), secondary prevention (before OA has caused significant clinical symptoms), and tertiary prevention (after the diagnosis has been established) (80).

While exercise is believed to be essential for improving patient outcomes, unfortunately, definitive recommendations cannot be made regarding optimal exercise programs for any joint region with OA since with only a select number of high-level randomized clinical trials available. The inability to make judgments about the best possible exercise program is further compounded by the wide range of exercise interventions that appear in the literature. Clinical and research designs have incorporated programs that focus on simple local joint exercises to complex, multijoint exercise plans that also include other modalities (81). Furthermore, contemporary research studies utilize different modes of exercise delivery, such that the exercises are taught to one individual or in a group setting programs, in a supervised clinical setting, community center, or on a home basis. The advantage of this array allows for the health care practitioner to select exercise regimens that may maximize a patient’s adherence.

**Flexibility**

ROM deficits are well-known sequelae of OA (76). Hip OA typically is accompanied by deficits in internal rotation, followed later by restricted abduction and flexion. Knee OA classically involves extension lag, but flexion may also be limited. The pathophysiology of ROM deficits is probably multifactorial, including articular changes within the joint as well as shortening of myotendinous structures in areas of pain and/or weakness. Joint stiffness is a common complaint of OA patients. Decreased ROM is often found not only at the OA joint but also at other joints within the same lower limb and even in the contralateral lower limb (76). When muscles are shorter than their ideal length, they are at a biomechanical disadvantage when they are required to generate force. Thus, a stretching program to address inflexibilities should probably be incorporated early in an exercise program for OA patients. ROM is also obviously important for functional activities. For example, various degrees of knee flexion are needed for mobility, for example, requiring at least 70 degrees for walking on level surfaces, 83 degrees for climbing stairs, and 93 degrees to get up from a chair. Neutral knee extension is required for all of these activities (76). Upper limb ROM deficits in OA patients are often seen in the fingers, for example, at Heberden’s node at the DIP joints, with functional implications for ADLs and fine motor skills.

A flexibility program often begins with patients gently moving their joints through the readily available ROM, to prevent further loss of range. Next, stretching should be added to reverse some of the previously accrued ROM deficits. OA patients should be taught how to perform the stretching program properly, generally with slow, gentle, and sustained stretching. Sustained stretching generally involves holding the stretch for at least 20 to 40 seconds, and perhaps longer, before relaxing and then repeating the stretch. Sudden, jerky or ballistic stretching should be avoided since it may cause exacerbation of the OA.

The rehabilitation specialist should provide the OA patient with guidance as to which muscle groups should be targeted by the stretching program. For patients with OA of the hip or knee, stretching of the quadriceps, hip flexors, and hamstrings muscles is crucial (76). Of course, any specific ROM deficits that have been identified as problematic within a specific patient should be addressed when formulating the stretching program. While some patients may be able to begin a flexibility program with little hands-on instruction, many will benefit from learning proper techniques from a qualified therapist and then transitioning to a home exercise program.
Strengthening

Muscle weakness has long been considered to be the best correlate of functional limitations in individuals with OA, especially knee OA (74). Some studies of patients with knee OA have shown that quadriceps weakness, in particular, is the single most important predictor of lower limb functional limitations. Quadriceps weakness was a stronger predictor of disability than was pain, radiographic findings, or other factors (82).

There are many potential reasons why the quadriceps muscles are important in knee OA. It was previously theorized that quadriceps weakness in patients with knee OA was due to deconditioning from disuse, perhaps secondary to the pain of the OA. But more recent studies have shown that quadriceps weakness actually seems to precede the development of knee OA, that is, apparently serving as a causative factor for knee OA rather than resultant sequelae. Explanations for this include the role of the quadriceps muscles in knee joint protection. Presumably, repeated lower limb impulse loading may contribute to the onset or progression of knee OA. Strong knee extensors can decrease the impulse loading of the lower limb by slowing the deceleration phase before heel strike (74).

Given that the quadriceps are the muscles most often associated with knee OA, it is not surprising that they are also the most well-studied muscle group within therapeutic strengthening programs for patients with knee OA (45). A recent review by Baker and McAlindon showed that despite many differences in research designs and strengthening programs, overall the vast majority of studies assessing quadriceps strengthening have found positive outcomes in reducing pain and/or disability (74).

Strengthening exercises are often classified as closed kinetic chain (where the distal aspect of the limb is against a fixed source of resistance, e.g., the foot on the floor during squatting) or open kinetic chain (where the distal part of the limb is free in space, e.g., knee extension while seated or recumbent). Open kinetic chain exercises may allow the patient to more specifically strengthen just one movement at one joint (e.g., isolated knee extension), but these exercises increase shear forces at joints, which may exacerbate OA. Conversely, closed chain exercises seem preferable because they cause less shear forces and may more closely mimic the many synergistic movements and firing patterns required for everyday functional tasks (76).

Strengthening exercises are often classified as isometric, isotonic, and isokinetic. Isometric exercises involve contraction of muscles while they are at a fixed length, such that little or no joint motion is taking place. Isometric exercises are a reasonable initial approach for strengthening in patients who cannot tolerate repetitive joint motion, for example, with painful, inflamed joints (76). Isometric contractions held for 6 seconds, performed at 70% of the maximal voluntary contraction (MVC) and repeated 5 to 10 times daily, have been shown to increase strength, but contractions greater than 50% of MVC can produce postexercise soreness (83). While isometric contractions can be learned quickly and can result in rapid strength gains, functional benefits from isometric contraction exercises might be limited to a small range around the joint angle of training (74,76). Therefore, these are suggested to be used prior to or in conjunction with other forms of exercises (83). In a recent study that compared the functional benefits of either a dynamic exercise program or an isometric resistive program to a control group, subjects who engaged in either exercise program increased physical function and decreased knee pain more than those participants who did not exercise (84).

Isotonic exercises are often used to maximize strength (76). By definition, isotonic refers to joint movement through its ROM, against a constant weight or resistance. Isotonic work consists of concentric and eccentric contractions, depending on whether the muscle length is shortening or elongating during the contractions, respectively. Eccentric contraction is more stressful than concentric but also results in greater strength gains. For the patient with OA, strengthening exercises should be performed within a pain-free range and progress toward the outer ROM (83).

Isokinetic exercises involve constant speed of motion throughout the joint range during muscle contraction, while the amount of resistance may vary throughout the range. Isokinetic exercises are infrequently used, due to equipment requirements and uncertain correlation to functional activities. Instead, functional training has replaced isokinetics as a component of a comprehensive exercise regime.

In summary, isometric and closed chain exercises are probably the most beneficial therapeutic exercises in OA patients, but individuals who find these too stressful may perform isometric exercises. The goal should be to progress to isotonic and functional exercises that include a combination of open and closed chain activities (74,81).

Balance and Proprioceptive Training

If lower extremity proprioception is suboptimal, the force of impact transmitted up to the hip and knee will be increased during weight-bearing activities (76). Repetitively, such forces may promote progression of OA and the associated symptoms. Similar to strengthening exercise prescription, the best method of optimizing balance in OA patients is not fully clear. Some older studies have shown that postural stability in OA patients is improved by strength training or especially by aerobic walking programs, while other studies have found, ironically, that exercises specifically designed to improve balance had no such effect (85). More recently, Fitzgerald investigated the use of agility training and balance/perturbation training techniques for a 73-year-old woman with complaints of instability with her bilateral knee OA. After 12 sessions, the subject reported no further incidents of instability and returned to recreational walking, tennis, and golf (86).

Joint Protection

Joint protection is one goal of exercise in a patient with OA. Flexibility, strength, and proprioception are optimized in hopes of reducing joint stresses, decreasing shock impacts to the joint, and maximizing joint movement and alignment (76).
Recently, the uses of patellofemoral taping techniques have been applied to relieve knee pain and improve reported physical function. While the taping group reported significantly less pain and improved function than those with no tape, the placebo tape group also reported benefit (86). Energy conservation techniques may help in minimizing unnecessary activities and joint stressors, so that the OA patient will be able to use his or her sometimes limited abilities to gain the most functional independence possible.

Modalities
Various passive modalities have been used in OA patients, with varying degrees of support within the medical literature. Application of therapeutic heat, through use of hydrotherapy, hydrocollator packs, or paraffin, seems to be effective at improving myotendinous flexibility, allowing enhanced subsequent stretching or promoting relaxation leading to pain relief. Similarly, therapeutic cold is well-tolerated and highly effective analgesic modality. A recent review of the evidence indicates that two randomized studies found that ice massage or ice packs applied to the knee improved muscle strength and ROM or decreased swelling, respectively (87). Some higher-level studies support the benefits of repeated use (4 weeks or more) of high-intensity burst modes and acupuncture-like transcutaneous electrical nerve stimulation (TENS) in relieving pain and improving function in OA patients (76,86). In a meta-analysis that specifically addressed the use of TENS for knee OA, the authors concluded that the data supported its use and reported a moderate effect size of 0.45 (81). Subsequently, TENS has been recommended as the only electrophysical agent by the American Physical Therapy Association in its recent clinical practice guidelines for patients with knee pain (88). Conversely, the medical literature contains little or no support for the routine use of electrical stimulation, iontophoresis, or ultrasound in OA patients (76,89). Possible reasons for the ineffectiveness include the limited treatment length, inconsistent dosage, and uncontrolled treatment area size and mechanical frequencies (81). In general, modalities are probably most appropriate when used as a means of facilitating more active components of the exercise program, rather than as treatments given in isolation. They may also be used for short-term relief of acute exacerbations.

Manual therapy is another intervention that physical therapists employ in conjunction with exercise to treat patients with arthritis. It is a general term used to describe passive movements applied by the physical therapist and may include passive ROM, passive accessory joint motions, soft-tissue mobilization, or massage techniques (88). These techniques aim to increase joint motion or reduce joint stiffness. Deyo et al. found that subjects who received manual therapy to the hip, knee, or ankle in addition to exercise reported significant improvements in pain, 6-minute walk test, and self-reported function when compared with a control group (90).

Compliance
Obviously, no exercise program can be expected to be beneficial if the patient is not compliant with performing the therapeutic activities. Health care workers treating OA patients should encourage patients to be actively engaged in performing their therapeutic exercises. Barriers to compliance should be actively elicited so that obstacles can be addressed and surmounted. Some clinicians advocate patient diaries to track compliance with recommended exercises.

A recent study on compliance in OA patients found that demographic, fitness, psychosocial, and disability-related measures did not predict compliance with any consistency. The researchers found that activity during the first few months of an exercise program was the strongest predictor of longer-term exercise compliance. Results also suggested that exercise should be prescribed at least three times each week, for approximately 35-minute sessions (91). To retain these treatment benefits longer than 6 months, the findings of a systematic review indicate that moderate evidence exists for when exercise programs contain booster sessions (92). Community-based exercise classes serve as excellent adjuncts for the long-term maintenance of the benefits obtained from the therapeutic program (81).

Lower Limb Orthotic Intervention
Background
Pain relief and joint protection via structural support and realignment are principle benefits that can be realized by the OA patient from the use of orthotic devices. The institution of a lower limb orthosis is often a team effort, involving the physician, physical therapist, and the orthotist. While the physician can direct care and prescribe bracing intervention, physical therapists can also recommend bracing during the course of treatment, and the orthotist will subsequently evaluate the patient for appropriate device function, fit, and usage. After the orthosis has been issued, the physical therapist can continue to work with the patient to help ensure proper usage and short-term compliance while the physician can monitor for long-term compliance.

Lower limb OA braces are prescribed to inhibit multiplane joint mobility so as to relieve pain associated with weight bearing. As will be discussed, pain reduction is achieved by supporting the affected joint, reducing the muscular force needed to stabilize the joint, and redirecting axial loads, which lead to intra-articular bone-on-bone forces. The clinician working with the OA patient should provide proper education on brace function and application. In particular, the prescribing physician should set proper outcomes expectations by explaining to the patient that the brace will not correct current “damage to the joint,” rather it will provide support, potential pain relief, and hopefully prevent further joint injury.

Clinically, the chief complaint of a knee OA patient is often pain within the knee and/or the surrounding tissues. Unicompartamental gonarthrosis of the medial compartment is most often seen and can be attributed to the forces present during the varus moment that is present during mid-stance of normal human gait. In order to be most effective, a knee OA brace must act to decrease this varus-related compressive force.

The early history of research as it relates to knee OA and the application of an orthotic appliance to alter forces and produce
pain relief includes two main studies. Smith et al. described the emphasis of their studies as follows: “It was often possible during weight bearing for the physician to control instability of the knee by applying forces with his own hands, and in so doing relieve the associated pain. This relief can occur even though the disease is still present, a fact emphasizing the role of mechanical stress in pain production. If pain can be reduced by the proper application of forces manually, it may also be possible to relieve it with a properly designed brace” (93). In 1975, the Canadian Arthritis Research Symposium—University of British Columbia knee orthosis (CARS UBC), as described by Cousins and Foort, was designed specifically for the treatment of knee OA (94). The brace was comprised of plastic thigh and calf shells, utilized universal hinges and a telescoping tube assembly, and required a waistband for suspension.

Knee bracing for medial compartment gonarthrosis currently involves application of a three-point force system across the coronal plane of the knee joint, that is, valgus bracing that “unloads” the medial compartment. Current bracing terminology refers to this type of orthosis as an “unloading brace” (Fig. 31-1). Manufacturing trends have focused on single and double upright brace designs that are fabricated from high strength, very lightweight materials.

A more current study showed that the external varus moment at the knee was reduced during initial contact and loading response phases of the gait cycle (95). The interval between 10% and 15%, or loading response, is the point of peak varus loading; therefore, a reduction in the external varus moment about the knee during this interval is felt to be important for diminishing load and subsequent pain reduction. There was, in fact, improvement in pain and function with use of the brace as measured by a visual analog scale (VAS). A second study showed that the mean femorotibial angle decreased from 185.1 degrees to 183.7 degrees. In addition, isokinetic quadriceps muscle strength increased from an average of 36.8 to 42.8 Nm (96). Another study was performed during which patients were asked to walk with and without a brace, and condylar separation angles were measured under fluoroscopic guidance. Twelve out of fifteen patients demonstrated an average condylar separation angle of 2.2 degrees, while the three subjects without separation were obese (97). Lindenfeld et al. evaluated the biomechanics of an unloader brace on patients with medial compartment involvement and found that the adduction moment at the knee was altered when the brace was worn (98). This change may reduce the load transmitted through the medial compartment, resulting in less pain. Lastly, the clinical knee score and bone mineral density of the proximal tibia were assessed in an attempt to evaluate the efficacy of valgus knee bracing. The knee score improved after 3 months, and increases in bone mineral density were seen more in the lateral tibial condyle than in the medial, suggesting that the brace transmits forces across the knee joint from the medial to the lateral side (99).

A clinician should consider using a simple knee sleeve for those patients who present with mild OA-related knee pain but without any significant angular deformity. Although there is still no scientific evidence from clinical studies that highlight positive results using only a sleeve, there are many anecdotal patient accounts. Since the nonjointed sleeve provides no significant structural support to the knee, patient-reported feelings of improved stability and reduced pain are likely due to an improvement in proprioception that theoretically occurs with use of the sleeve. In 1999, a study was performed in order to measure the efficacy of a sleeve along with pharmacological treatment versus an unloader brace and medication as compared to the control group of patients who were given medication only. A significant improvement in function, in addition to disease-specific quality of life ($p = 0.001$), was found in both the neoprene-sleeve and unloader brace groups as compared to the control group. The unloader brace was on average more effective than the neoprene sleeve (100).

Wedged shoe soles/insoles are additional orthotic devices that can possibly be used to counter OA-related knee joint compressive forces. Patients have reported pain relief in the medial compartment after applying a lateral wedge under the calcaneus. Sasaki and Yasuda examined the effects of using lateral heel wedges with load transducers and roentgenograms to observe the biomechanical effects (101). The variables that were under study included the ground reaction force (GRF), the tibiofemoral angle, and the tibiocalcaneal angle. Although the GRF and the tibiofemoral angle did not exhibit a change, the lower limb mechanical axis, however, approached an upright position, and the calcaneal angle corrected in the valgus direction with reference to the subtalar joint. The observed changes were believed to reduce the axial load bearing on the medial joint surface. The use of the wedged insole at a certain angle serves as a “stimulant” to the patient, because it causes an inclination of the floor to which the patient must compensate.
Patients must adapt to the wedge via changes in muscle activity and posture, ultimately modifying their standing and walking mechanics. In addition to their effects on knee pain and forces acting at the knee, other investigators have examined the effects of lateral wedged insoles on knee kinetics. In this study, three-dimensional analysis was performed on 17 subjects with and without them wearing a 5-degree lateral wedged insole. They found no significant differences in temporal and spatial parameters; joint angles at the hip, knee, and ankle; or kinetics at the hip and ankle. The external varus moment and estimated medial compartment load at the knee were reduced significantly (102). One can interpret the results to indicate that lateral wedged insoles are a potential effective component of a conservative treatment regimen for patients with mild medial compartment OA.

Viscoelastic shoe inserts presumably provide shock absorption at the knee and as such are another intervention at the foot and ankle for the treatment of painful knee OA. Walking has been linked to the generation of impulsive loading on the human locomotor system. Impulsive loading from normal gait has been associated with “damage” to some elements of the locomotor system. Specifically, boney vibratory forces during normal gait have been detected at 25 to 100 cycles/second at initial contact and are believed to lead to articular cartilage deterioration (103). The feet, menisci of the knees, intervertebral discs, and vertebral bodies are the so-called “natural shock absorbers,” as one of their functions is to absorb and dissipate the energy that acts upon the body during heel strike. Depending on gait velocity and the foot’s viscoelastic properties, the dynamic forces to which the foot is subjected during gait may be considerably higher than the body weight. Under normal physiologic conditions, the intermittent and continuous shock and load experienced during gait tends to cause a slowly progressive weakening of the natural shock absorbers and may later lead to articular cartilage degeneration and OA. Insufficient damping of incoming shock waves can overload the proximal joint and contribute to the process of joint degeneration (103,104).

Voloshin and Wosk clearly showed that viscoelastic inserts reduce the amplitude of incoming shock waves by 42%. They, therefore, can dampen overloading and in a sense possibly prolong the life of articular cartilage. Conservative treatment, using viscoelastic shoe inserts, for patients with clinical symptoms of knee OA too yielded an improvement in symptoms by 78% (103).

After medial knee compartment involvement, patellofemoral compartment involvement is the second most common site of involvement in knee OA. The principal biomechanical factor noted to cause patellofemoral joint pain is felt to be patellar maltracking. Specifically, it is believed that there is abnormal lateral patellar tracking, thus increasing lateral patellar facet contact forces (105). The purpose of orthotic intervention is to help achieve and maintain optimal patellar tracking in order to prevent abnormal compression forces and degenerative changes. Knee braces used for patellofemoral compartment OA are specifically designed to reduce patellar compression as well as to prevent excessive lateral shifting. Criteria evaluated at time of prescription are patient vocation/avocation, likely compliance, and cost. Patellar-stabilizing braces, infra-patellar straps, and patellar taping are some of the current orthotic treatment options.

The neoprene-sleeve patellar stabilizing brace consists of a patellar cutout and force inducing buttress pads around the inferior and lateral aspects of the patella. The brace has two circumferentially wrapped rubber straps that apply dynamic tension to a crescent-shaped lateral patellar pad. Palumbo found this sleeve useful in the treatment of patients with patellofemoral arthritis (106). Cushnaghan et al. demonstrated that medial taping of the patella resulted in a 25% reduction in knee pain compared to lateral or neutral taping (107).

Although hip OA is also extremely common, there is a much smaller body of literature on orthotic intervention for this body region. This is not surprising given the practical difficulties encountered with bracing at the hip joint. Ankle OA, by contrast, is much less common than hip and knee OA. Literature pertaining to orthotic intervention for ankle OA is quite limited and largely anecdotal.

**Conclusion**

The inclusion of an orthosis into the overall treatment plan should be considered for the OA patient. Patients can benefit by using a device daily, particularly during certain activities that stress the lower limb joints. The pain that is associated with the degenerative condition can be reduced or eliminated, thereby increasing the activity level. Furthermore, more invasive treatment can be delayed or potentially avoided all together as a result of the benefit of joint protection from damaging forces. The orthosis should be used as part of a comprehensive management plan, which could include simultaneous pharmacologic intervention, joint injections, and exercise prescription. Other factors that can help to maximize the chance for success include proper patient assessment, explaining treatment goals to the patient, achieving appropriate device fit, and utilizing the services of a licensed orthotist to make modifications when necessary.

**Hand and Occupational Therapy for OA patients**

In addition to the many potential benefits from PT interventions, OT is generally believed but not proven to be of benefit in select patients (1).

**OT for Upper Limb OA**

OA can affect a variety of upper limb structures including the shoulder girdle region, the elbow, the wrist, and the hand. OA in the shoulder region may involve the glenohumeral joint, the acromioclavicular joint, and rarely the sternoclavicular joint and nonarticular locations such as the subacromial space due to impingement of the rotator cuff tendons and/or the subacromial bursa by acromion bone spurs. Elbow region OA may be secondary to prior trauma or as a result of inflammatory joint disease or may be primary due to repetitive overuse. OA of the wrist region can occur due to prior trauma, inflammatory joint disease, prior infection (especially prior gonococcal or Lyme infection), or perhaps due to overuse. OA of the hand region...
is very common and especially affects the first CMC joint, the proximal interphalangeal (PIP) joints, and the DIP joints.

Hand OA can be associated with significant pain and disability, especially in women (108). Upper limb weakness due to deconditioning from pain-related disuse and stiffness is common in OA patients. The Arthritis Hand Function Test can provide reliable, standardized measurements of strength and dexterity in adults with OA (109).

An OT program can be targeted toward the associated pain, disability, and weakness (110). Modalities, including therapeutic heat or cold, may provide pain relief, especially during an exacerbation. Adaptive equipment may include devices that empower the patient to independently perform activities such as opening jars, buttoning clothing, or holding utensils. Occupational therapists often prescribe assistive devices to assist with ADLs and train patients in their use. Although it is believed that patients with OA are frequent users of such assistive devices, there is a scarcity of scientific evidence for the prescription, provision, and use of these in OA patients (111).

Splinting may allow inflamed joints to rest, provide support to weak muscles, control ROM, and theoretically prevent deformities (110). Prefabricated or custom-made finger splints can be used to immobilize joints that are inflamed or painful, while allowing movement at adjacent joints. OA often involves the DIP and PIP joints in the form of Heberden’s and Bouchard’s nodes, respectively, as well as the first CMC joint. Various finger splints are available. Some of these may offload the DIP or the PIP joint, whereas a thumb spica splint can immobilize the CMC joint and the MCP joint. A CMC splint (C splint) supports the first CMC joint while allowing movement at the wrist and at the interphalangeal (IP) joint of the thumb. A systematic review of the literature found that although there was fair evidence for the effectiveness of CMC splinting to relieve pain and improve function, further study is needed (112). Patient preference regarding type of splint varied, and there was no clear evidence of the superiority of one type of splint over another. There is also some disagreement regarding whether splinting for first CMC joint OA should also extend proximal to the wrist and/or perhaps even distal to the IP joint of the thumb (110). In general, the prescribed splint should probably be the least cumbersome brace that will immobilize the necessary joints. Any extension of the splinting beyond the primary joint(s) being targeted should be pursued only if there are symptomatic or functional gains that will offset the increased splint size. Ideally, hand and finger splints should still allow reasonably good hand dexterity for functional activities, although some patients may initially find them awkward.

The strengthening component of the OT program can guide the patient in finding the right balance between avoiding disuse and avoiding excessive repetitive stress.

OT can also help address underlying ergonomic deficiencies that may be exacerbating the upper limb OA. Ergonomic modifications can be made to workstations and home environments so as to minimize any upper limb cumulative trauma disorders that may be exacerbating upper limb OA (113). Furthermore, in patients with lower limb OA, OT can address whether assistive devices such as canes and walkers are inadvertently causing excessive loading of the upper limbs (110). Modifications of the handles and/or heights of the walking aids can minimize detrimental impact on the upper limbs and thereby allow the patient to continue using the device.

While the literature on the effect of a course of OT for hand OA patients in general is extremely sparse, one study documented a significant improvement in mean disability score in 77 patients who underwent an OT program that included hand exercises and splints (108). A 7-year prospective study found that 70% of 33 patients with first CMC OA who were awaiting surgery for this condition no longer wanted surgery after participation in a program of modalities, splints, and ADL training (114). Referral of hand OA patients to an occupational therapist should therefore be considered, especially for first CMC OA.

**OT for Lower Limb OA**

Some lower limb OA patients may benefit from OT, if there are deficits in functional activities such as lower extremity dressing, bathing, and home safety. OT can provide appropriate assistive devices to help maximize the patient’s independence and function for a wide variety of tasks. Items as seemingly simple as reachers, sock donners, long-handled shoe horns, or elastic shoe laces may make the difference in whether a patient with limited lower limb ROM can get dressed without additional help (110).

**Other Benefits of OT for OA**

Since endurance is believed to be often impaired in OA individuals, energy conservation techniques can play an important role in maximizing functional independence. Energy conservation techniques involve shaping the patient’s habits and environment so that the most function can be accomplished with the least amount of effort. The goal is not to decrease the activity level of the individual but rather to harness that activity level as efficiently as possible, to maximize independence and quality of life. Finally, an OT can help to teach the concept of joint protection techniques to the patient. Joint protection techniques include the following principles:

1. Respect pain
2. Avoid activities that hurt affected joints
3. Use appropriate assistive devices
4. Use largest and strongest joints and muscles
5. Avoid staying in one position for too long
6. Balance activity and rest
7. Avoid prolonged periods of immobility
8. Reduce excess body weight
9. Simplify, plan, and organize

**Aerobic Conditioning and Aquatic Therapy (Aquaatherapy)**

The final nonpharmacologic treatment guideline is aerobic conditioning exercises, especially aquatic therapy (aquaetherapy) if available. In practice, although this is listed as a separate guideline, it is often incorporated as part of a formal PT program and then conducted within a long-term home exercise program. Aerobic
exercise for OA patients commonly includes a daily walking program since physical activity levels are often reduced. Using the guideline of 30 minutes of accumulated of moderate activity on most days is an excellent goal for those patients whose lifestyles are sedentary (115). In addition to the aerobic effects, walking adds some degree of joint loading, which helps the cartilage to obtain necessary synovial nutrients via imbibition. In a review of controlled trials evaluating the role of aerobic training in patients with knee OA, Baker and McAlindon concluded that aerobic exercise effectively minimizes further progression of disability (and possibly pain) in individuals with knee OA (74). They also concluded that supervised aerobic training seemed to yield better results than unsupervised training, although even home-based programs can prevent further decline over the long term. Cycling, both indoor and outdoor, is an excellent form of aerobic exercise for most people with arthritis. This low impact activity uses the large leg muscles that are often weakened due to OA (116). Mangione evaluated the effects of high-intensity and low-intensity aerobic cycling exercises in older adults with knee OA. Both forms of aerobic activity resulted in improved functional status, pain, gait, and aerobic capacity in this population.

Water is an amazing medium for patients with arthritis to wish to exercise. The increased sensory input, relaxation from warm water, and decreased joint compression often allow individuals to move with less significant pain than when on land. The hydrodynamics can support, assist, or resist movements so that the exercise program can easily be progressed. In 1983, the Arthritis Foundation and the Young Men Christian Association (YMCA) created a warm-water, community-based exercise program: The AFAP. Since its origin, the program has undergone three revisions. From 1989 to 1990, the program added an advanced “plus” level class (AFAP/AFAP PLUS) to expand the aerobic conditioning component to the basic AFAP program. A 1995 revision task force updated safety guidelines for endurance exercise, added joint-specific precautions, and supplemented the existing program with new exercises and special activities. A subsequent modification in April 2002 added two new versions of the aquatic program: deep-water exercise and exercise for children with arthritis. The latest revision in 2008 incorporates the advances in aquatic exercise and equipment to further enhance this widely popular program.

The basic AFAP program consists of 68 ROM and muscle-strengthening exercises for each major region of the body and is performed for up to 30 to 40 minutes. The instructor may add an optional, short 5- to 10-minutes endurance segment. By contrast, the AFAP PLUS program adds 5 to 20 minutes of endurance-building activities and additional muscle-strengthening activities to the basic program. The Deep Water AFAP classes utilize floatation devices to suspend participants above the pool floor while exercising. Due to the physical properties of the aquatic environment, most people with OA are capable of performing each of these program exercises, and the ability to swim is not necessary. Participants are encouraged to contact their physicians to discuss any exercise-related precautions or contraindications prior to beginning.

Since AFAP’s inception, over a thousand aquatic facilities offered the program. Even as a recreational program, participants reported physical benefits such as decreased pain and stiffness. The friendly and supportive atmosphere promotes peer interaction, helping to decrease feelings of depression and isolation. Successful pain-free participation results in new self-efficacious behaviors, and the progression of exercise tolerance and activity level may improve functional independence and improved self-esteem.

**PHARMACOTHERAPY OVERVIEW**

In addition to the nonpharmacologic guidelines discussed above, pharmacologic guidelines for OA of the knee and hip have also been published by the ACR (1). An algorithm based upon these guidelines is shown in Figure 31-2. These

---

**FIGURE 31-2.** Basic medication algorithm for knee or hip OA.
studies have compared acetaminophen to NSAIDs, including remains a logical choice for initial treatment of hip OA. Several impossible, to determine on physical exam alone due to the to be present. Hip inflammation is much more difficult, if not knee, especially in early OA when inflammation is less likely choice is acetaminophen. Acetaminophen remains a logical choice for relief of mild to moderate OA pain of the hip and if inflammation is not evident, then the initial medication of comitantly, as discussed above. The first question pertains specifically to the knee and que- tion decisions based upon the algorithm. For the sake of simplicy, the notion that the clinician can and at times should prescribe more than one of the above medications simultane- ously is not shown in the above algorithm. This treatment strategy of medication coadministration can take advantage of the fact that analgesics from different medication classes sometimes act synergistically and therefore limit the potential for dose-related toxicity. By contrast, as the clinical scenario changes, portions of the algorithm that had previously failed or were skipped over can be retried. For example, a patient whose knee inflammation has subsided after a corticosteroid injection but was not also receiving acetaminophen should be considered as a candidate for acetaminophen particularly if the knee inflammation was situational and related to unac- customed exertion that is no longer present. Finally, patients ideally should be receiving nonpharmacologic treatment con- comitantly, as discussed above.

Certain key questions are considered in the algorithm. The first question pertains specifically to the knee and queries whether or not there is evidence of knee inflammation. If inflammation is not evident, then the initial medication of choice is acetaminophen. Acetaminophen remains a logical choice for relief of mild to moderate OA pain of the hip and knee, especially in early OA when inflammation is less likely to be present. Hip inflammation is much more difficult, if not impossible, to determine on physical exam alone due to the deep lying position of this joint. As a result, acetaminophen remains a logical choice for initial treatment of hip OA. Several studies have compared acetaminophen to NSAIDs, including ibuprofen and naproxen, for early knee OA (117,118). In these studies, acetaminophen was found to be equally as effective as the NSAIDs but was associated with fewer side effects, mak- ing it the initial drug of choice for treatment of early knee OA pain. This evidence has since been refuted by recent meta- analyses that have found acetaminophen to be less effective than NSAIDs but more effective than placebo for relieving OA pain (119–122). Furthermore, a survey of OA and other rheu- matic disease patients found that there was a considerable preference for NSAIDs compared to acetaminophen (123,124). Although acetaminophen is less effective than NSAIDs, it is also less toxic than systemic NSAIDs, and therefore, acetaminophen is the most appropriate first-line oral analgesic for the treatment of OA (125).

Acetaminophen extended release, such as Tylenol Arthritis, provides convenience of use due to its tid rather than q4-6 dos- ing and may increase patient compliance. The recommended dose is one to two 650-mg extended-release tablets every 8 hours rather than one to two 325-mg standard-release tablets every 4 hours. A recent randomized controlled trial concluded that acetaminophen extended release (1,300 mg tid for a total daily dose of 3,900 mg) administered for up to 12 weeks was effective in treating moderate to moderately severe chronic OA pain of the hip or knee and was generally well tolerated (126).

Acetaminophen continues to be well tolerated by patients, but they should be made aware of acetaminophen dosing limi- tations due to potential hepatotoxicity, particularly if taken at excess doses for prolonged time periods as could occur in patients with a chronic, progressive disorder. Specifically, for patients with normal hepatic function, the safe dosing limit on average is felt to be 4 g/24 hours. This limit is lower in those with preexisting hepatic disease such that the maximum dose is tentatively set at one half of the maximum dose (i.e., 2 g) for those patients with normal hepatic function, and regular LFT monitoring is advisable. Since chronic alcohol abusers (defined as >3 drinks/day) are at increased risk of liver tox- icity from excessive acetaminophen use, these patients should be considered for the above dose adjustment (1). This recom- mendation may be overly cautious as reports of severe hepa- toxicity almost invariably involve acetaminophen doses that exceed the recommended ones and have often involved sub- stantial overdose.

Another important advantage of acetaminophen is that it rarely causes medication interactions. One important excep- tion is acetaminophen and the concomitant use of warfarin as acetaminophen can prolong the half-life of warfarin (127). Although this might only be clinically significant in patients at higher levels of anticoagulation who are chronically taking acetaminophen, the INR should be closely monitored (128). As the algorithm implies and epidemiologic studies have shown, eventually the knee or hip OA patient will require alternative or additional analgesia as the disease progresses and pain increases (129).

In summary, acetaminophen remains an inexpensive and well-tolerated medication (few side effects, toxicities, or drug- drug interactions) that provides modest pain relief for those with OA, and as a result, it should be considered a first-line

<table>
<thead>
<tr>
<th>TABLE 31.3</th>
<th>Risk Factors for Renal Toxicity in Patients Requiring Chronic NSAID or COX-2 Inhibitors</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Serum creatinine ≥ 2.0 and any one or more of the following:</td>
<td>• Hypertension</td>
</tr>
<tr>
<td>• Age &gt; 65</td>
<td>• Congestive heart failure</td>
</tr>
<tr>
<td>• Comorbid medical condition that adversely affects renal blood flow</td>
<td>• Diuretics</td>
</tr>
<tr>
<td>• Concomitant use of medications that alter renal prostaglandin synthesis</td>
<td>• ACE inhibitors</td>
</tr>
</tbody>
</table>
analgesic in OA and if necessary can be used as an adjuvant medication in more advanced diseases (130).

For patients with hip or knee OA but without risk factors for renal dysfunction, an oral anti-inflammatory agent can be considered. Trials and reviews have shown that NSAIDs have better efficacy than acetaminophen or placebo for OA pain (121,131–133), especially in patients in whom acetaminophen has been previously trialed without significant relief.

An important question is whether the patient is taking Coumadin. If yes, then celecoxib (Celebrex), the only COX-2 inhibitor that is currently available at the time of this writing, should be used instead of a traditional NSAID. Celecoxib does not inhibit thromboxane synthesis and therefore lacks an antiplatelet effect (134). However, since celecoxib prolongs the half-life of Coumadin on average by 11%, the INR must be followed appropriately, particularly during the first 2 weeks when celecoxib is begun or when the dose is changed.

A patient’s risk of gastrointestinal (GI) bleeding (Table 31-4) must be considered when deciding between an NSAID with or without gastroprotection and a celecoxib. Celecoxib is not as likely to cause GI toxicity because selective COX-2 inhibitors suppress COX-2 enzymes responsible for pain without affecting the COX-1 enzymes that regulate endogenous prostaglandin production that provides GI protection.

If GI bleeding risk factors are present, the physiatrist has three options to consider (see Fig. 31-2). One option is celecoxib since trials have found it to be as effective as nonselective NSAIDs (e.g., naproxen and diclofenac) but cause significantly fewer serious upper GI events (135). Another factor is whether the patient has a history of a sulfonamide allergy since this would eliminate celecoxib as an option. A second option is to combine an NSAID with a form of gastro-protection, which would eliminate celecoxib as an option. A second option would be to use a traditional NSAID (diclofenac) and Cytotec (misoprostyl) along with an enteric coating and should be considered if the decision has been made to pair a traditional NSAID with misoprostyl.

With respect to PPIs, comparisons between these found that they provide equivalent, effective, and well-tolerated prophylaxis against NSAID-induced GI lesions (140).

The third option is to choose a nonacetylated salicylate, which will be discussed later in the chapter. A final consideration in patients who need anti-inflammatory analgesics but are at very high risk for recurrent GI ulcer bleeding is combination therapy with a celecoxib and a PPI (141). The addition of a PPI to celecoxib should also be considered in patients older than 75 years, since additional GI protection is gained (142).

CV toxicity from traditional NSAIDs is another factor to consider when choosing this class of medications. One theory is that patients using propionic acid derivative NSAIDs (e.g., ibuprofen, naproxen) have the potential to interfere with the aspirin’s antiplatelet activity (143). Therefore, patients already on aspirin for known coronary artery disease may be at increased risk of myocardial infarction if taking a propionic acid derivative NSAID. Meta-analyses reveal that the risk of serious CV events exist for both selective and nonselective NSAIDs and that among the nonselective NSAIDs, diclofenac had the highest risk for an adverse CV event (144).

COX-2 inhibitors have received significant public attention due to their associated CV toxicity. Rofecoxib (Vioxx) was voluntarily removed from the market in 2004 after a clinical trial that was investigating the cancer prevention properties of Vioxx found an increased risk of myocardial infarction and stroke (145). Valdecoxib (Bextra) was then similarly removed due to increased CV risks. One possible explanation of increased CV toxicity is that selective COX-2 inhibition reduces levels of COX-2–derived prostacyclin that increases the risk of thrombosis (146). Celecoxib remains on the market, but it has two important black box warnings for CV and GI risks. Celecoxib has a lower CV risk profile but possibly an increased GI risk profile compared to previously available COX-2 inhibitors. Although this suggests lack of complete selectivity for COX-2 suppression with celecoxib, a meta-analysis revealed that celecoxib had a lower rate of GI bleeding and ulcers as compared to nonselective NSAIDs (147). In addition, celecoxib at supratherapeutic doses appears to have a safer cardiorenal profile as compared with traditional NSAIDs (e.g., ibuprofen or diclofenac) at standard doses (148). Research does suggest that COX-2 inhibitors may confer a decreased risk of CV events as compared with nonselective NSAIDs in aspirin users at high CV risk (149). Although COX-2 inhibitors appear to have a lower risk of CV events when compared to traditional NSAIDs, the recommendation is to use the lowest effective dose for the shortest duration (150), especially

<table>
<thead>
<tr>
<th>TABLE 31.4 GI Bleeding Risk Factors in Patients Chronically Taking NSAIDs</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Age &gt; 65</td>
</tr>
<tr>
<td>• History of PUD or UGI bleeding</td>
</tr>
<tr>
<td>• Concomitant corticosteroids or anticoagulants</td>
</tr>
<tr>
<td>• Smoking</td>
</tr>
<tr>
<td>• Alcohol use</td>
</tr>
<tr>
<td>• Poor general health</td>
</tr>
<tr>
<td>• Longer duration of NSAID treatment</td>
</tr>
</tbody>
</table>

in those patients with significant risk factors for CV events. Although current research provides some guidelines concerning the use of selective and nonselective NSAIDs and their inherent CV risks, future research will hopefully provide definitive answers regarding these issues.

Besides oral anti-inflammatory medications, two other options for delivery of anti-inflammatory medications include a topical NSAID and a topical NSAID patch. A topical prescription NSAID patch, Flector (diclofenac 1.3% eploline topical patch) patch, was FDA approved in 2008. Flector is indicated for the topical treatment of acute pain due to minor strains, sprains, and contusions. A randomized controlled trial also found that it was effective in the short-term treatment of pain due to symptomatic knee OA (151). Another diclofenac patch (diclofenac hydroxyethylpyrrolidine) was previously trialed and was found to be an effective and safe treatment for patients with symptomatic knee OA (151). Off-label use for patients with painful OA, particularly affecting the hands and knees, is likely to begin. Precautions and contraindications for Flector are similar to those for oral NSAIDs, and future research may reveal specific clinical situations where an NSAID patch is preferred to oral NSAIDs.

The algorithm in Figure 31-2 shows that for those patients whose pain and/or inflammation cannot be managed with oral anti-inflammatories, for those experiencing side effects that preclude further use of oral anti-inflammatory medications, or for those patients with renal risk factors that preclude the use of oral anti-inflammatory medications, the clinician can choose among several options. For patients with pain due to knee OA, these choices include viscosupplementation if it has not been previously tried or has been previously successful, or a corticosteroid injection with or without subsequent viscosupplementation, especially if there are signs of inflammation. Another option is tramadol (Ultram) or one of its variants such as its combination form with acetaminophen (Ultracet) or its once per day extended-release form (Ultram ER). Tramadol may be used in addition to NSAIDs or as an alternative to single therapy with high-dose NSAIDs (152).

The available literature to support use of tramadol compounds in the treatment of OA pain has grown since the last edition of this book. Studies have concluded that tramadol is a safe and effective option for the treatment of OA pain, more effective than placebo and comparable in efficacy to traditional NSAIDs (e.g., ibuprofen) (153,154). A Cochrane database systematic review concluded that tramadol or tramadol/acetaminophen decreases pain intensity, provides symptomatic relief, and improves function in patients with OA, but these benefits are somewhat small (155). It has been shown, however, to effectively treat breakthrough pain in OA that is moderately severe in intensity (156).

Tramadol’s relatively short duration of action (qid dosing) can be a disadvantage for a chronic painful condition, but tramadol extended release allows for qd dosing. Another advantage is that since the osteoarthritic patient population tends to be older, tramadol’s lack of GI bleeding potential and lack of renal toxicity are major advantages. In addition, it does not cause exacerbations of hypertension or congestive heart failure (157). GI side effects (e.g., constipation, nausea, and vomiting) and central nervous system (CNS) side effects (e.g., dizziness, somnolence, and headaches), however, were the most commonly reported adverse events (158,159). Dizziness remains a significant concern, especially for older patients, and can be detrimental to rehabilitative efforts. Dose titration does appear to lower the incidence and severity of these GI and CNS side effects and should be considered when prescribing tramadol, tramadol/acetaminophen, or tramadol extended release (160–162). Tramadol is associated with extremely low rates of abuse and dependence, no significant tolerance, and mild withdrawal symptoms when compared to other opioid compounds (163,164). Overall, the literature supports the use of tramadol for OA pain as it is a well-tolerated medication in comparison with other opioid compounds.

Ultracet (tramadol 37.5 mg and acetaminophen 325 mg) is another current option to consider. Because it was FDA approved in 2001 for the treatment of acute pain, it is not included in the 2000 ACR OA guidelines. It offers convenience for those patients who are taking both acetaminophen and tramadol. It provides similar levels of pain relief but with an improved side-effect profile as compared with stronger opioids (165–167). The literature also supports its use along with selective and nonselective NSAIDs. Adding tramadol/acetaminophen for patients with persistent OA pain despite ongoing therapy with nonselective NSAIDs or COX-2 inhibitors resulted in improved pain relief and improved physical function (168,169). Tramadol/acetaminophen was also shown to be an effective add-on therapy to help manage painful OA flares in the elderly and was generally well tolerated in this group of patients (170). As was noted above, a Cochrane systematic review concluded that tramadol/acetaminophen decreases pain intensity, provides symptomatic relief, and improves function in patients with OA, but these benefits are somewhat small (155).

Ultram ER or extended-release tramadol became available in 2005 and, therefore, was not included in the 2000 ACR guidelines on OA (1). Ultram ER is approved for the treatment of moderate to moderately severe chronic pain and is recommended in adults who require around-the-clock treatment for their pain for an extended time period (171). It is an option that offers a degree of convenience for those patients taking tramadol tid or qid. Ultram ER appears to be well tolerated for an opioid compound and has similar common adverse effects (dizziness, nausea, constipation, headache, and somnolence) as Ultram (171). It has proven to be safe for long-term use in patients with OA (172). It is well tolerated and more effective than placebo and is as effective as tramadol for treating pain associated with knee or hip OA (173–175). Overall, extended-release tramadol has been shown to effectively reduce pain and improve physical function in those with OA, and its convenient dosing schedule may allow it to play an increasing role in the treatment of OA (176,177).

If a patient continues to have pain despite one of the tramadol compounds and/or is intolerant to these, then opioid...
analgesics should be considered (see Fig. 31-2). Like tramadol, the lack of GI bleeding potential and lack of renal toxicity associated with pure opioid analgesics are a major advantage in elderly patients with OA. Certain groups of OA patients appear to be best suited for opioid analgesics, including those who are not good candidates for joint replacement surgery. For example, while some patients are too young for this surgery and would be best suited for surgery at an older age so that they would not be as likely to require revision surgery, others are “too old” and otherwise too limited in their daily activities to justify this form of surgery. Other patients are too medically ill for arthroplasty, and finally, some patients are too afraid to undergo surgery. Each of these patient groups should at least be considered for chronic opioid analgesics. Along these lines, the 2000 ACR guidelines have revised their wording from the previously published 1995 guidelines regarding this issue. The 1995 guidelines stated that: “Opioid analgesics, such as propoxyphene, codeine, or oxycodone should be avoided for long-term use, but short-term use may be helpful for the treatment of acute exacerbations of pain” (178). By contrast, the 2000 guidelines state: "Patients who do not respond to or cannot take tramadol and who continue to have severe pain may be considered candidates for more potent opioid therapy.” This change in wording is reflective of the general sentiment in musculoskeletal medicine that opioid analgesics may be appropriate in some chronic nonmalignant pain conditions. Although the consensus is that these medications can be used for chronic management of OA pain, a recent meta-analysis of randomized controlled trials concluded that the long-term efficacy and safety of these drugs for OA are yet to be determined due to the short mean trial duration (178).

Numerous studies have supported the safety and efficacy of opioid analgesics in patients with moderate to severe OA pain (179–182). They can be effectively used to treat OA pain that has not responded to combined acetaminophen, NSAID, or tramadol therapy. These patients may be started on low-dose opioids, and the dose may be titrated as needed and as tolerated (183). Unlike other medications that have a “therapeutic ceiling,” they can be titrated upward to achieve a satisfactory analgesic effect, as long as the patient can tolerate the dose (130). Another consideration is to add an opioid to ongoing NSAID therapy since this combination provides improved pain control and sleep quality (184,185). NSAIDs and opioids are also mutually dose sparing, and combining relatively low doses of a drug from each class provides a synergistic analgesia while limiting toxicity, which may improve patient tolerance and compliance with these medications (186).

There are a wide variety of choices in regard to opioid analgesics, including opioids of various durations and opioids of various routes of administration. One option that is gaining popularity in the management of chronic pain is topical patches, such as fentanyl transdermal (Duragesic). Clinical trials have demonstrated that transdermal fentanyl can significantly improve pain control, function, and quality of life for patients with OA of the knee or hip (187,188). Additional research has shown that therapy with transdermal fentanyl may be initiated in patients who were pretreated with nonopioid analgesics or weak opioids and can be tapered off without major complications (189).

Intra-articular opioids are another option to consider in the setting of an acute exacerbation of pain in patients with chronic OA. Previous studies have demonstrated significant pain reduction with intra-articular morphine as compared with intra-articular saline in the setting of chronic pain due to OA (190,191). Although effective, this option has only been shown to provide a short duration of benefit (5 to 7 days). Therefore, unless a longer duration of benefit can be shown, intra-articular morphine should be considered in the setting of an acute exacerbation of OA pain in patients in whom previous corticosteroid injections were ineffective or contraindicated (130).

A final option to consider for knee OA and an option for hand OA includes topical analgesics (see Fig. 31-2). The topical application of analgesic medications is a viable supplementary treatment that is available to patients with OA of joints that are superficial enough to allow medication penetration. By contrast, these medications are relatively inaccessible to deeper lying joints such as the hip and spine. Topical medications can be used as adjunctive therapy or sometimes as an alternative to oral medications. Topical medication administration has the advantage of local medication delivery to the joint at high concentrations with minimal systemic side effects. By contrast, only a small percentage of medication administered systemically reaches the joint. Capsaicin, topical salicylate, and topical NSAIDs are the three medication categories cited in the ACR guidelines. Since the time that those guidelines were published, topical lidocaine (Lidoderm patch) has become an off-label option.

Capsaicin, derived from jalapeno (aka “hot chili”) peppers, provides analgesia by reversible depletion of substance P and neuropeptide from unmyelinated C-fiber afferent nerve endings. There are various capsaicin-containing preparations, some of which contain other topically active substances. Because it often takes at least 2 weeks of substance P depletion before analgesia occurs and there is initial stinging from a flare response that occurs when neuropeptides are depleted from nerve endings, initial compliance during the first 2 weeks of treatment can be poor. There is some evidence that substance P enhances cartilage degradation (192), and therefore, an antisustance P agent such as capsaicin has been implicated as a potential disease-modifying agent in OA (193). Randomized controlled trials and a meta-analysis have found capsaicin superior to placebo for the treatment of OA pain, especially in patients with painful hand OA (194,195). Localized burning is the most commonly reported adverse effect and has been reported in approximately 50% of users (194,196). Because many patients cannot tolerate this adverse effect and as a result discontinue its use, capsaicin has been studied in combination with glyceryl trinitrate. Capsaicin combined with topical glyceryl trinitrate resulted in a decrease in analgesic consumption, likely due to the synergistic analgesic effect and improved tolerability of these two agents in combination as compared to either agent alone (197). Unfortunately, glyceryl trinitrate has been
shown to accelerate cartilage degeneration and subchondral bone sclerosis in the setting of ovine OA (198). Until additional evidence is available, it is unclear if capsaicin should be combined with glyceryl trinitrate. Another factor in patient compliance is the recommendation to apply topical capsaicin four times a day, which makes long-term compliance very difficult. Tapering the dose to twice a day has been suggested as a way of improving long-term compliance and reducing cost without losing efficacy (199). In summary, topical capsaicin remains an option as an adjunctive therapy in patients with OA of superficial joints (e.g., joints of the hand) if they are tolerant of its common adverse effect; although evidence-based support for its use in OA is limited, its potential role as a disease modifying agent is intriguing.

Over-the-counter topical salicylates have been used for years by patients with OA and other types of musculoskeletal pain. These preparations have been extensively marketed and include methyl salicylate (e.g., Bengay, Icy Hot, and Flexall) and trolamine salicylate (e.g., Sportscreme, Aspercream). Topical salicylates may be found alone or combined with other topical medications (e.g., capsaicin, camphor, and menthol) into topical creams, liniments, lotions, and rubs. In general, peer reviewed literature on the safety and efficacy in OA is limited. A systematic review of available clinical trials concluded that topical salicylates may be efficacious acute pain treatment, but they have only moderate to poor efficacy for musculoskeletal and arthritic pain (200). In addition, trolamine salicylate failed to demonstrate efficacy as compared with placebo in a randomized double blind crossover study in patients with knee OA (201). Overall, topical salicylates are well tolerated and adverse events have been rare in studies of acute pain; no adverse events occurred in the previously listed trials on knee OA, but topical methyl salicylate preparations may cause irritant or allergic contact dermatitis and anaphylactic reactions. Therefore, physicians should be aware of potential dangers from topical preparations containing methyl salicylate (202). In summary, topical salicylates when used properly are safe and may be beneficial in OA, although evidence-based support is very limited. The available literature also suggests that methyl salicylates are probably more effective than trolamine salicylates, especially when combined with penetration-enhancing substances (e.g., menthol and camphor).

The ACR’s inclusion of topical NSAIDs among its list of recognized medication categories is based in part on its efficacy in multiple placebo-controlled trials of topical NSAIDs in the treatment of osteoarthritic pain (203,204). Several short-term studies (approximately 2 weeks in duration) have concluded that topical NSAIDs help relieve knee OA pain compared with placebo (205–208). In addition, a meta-analysis concluded that topical NSAIDs are effective for a longer duration (>4 weeks) of pain relief in knee OA (209). Although there are significantly fewer controlled trials comparing topical NSAIDs with oral NSAIDs, these have generally shown equal efficacy and less GI side effects (210). In contrast to their relatively recent recognition in the United States, topical NSAIDs have been available in Europe for many years (211,212). In the United States, they have been used infrequently, and this may be because previously they were available only through compounding pharmacies and were rarely covered through traditional prescription medication plans. Voltaren gel (Diclofenac topical) received FDA approval in 2007 for the treatment of OA pain. Voltaren Gel is recommended for the treatment of OA pain in joints amendable to treatment with topical preparations, such as the knees and hands. Similar to capsaicin, it is to be applied four times a day. A dosing card onto which the patient squeezes the dose can help the patient to comply with the recommended dose of 4 g/day per lower limb joint and 2 g/day per upper limb joint. The most common adverse reactions were application site reactions (e.g., dermatitis), but the same warnings and precautions that exist for oral NSAIDs (e.g., CV, GI, and renal risks) exist for this topical preparation (213). Voltaren Gel has limited peer reviewed literature available to support its use, but previous studies on other diclofenac preparations have demonstrated safety and efficacy for the treatment of knee OA (214–216).

Another diclofenac product, diclofenac 1.3% (Flector patch), is indicated for minor sprains, strains, and contusions. Although it is not officially indicated for knee OA pain, a randomized controlled trial found it to be effective in short-term treatment of symptomatic knee OA (217).

Topical NSAID use in the United States will likely increase and gain increased acceptance amongst physicians and patients in part due to our aging population who will need treatment options for OA that offer a lower risk of adverse events. In a recent randomized controlled trial, the authors concluded that topical NSAIDs may be a useful alternative to oral NSAIDs for treatment of knee OA in the elderly since efficacy is equivalent and there are fewer minor side effects (218). As a result, physicians who treat OA should be very aware of topical NSAIDs as a possible treatment option.

The final topical option to consider is 5% topical lidocaine patch (Lidoderm patch). It was FDA approved in 1999 for the treatment of pain associated with postherpetic neuralgia but now is being used in an off-label fashion for a variety of painful musculoskeletal conditions, including OA. The lidocaine patch 5% releases lidocaine that penetrates into the skin and provides an analgesic effect without local anesthesia; and unlike transdermal medications, it does not result in clinically significant serum levels even in the setting of chronic use (219,220). Although its package label indicates that a maximum of three patches may be applied at a time and that it can be applied for 12 hours/day, patients report using it for up to 24 hours without adverse effects on their skin (221). In one published study on knee OA, the lidocaine patch 5% was effective as an add-on therapy for reducing pain and stiffness and improving physical function in a group of patients with knee OA pain that had not responded completely to prior analgesic therapy (222). Another study found that 2 weeks of treatment with the lidocaine patch 5% significantly improved the quality of pain, as assessed by a neuropathic pain scale, in patients with moderate to severe...
knee OA pain (223). In regard to patient tolerance, it is well tolerated with the main side effect reported in these clinical OA trials being an application-site local skin irritation, which was generally mild and usually did not require medication discontinuation. In regard to safety, no serious adverse events were observed during more than 120,000 patch hours involving patients who have used the patch in clinical trials for up to 8.7 years (219,224). In addition, no drug-drug interactions have been noted in clinical trials (220), but the manufacturer still recommends that the lidocaine patch be used cautiously in patients already on a class I antiarrhythmic (221). Controlled clinical trials are still needed to determine whether this agent is indeed effective in patients with OA of the knee and/or other joints, but it does appear to be reasonable to trial a topical lidocaine patch as an adjunctive agent in patients with OA who have failed to respond adequately to the previously listed agents.

The final topic of discussion for oral medications is that of pharmacologic agents that could play a role in the future management of OA. New NSAIDs, both oral and topical, and additional COX-2 inhibitors, if they in fact receive FDA approval, could provide additional treatment options. A new anti-inflammatory class, known as lipooxygenase (LOX)/cyclooxygenase (COX) inhibitors (e.g., licofelone), inhibits leukotriene synthesis so as to theoretically increase the anti-inflammatory effect while reducing the risk of prostanoid inhibition–related GI toxicity (225). If LOX/COX inhibitors become available in the United States, they may prove to be as effective yet safer than COX-2 inhibitors and/or traditional NSAIDs. Intra-articular (IA) injections that are currently available (e.g., corticosteroids, hyaluronic acids) may be redesigned into suspensions or hydrogels to achieve sustained release of the active ingredient to provide long-term pain relief and potential functional improvement (226). The most intriguing agents being investigated are disease modifying OA drugs (DMOADs), which target various steps in the cascade of events leading to arthritic changes in the joint. Clinical trials are currently underway for DMOADs, such as interleukin-1b inhibitors, and they may prove to be the most promising new agents available (227), since their goal is to halt the progression of OA.

Injectable Treatment Options
Corticosteroid injections should be considered for patients with hip and/or knee OA who continue to have pain despite acetaminophen, particularly if they are at particular risk for renal dysfunction (see Table 31-4; Fig. 31-2).

If the patient is deemed to have any of these risk factors for renal dysfunction, then consideration should be given for corticosteroid injections, especially for patients with hip or knee OA, and for viscosupplementation (the intra-articular administration of a hyaluronate-containing solution [aka hyaluronate, sodium hyaluronate, HA] into osteoarthritic joint synovial fluid). For knee OA patients, corticosteroid injections can be potentially used in conjunction with a course of viscosupplementation (see Fig. 31-2). The need for fluoroscopic or ultrasound guidance for hip corticosteroid injection very frequently precluded these injections in the past. However, with image guidance becoming increasingly available to musculoskeletal physiatrists, hip joint corticosteroid injections will probably be utilized more frequently.

A Cochrane review of 26 trials of IA corticosteroids for knee OA involving 1,721 patients found short-term efficacy with few side effects (228). IA corticosteroids were compared with placebo and viscosupplementation. IA corticosteroids were found to be more effective for pain reduction than IA placebo up to 3 weeks postinjection, but not from 4 to 24 weeks. No statistically significant differences were found between IA corticosteroids and viscosupplementation at 1 to 4 weeks postinjection. From 5 to 13 weeks, viscosupplementation showed superior efficacy for several efficacy parameters. These findings suggest a similar onset of action for IA corticosteroids and viscosupplementation but that viscosupplementation provides greater long-lasting benefit. Another meta-analysis on the efficacy of IA corticosteroids for knee OA showed benefit that lasted for only 2 weeks (229). Additional data from this meta-analysis suggested that high corticosteroid doses (i.e., doses of 50 mg equivalent of prednisone) may be required to achieve a longer response.

The analgesic mechanism of action of corticosteroid injections is not completely understood despite detailed study of this phenomenon (230–239). Furthermore, aspiration of an associated joint effusion as is generally performed at the time of the corticosteroid injection also probably contributes to the analgesic effect through removal of inflammatory mediators and a decrease in mechanical pressure from capsular distension. Although recommended intra-articular corticosteroid dosages have not been determined with certainty for any joints as intrasynovial steroid dose-response curves have not been developed, a generally accepted dosage range for the knee is 40 to 80 mg of methylprednisolone acetate (Depo-Medrol) or its equivalent. “Steroid arthropathy” (accelerated OA of a synovial joint) can potentially occur as a result of repeat injections but remains a very controversial topic with respect to incidence, mechanism of action, and the number of injections that should not be exceeded (240). There is laboratory evidence from studies in nonprimate OA animal models of a deleterious effect on proteoglycan synthesis that is believed to make cartilage more susceptible to injury (241–243). Another proposed mechanism is algiesia leading to a Charcot-like joint (243). The most common dosing limit quoted in the literature is 3 per year and 20 in a lifetime into any given joint (244).

In patients with knee OA who also have signs of inflammation, the sequential performance of a corticosteroid injection followed by a course of viscosupplementation is an option (245,246). The rationale is that the reduction of inflammatory mediators by the corticosteroid injection limits the degradative effect that these mediators subsequently have on the viscosupplement. Alternatively one can simply consider a course of viscosupplementation as there appears to be an anti-inflammatory effect from viscosupplementation. Basic science research has led to the discovery of several cell surface proteins with specific HA-binding sites that mediate biologic responses, and the
author as well as others have noted the “drying up” of initially inflamed knee joints during a viscosupplementation cycle (247–249).

Viscosupplementation is FDA approved for patients with knee OA who continue to have pain despite regular doses of a simple analgesic such as acetaminophen and nonpharmacologic therapy (see Fig. 31-2). Although there is evidence supporting several different potential mechanisms of action, which have been described in detail elsewhere, its exact overall mechanism is not known with certainty (Table 31-5) (250–311).

It remains controversial if viscosupplements exert a chondroprotective effect, even though there is evidence from in vitro studies, animal studies, and limited evidence in humans (312,313). Less controversial are the conclusions from several different investigators that viscosupplementation improves pain and function (314–317). There is also some evidence that the efficacy of a viscosupplement may be enhanced by concomitant therapeutic exercise, particularly quadriceps strengthening (318,319). Pain relief from viscosupplements has been shown to be of longer duration compared with other pain-relieving interventions such as arthrocentesis and corticosteroid injections (316,320). The only contraindication to viscosupplementation with a nonbacterially fermented viscosupplement (all current viscosupplements other than Euflexxa) aside from the contraindications to joint injection procedures in general includes the rare circumstance of true allergy to a viscosupplement. This situation is predicted by an allergy to avian proteins such as eggs or feathers.

At the time of this writing, there are five viscosupplements available for use in the United States. These include hylan GF-20 (Synvisc); three formulations of sodium hyaluronate that differ somewhat in molecular weight including Hyalgan, Supartz, and Orthovisc; and a bacterially fermented sodium hyaluronate preparation known as Euflexxa. Synvisc is chemically unique compared to the other viscosupplements in that its molecular weight is higher as it is chemically cross-linked with formaldehyde and divinyl sulfone. The question as to whether there is a correlation between clinical efficacy and higher molecular weight is controversial and remains unsettled by both individual studies and by meta-analysis, which noted significant interstudy heterogeneity in estimating efficacy (321–326). Cross-linking may be a disadvantage given the severe acute inflammatory reactions (SAIRs) seen with Synvisc that have not been reported with the noncross-linked viscosupplements. The exact mechanism of this side effect is unclear but possibly relates to the divinyl sulfone and/or formaldehyde cross-links.

Although efficacy of viscosupplementation is in part related to the radiographic stage of OA, with less efficacy seen for higher grades of OA (91% of grade I OA patients compared to 58% of grade IV OA patients reported that they felt much better or better), this is likely to be true of any nonsurgical intervention (327). Furthermore, there is evidence that viscosupplementation can delay the need for TKR in advanced OA patients (328–330).

### Other Image-Guided Injections for OA

Physiatrists are increasingly performing image-guided injection procedures, including fluoroscopy-guided and ultrasound-guided procedures, both into the spine and into the axial skeleton. The spinal injections that are performed in OA patients with osteoarthritic-related spinal pain include epidural injections and facet-related procedures. For example, patients with cervical and lumbosacral spinal stenosis often receive these injections as a symptomatic treatment for associated radicular and axial pain. Facet joint injections, medial branch blocks, and radiofrequency ablation are performed on patients with facet joint OA.

While most of the spinal injection procedures in OA patients utilize fluoroscopic guidance, ultrasound is becoming increasingly available. Thus, injection procedures on peripheral osteoarthritic joints such as the hip and knee are now being performed utilizing either ultrasound or fluoroscopic guidance. Image guidance is also particularly helpful into other peripheral joints that may be otherwise difficult to inject due to severe osteoarthritic narrowing.

There are several reasons why it is likely that image-guided injection procedures will become increasingly performed by physiatrists who treat OA patients. First, as the population continues to age, the incidence of OA is also increasing. Second, physiatrists are becoming more adept at image-guided injections as part of their treatment armamentarium. Finally, studies on viscosupplementation into other osteoarthritic joints continue to be conducted. If the FDA eventually approves viscosupplementation for other joints besides the knee, this will most likely lead to an increased performance of image-guided hip joint injections.

### REFERENCES


245. Donnelly S. Personal communication; 1999.


249. Schlesinger N. Personal communication; 1999.


Disorders of the Cervical Spine

The intricately designed and highly mobile cervical spine is routinely subject to unique mechanical stresses that predispose the spinal elements to degenerative change. Three common symptom complexes are associated with the degenerative cascade of the cervical spine. These include cervical axial pain, radiculopathy, and myelopathy (1,2). These syndromes may occur in isolation or in combination. Axial pain, arising from degenerative discogenic or zygapophyseal joint pain generators, is more commonly described in the middle decades of life (3). Spinal nerve root pathology and radicular syndromes may arise from an acutely offending disc injury or more gradually evolving and degenerative neural foraminal compromise (4,5). Myelopathic symptoms occur when the spinal cord is affected by central canal stenosis (6–8). This chapter will review these common degenerative cervical syndromes, the cervical spine in rheumatoid arthritis (RA), and rehabilitative approaches.

ANATOMY AND BIOMECHANICS

The cervical spine is composed of seven vertebrae and five intervertebral discs. The C1 and C2 segments are anatomically unique. The C2 through C7 bodies articulate anteriorly through the intervertebral discs and uncovertebral joints, or joints of Luschka, and posteriorly through the zygapophyseal joints (Fig. 32-1). The occipito-atlanto-axial or Oc-C1-C2 complex is a specialized upper cervical segment that allows for a significant range of motion between the head and upper torso (9). The articulations of this segment are composed of synovial joints with no intervertebral discs. The basic structure of this segment is a biconcave ring of C1 interposed between the convex condyles of the occiput (Oc) above and the C2 lateral masses below. The odontoid, or dens, projects upward from the C2 vertebrae, providing a post to which the C1 ring and Oc are directly anchored (9). The Oc is attached to the odontoid predominantly by the alar ligaments and also by the less significant apical and upper arms of the cruciform ligament (10). The C1 ring is bound to the odontoid by the sturdy transverse arm of the cruciform ligament as well as the accessory C1-2 ligaments and the C1-2 joint capsules (11) (Fig. 32-2).

In the lower cervical spine, the intervertebral discs contribute to approximately one fourth of the height of the cervical column (12,13). The dimensions of the cervical disc, which is thicker anteriorly than posteriorly, contribute to the cervical spine’s lordotic curvature (14). The cervical disc allows for a greater degree of motion than do the discs in the lumbar region, as the disc-to-vertebral body height ratio in the cervical region is 2:5, compared with a 1:3 ratio in the lumbar spine. The anterior longitudinal ligament runs along the anterior vertebral body and discs, providing structural support in limiting cervical extension. The posterior longitudinal ligament supports the disc and body posteriorly, stretching in flexion and relaxing with extension. The superficial ligamentum nuchae is a dense midline band that extends from the Oc to the spinous process of C7. Proceeding ventrally in the midsagittal plane, the supraspinous ligament, interspinous ligament, and ligamentum flavum are encountered in order, each contributing to stability in flexion (9).

The external component of the intervertebral disc is comprised of the annulus fibrosus, which consists of multiple lamellae of type I and II collagen. The annulus encapsulates the internal nucleus pulposus. The fiber direction of the annular lamellae alternates, with each consecutive layer oriented opposite to the adjacent fibers (14). This structural arrangement uniquely allows the annulus to accommodate angular motion while providing stability against torsion and shear. The intervertebral discs are separated from the vertebral bodies by the end plates (13,15). The end plates are composed of hyaline and fibrocartilage and form a permeable surface through which nutrients may pass between the cancellous bone of the vertebral body and the intervertebral discs. The end plates are vascularized during fetal life, but their vessels involute during the first 10 to 15 years of development (16). Thereafter, the essentially avascular intervertebral disc receives its nourishment through the end plates and vessels circumscribing the outer annular fibers (12,14). The nucleus pulposus is a semifluid gel, composed predominantly of water, which represents 40% to 60% of the intervertebral disc and allows for deformation and accommodation of movement and compressive loading. With aging, the nucleus is slowly replaced by fibrocartilage, and by the end of the first decade of life, the distinction between the nucleus and surrounding annulus begins to lessen (12). The innervation of the outer one third to half of the annulus fibrosus has been demonstrated (17,18). The anterior disc is innervated by branches of the vertebral nerves and sympathetic...
trunks. The posterior disc, posterior longitudinal ligament, anterior dura, and dural root sleeves are innervated by a posterior plexus derived from the sinuvertebral nerves (19) (Fig. 32-3).

The uncovertebral joints or joints of Luschka are anterior articulations found between the third through seventh vertebrae that share in load bearing and are often affected by degenerative change (12). The zygapophyseal joints are planar synovial joints, which are created by the inferior and superior articular processes of adjacent vertebrae. Each joint is composed of circular or ovoid facets that are covered by articular cartilage and enclosed by a fibrous joint capsule (20). The zygapophyseal joint may include fibroadipose menisci located at the ventral and dorsal joint poles. These menisci are thought to be drawn from the joint cavity during joint gliding to cover those articular surfaces that become exposed during such motions (21) (Fig. 32-4). The cervical zygapophyseal joints are innervated by articular branches originating from the medial branches of the cervical dorsal rami. Each cervical zygapophyseal joint receives a dual innervation from the medial branches arising from the dorsal rami above and below. The medial branches of the C3 dorsal ramus are unique. A deep C3 medial branch contributes to the C3-4 joint innervations, whereas the larger superficial medial branch of C3, also known as the third occipital nerve, innervates the C2-3 joint (20). Beyond the C2-3 joint, the third occipital nerve innervates the semispinalis capitis and supplies cutaneous sensation to the suboccipital region. This is the only dorsal ramus below C2 that has a cutaneous distribution (22).

From C2-3 through C5-6, the zygapophyseal joints are angled at approximately 45 degrees, with the longitudinal axis of the cervical spine, with a typically steeper angle at C6-7 (23). The superior articular processes are progressively taller at the more caudal segments, with the upper ends of the joints extending further above the level of the segmental intervertebral disc. The orientation of the zygapophyseal joints allows them to resist both forward and downward displacement of the vertebral body. As the zygapophyseal joint of the upper cervical segments are relatively more horizontally oriented, they contribute more to bearing axial loads than do the more caudal segments. The orientation of the zygapophyseal joints dictates the nature and magnitudes of movement in the cervical spine (20).

The primary motion of the Oc-C1-C2 complex is flexion and extension at both Oc-C1 and C1-2, axial rotation only at C1-2, and minimal lateral bending at Oc-C1 (9). The Oc-C1 articulation allows for approximately 10 degrees of flexion and 25 degrees of extension. Rotation of 45 degrees in either rotation can be observed at the C1-2 joint (24,25). A useful approximation of the range of motion allowed by each segment between C2-3 and C7-T1 is 10 degrees of total range from flexion to extension, as well as 10 degrees of lateral bending and axial...
rotation to each side. Flexion and extension and axial rotation tend to be the greatest in the midcervical region, decreasing both above and below. Lateral bending is typically greatest at C2-3 and diminishes caudally (26). Instability of the cervical region has been defined as a horizontal displacement of the vertebral body during flexion and extension of greater than 3.5 mm or angular change of greater than 11 degrees (27).

The neural elements of the cervical spine include the spinal cord, dorsal and ventral roots, spinal nerves, and dorsal and ventral rami. The dural and arachnoid mater sleeve contain the ventral and dorsal roots and ultimately blend with the spinal nerve epineurium. The spinal nerves are formed by the joining of the ventral and dorsal roots. Each spinal nerve exits the spinal canal through the intervertebral foramen, which is bordered by the uncovertebral joint anteromedially, the zygapophyseal joint posteriorly, and superiorly and inferiorly by the pedicles of the respective vertebral bodies. The most cephalad and first true neural foramen is located at the C2-3 level. This foramen has the largest area, with a progressive reduction in foraminal size observed more caudally (13). The spinal nerve is a mixed nerve that resides within the neural foramen and is accompanied by the radicular arteries and veins (28). The C3 through C7 nerves exit above their respective pedicles, whereas the C8 spinal nerve exits beneath the pedicle of C7. The C1 nerve rests on the posterior arch of C1, where it divides into dorsal and ventral rami, and the C2 nerve exits the thecal sac and descends obliquely across the dorsal aspect of the atlantoaxial joint (29). Upon exiting the intervertebral foramen, the spinal nerves divide into dorsal and ventral rami. The C5-T1 ventral rami contribute to the brachial plexus. The C1-4 ventral rami comprise the cervical plexus, which innervates the cervical musculature and cutaneous structures of the ear, face, and neck. The C1 and C2 ventral rami innervate the atlanto-occipital and atlantoaxial joints, respectively (19).

**AXIAL PAIN AND SYMPTOM REFERRAL**

**Epidemiology**

Cervical spondylosis describes the degenerative change that affects the five articulations of the cervical segment, including the intervertebral discs, the bilateral zygapophyseal joints, and the uncovertebral joints of Luschka (30). The degenerative cascade is thought to begin with disc desiccation and loss of disc height, which is followed by an approximation of the uncovertebral joints and a disruption of the normal zygapophyseal joint biomechanics. Uncovertebral and zygapophyseal joint hypertrophy, osteophyte formation, anular disruption, and ligamentum flavum hypertrophy can comprise the ensuing phases of degenerative change (31–33). Radiographic evidence of degenerative change is appreciated in 10% of individuals by the age of 25, 35% by age 40, and in up to 95% by the age of 65 (34–36). Plain film imaging in asymptomatic subjects reveals degenerative change affecting the cervical spine in 70% of women and 95% of men between the ages of 60 and 65 (37). It is essential to remember that although there is typical radiographic evidence of degenerative disease, including loss of disc space height and age-dependent osteophytes, in individuals presenting with cervical pain,...
all individuals with degenerative findings are symptomatic (3,37). The incidence of neck pain similarly increases with age, progressing in a fairly linear fashion from age 20 through 60 (3,38). Approximately 95% of individuals will experience cervical pain by the age of 65 (35). Severe episodes of neck discomfort will affect 10% of the population at some point in their lives. The annual incidence of significant cervical pain has been estimated at 12.3% (36).

Referred pain is pain perceived in a region innervated by nerves other than those innervating the true source of the pain. Pain referral patterns arise as the brain is unable to decipher the true pain origin secondary to convergence at the level of the spinal cord and thalamus. Referred pain is typically experienced as deep, diffuse, and poorly localized pain (19). Any of the innervated structures of the cervical spine can contribute to local pain as well as symptom referral. Although injury to the cervical musculature and ligamentous structures can result in local and regional pain (39,40), the intervertebral disc and zygapophyseal joints have been more extensively investigated as pain generators.

**Intervertebral Disc**

Degenerative disc disease is most commonly observed at C5-6, with C6-7 involvement the next most prevalent. These intervertebral discs might be more commonly affected because of increased segmental motion at these spinal segments (41,42). An injured disc, resulting from degenerative change or more acute trauma, can produce local and referred symptoms (17). The outer third of the annulus fibrosus houses nerve endings that can be stimulated during injury. It has been suggested that degenerative or traumatic alteration of the internal architecture of the annulus can result in pain production through stimulation of local mechanoreceptors and nociceptors (17,41). The ability of the annulus to mediate pain has been demonstrated during surgery through mechanical and electrical stimulation of cervical intervertebral discs (42). In addition to this mechanical component of disc injury, discogenic pain might also be biochemical in origin. An annular defect could allow for the migration of nuclear material, which might then stimulate the outer annulus, dura mater, posterior longitudinal ligament, dorsal root ganglion, or spinal nerve (43). Increased levels of inflammatory mediators have been identified in degenerative and herniated discs when compared with asymptomatic controls (44).

Using cervical discography, characteristic discogenic pain patterns have been described (42,45,46). The correlation between cervical disc magnetic resonance imaging (MRI) abnormalities and pain-generating potential observed through discography has been shown to be poor (46). In a study of 807 disc injections, 404 concordant pain responses were used to describe pain referral patterns. Pain referral to the scapular region was observed to arise from disc stimulation at the C3-4 through C7-T1 levels. The C5-6 and more caudal segments were noted to refer symptoms to the upper limb, and C6-7 stimulation was unique in that pain was referred to the anterior chest wall (45) (Fig. 32-5).

**Zygapophyseal Joints**

The zygapophyseal joints can also become active pain generators in the setting of degenerative change or following trauma.
These posterior elements are particularly vulnerable to injury during a whiplash event. Although often not observed radiographically, zygapophyseal joint fractures, intra-articular hemorrhage, and capsular tears have been observed in pathologic studies (47–53). The prevalence of whiplash-induced zygapophyseal joint pain has been described (54–56). Using diagnostic blocks, chronic cervical zygapophyseal joint pain has an estimated frequency of 54% to 60% (55,56). In this chronic neck pain population, the C2-3 joint has been observed to be most frequently symptomatic, followed by C5-6 (55,56). Of those patients with chronic neck pain, 58% to 88% describe significant headaches (54–56). The prevalence of C2-3 joint-mediated headaches has been estimated at 50% to 53% in those individuals with a chief symptom of headache following whiplash (54,56). The Oc-C1 joints (54,57–59) and C1-2 joints (57,58) have similarly been described as active nociceptors in cervicogenic headaches. Studies in a limited number of asymptomatic and symptomatic volunteers have suggested particular resultant pain patterns from intra-articular zygapophyseal joint stimulation and anesthetization (60–62). Pain arising from C2-3 stimulation was described as referred to the upper neck and occipital region, whereas C3-4 and C4-5 joint pain was observed to affect the mid- and lower cervical region, extending toward the superior scapular border. Pain from C5-6 stimulation was observed to extend toward the shoulder, whereas C6-7 pain radiated over the more caudal scapular region (Fig. 32-6). These degenerative or posttraumatic pain referral patterns arising from primary axial zygapophyseal joint or discogenic pain generators can be difficult to distinguish clinically from one another and from radicular pain presenting without a more distal component or neurologic correlate.

A randomized controlled study investigating the therapeutic benefit of intra-articular zygaphophyseal joint injections using corticosteroid has suggested that such treatment is ineffective (63). A retrospective uncontrolled study of intra-articular C2-3 joint injection in the treatment of posttraumatic headaches has suggested a significant therapeutic response. A randomized placebo controlled trial of radiofrequency neurotomy in patients with unremitting cervical zygapophyseal joint pain has demonstrated a significant treatment effect (64). Pain is expected to return following neurotomy, as distal axons innervating the zygapophyseal joint are likely to regenerate.

**Radiculopathy**

**Epidemiology**

Cervical radiculopathy can arise from pathologic compressive processes affecting the nerve root, including acute disc herniation, degenerative foraminal stenosis, trauma, or tumor (65). In the 1920s, cervical nerve root compression was described as a cause of anginoid pain (66). Cervical disc herniation resulting in cord compression was initially recognized as a syndrome of spinal cord tumors or chondromas. This syndrome of cord compression was subsequently defined by Mixter and Ayer in 1935 (67) as arising from a ruptured intervertebral disc. In 1936, cervical spondylosis and neural irritation was described as a cause of shoulder girdle and arm pain (68). In the early 1940s, cervical nerve root irritation arising from disc injury was defined (69).

The annual incidence of cervical radicular pain is 5.5/100,000 (70). These injuries represent 5.3% of all nerve root injuries caused by disc pathology. The nerve roots most commonly involved by cervical radiculopathy are C7 and C6, with most studies suggesting C7 syndromes as the most common, followed by radiculopathy of C5 and C8 (65,71–74). Patients younger than 55 years are more likely to present with radiculopathy arising from acute disc herniations, whereas those older than 55 years are more likely to demonstrate symptoms arising from degenerative foraminal or central canal stenosis (30).

**Pathophysiology**

The cervical nerve root can be compressed in the neural foramen by the intervertebral disc, degenerative change affecting the zygaphophyseal or uncovertebral joint, or in a combined fashion. Cervical radiculopathy arising from a herniated disc is more common than radiculopathy arising from the more slowly evolving degenerative foraminal stenosis typically observed in older individuals (75–78) (Fig. 32-7). Disc herniations have been classified as “soft” and “hard.” “Soft” refers to more acute disc injuries and displacement of nucleus pulposus
FIGURE 32-7. T2-weighted axial MRI section at C5-6 demonstrating degenerative spondylosis and resultant advanced right-sided neural foraminal stenosis, delineated by arrowheads, in a patient with a C6 radiculopathy.

material, whereas “hard” disc material describes a calcified and spondylotic ridge (72,79,80). With disc degeneration, tears in the annulus can allow for the displacement of the nucleus. Anterior disc herniations are less common as the anterior longitudinal ligament is wider and stronger than the posterior longitudinal ligament (30).

Disc herniations have been characterized based on location (81,82). The three locations described are intraforaminal, posterolateral, and central. Intraforaminal herniations are the most common and may result in an acute radiculopathy affecting the nerve root exiting through the respective foramen. Therefore, an intraforaminal herniation of disc material at the C4-5 level might result in a C5 radiculopathy. The posterolateral disc herniation occurs because the rhomboidal shape of the posterior longitudinal ligament tends to direct disc material to one side or the other. These herniations are located between the lateral edge of the posterior longitudinal ligament and the posterior aspect of the uncinate process. The central disc herniation passes through the substance of the posterior longitudinal ligament. Such lesions are more likely to result in central canal compromise and spinal cord compression. Central herniations are believed to be more likely to occur in the later stages of cervical degeneration when uncovertebral joint hypertrophy serves as a barrier to the more lateral migration of disc material (83) (Fig. 32-8).

The resultant injury to the nerve roots likely arises from combined mechanical and biochemical pathophysiologic processes. A study in which cervical disc specimens removed surgically from patients with radiculopathy were compared with discs from a traumatic control group found significantly increased levels of matrix metalloproteinase, nitric oxide, prostaglandins, and interleukins in the radiculopathy group (44). These findings, in conjunction with a plethora of literature describing the combined mechanochemical nature of lumbar radicular syndromes, suggest that these biochemical markers are likely involved in the injury process of cervical radiculopathy (84–88).

Radiologic Findings

Plain films can be used as an initial screening tool but are not necessary in every individual presenting with cervical or radicular pain. In addition to anteroposterior and lateral images, flexion and extension views can prove of value in patients with recent trauma, suspected instability, or in cases of ankylosing spondylitis or RA (65).

MRI is considered the imaging modality of choice in cervical radiculopathy (89). Earlier comparative studies demonstrated a close correlation between findings on cervical MRI, computed tomography (CT) scan, and myelography (90–92). More recent studies have suggested a higher correlation between radiographic and surgical pathology when MRI was used than with either myelography (93) or CT (94). A high incidence of abnormalities has been observed on cervical MRI imaging of asymptomatic individuals. Boden et al. (95) observed significant abnormalities in 19% of 63 asymptomatic patients.

In those younger than age 40, 14% had a herniated disc or neural foraminal stenosis, and similar abnormalities were observed in 25% of individuals older than age 40 (83). An investigation of cervical myelography has identified nerve root filling defects in 21% of asymptomatic subjects (96).

As there is a significant prevalence of such abnormalities in asymptomatic subjects, advanced cervical spine imaging in symptomatic individuals must be interpreted carefully to determine if a correlation exists between radiographic pathology and clinical findings (97,98). Also of interest is the natural radiographic history of symptomatic disc pathology in the cervical spine. A study following “soft” disc herniations in 21 patients with radiculopathy found that the largest lesions decreased in size the most, with 16 of 21 lesions reducing in size by 35% to 100% at 15-month follow-up CT (99). In similar studies observing patients with serial MRI, 40% to 92% of patients demonstrated radiographic regression (100,101). More lateral and extruded discs were observed to be more likely to resolve (Fig. 32-9).

**Electrodiagnostics**

Electromyography (EMG) studies can help to localize the level of a cervical radiculopathy and differentiate such a presentation from a brachial plexopathy, more distal entrapment, or peripheral neuropathic process. Such testing is not necessary in all patients presenting with cervical radicular syndromes. Electrodiagnostic studies might prove particularly useful in individuals with multiple levels of radiographic pathology whose physical examination findings are less conclusive in identifying the segmental pathology of clinical significance. In other cases, a patient might present with clinical findings suggestive of radiculopathy but without a clear radiologic correlate. In these cases, electrodiagnostic studies might similarly prove useful. The earliest abnormality that might be observed in the setting of motor root compromise is reduced voluntary recruitment observed during needle exam. Abnormal activity at rest in the form of positive sharp waves or fibrillation potentials, indicative of axonal loss, might not be observed until at least 18 to 21 days after the onset of a radicular syndrome (102).

---

**FIGURE 32-9.** Sagittal (A) and successively sliced axial (B) MRI performed 1 month (1M), 3 months (3M), and 5 months (5M) after the onset of a C5 radiculopathy in a 56-year-old man. Radiographic regression of a lateral extrusion with superior and inferior extension is demonstrated. (From Mochida K, Komori H, Okawa A, et al. Regression of cervical disc herniation observed on magnetic resonance images. *Spine*. 1998;23(9):990–997, with permission.)
EMG abnormalities have been shown to correlate well with myelographic abnormalities and operative pathology (103). Needle examination of seven limb muscles is likely adequate for the identification of segmental pathology (104). The segmental sensitivity and specificity of EMG abnormalities might only approach 67% and 50%, and EMG abnormalities can be appreciated at asymptomatic levels (105). The sensitivity and specificity of EMG abnormalities have not been studied against an agreed-upon gold standard. The false-negative rate of electrodiagnostic studies is likely higher than desired (106).

A diagnosis of cervical radiculopathy is established in a patient when corroborative radiographic findings are observed in association with a suggestive pain distribution and/or a myotomal strength deficit, reflex abnormality, or sensory loss. If the cervical radiographs and clinical findings are less conclusive, electrodiagnostic studies might prove helpful in confirming a level of radicular pathology when a myotomal denervation pattern is observed. If radicular pain remains within the differential and these diagnostic criteria are not satisfied, a diagnostic selective nerve root block might be employed. Although the specificity and sensitivity of lumbar diagnostic injections have been described as ranging from 87% to 100% (107,108), the utility of diagnostic selective cervical injections has not been as well-defined (109).

## History

Patients with radiculopathy typically present with a chief report of pain, weakness, paresthesias, or a combination of sensorimotor deficits (65). The majority of patients describe cervical and upper-extremity pain that began without trauma or a particular inciting event (72,73,110). Most studies suggest that pain affects the upper limb more commonly than the neck, although most often both are involved (73,79). Coughing, sneezing, or a Valsalva maneuver may lead to symptomatic worsening. The reports of pain might also incorporate the anterior chest wall, resulting in a syndrome of pseudoangina pectoris (67,111–113). It is often difficult to determine the symptomatic level based on the patient’s pain description alone. The distribution of symptoms might not follow the classic dermatomal patterns as defined by Keegan and Garrett (114) or Foerster (115). The limitations of these mapping studies include the assignment of dermatomal distribution based predominantly on the loss of sensation or hyposensitivity arising from compressive injuries. Each of these classic studies describes considerable dermatomal overlap. Dermatomal pain patterns, arising from selective cervical spinal nerve stimulation, have recently been described and may be more representative of the pain and paresthesia patterns clinically observed in radiculopathy (116). Using fluoroscopically guided cervical nerve root stimulation, pain patterns were described after the stimulation of 134 roots in 87 subjects. Some of the highlights of this study included pain distribution distal to the elbow in only 14% of C5 stimulations, C6 affecting the ulnar hand in 67%, C7 uniquely resulting in anterior head symptoms and most commonly referring pain to the chest, and C8 affecting the thumb in 14%. No nerve root pain distribution was observed to be confined to the classically defined dermatomes in greater than 50% of stimulations (116) (Fig. 32-10). Some of the discrepancy between cervical dynatome and the more classic dermatomal maps might also be explained by intrathecal cervical dorsal root intersegmental anastomosis, which can be present in as many as 61% of individuals (117–119) (Fig. 32-11). Additionally, pain in radiculopathy might arise in part from a myotomal or myalgic pain component arising from ventral nerve root irritation (120).

## Physical Examination

When examining the patient with suspected cervical radiculopathy, it is essential that the clinician remains mindful of extraspinal disorders that can mimic nerve root pathology. A brachial plexopathy or neuralgic amyotrophy with upper-trunk involvement can appear similar to a radiculopathy of C5 or C6 origin (121–123). More localized disorders affecting the upper limb can similarly result in regional pain, including subacromial
bursitis, lateral epicondylitis, and bicipital tendinitis (65). More distal neural entrapments affecting the median, radial, or ulnar nerve should also be considered. A median neuropathy at the wrist can result in pain extending as proximal as the cervical region (65). When cervical radiculopathy and myelopathy coexist, such as in amyotrophic lateral sclerosis and syringomyelia, the examiner should seek signs of possible combined cord and root compression (65,124).

Muscle girth should be carefully inspected for any evidence of atrophy that might be consistent with a more chronic radiculopathy. Although C5 and C6 involvements can lead to wasting of the biceps and periscapular musculature, loss of triceps girth can be observed in the setting of a C7 radiculopathy. Atrophy of the hand intrinsics might be observed in patients with C8 or T1 pathology. Active cervical range of motion can be recorded and observed for symptom provocation in a particular plane. In particular, extension and ipsilateral side bending are more likely to result in a reproduction of radicular symptoms. Normal range of motion of the neck, using an inclinometer, has been estimated at 60 degrees of flexion, 75 degrees of extension, and 45 and 80 degrees of bilateral side-bending and rotation (125).

Manual muscle testing has been suggested to be a more sensitive means of isolating a level of segmental pathology than either reflex or sensory testing. Variations in the contributions of the spinal nerves to the cords and branches of the brachial plexus can lead to an altered pattern of muscular innervation in the upper limb and perhaps more perplexing motor testing findings (65,74). Reflex testing of the upper extremity should include the bicep, tricep, brachioradialis, and Hoffman’s response. Lower-extremity reflex testing should also be included to rule out hyperreflexia or long tract signs. Gait should similarly be observed to rule out ataxia, which might be observed in the setting of cervical cord compromise.

Spurling’s maneuver incorporates rotation and side-bending of the head toward the affected side, and this can be combined with axial compression. The goal of this maneuver is to diminish the foraminal area and reproduce radicular symptoms (30). This neck compression test was initially defined by Spurling and Scoville in the 1940s at a time when cervical radicular syndromes were first recognized (126). This maneuver has been demonstrated to have a high specificity but a low sensitivity (110). Such maneuvers should be performed with caution in patients with a suspected radicular syndrome, as this could lead to further neural irritation. In those patients demonstrating a reproduction of radicular symptoms with active cervical extension or ipsilateral side-bending, this provocative test should be avoided (65). The shoulder abduction relief sign has the patient place the affected hand on the top of the head. This position of upper-extremity abduction reduces stress on the nerve root affected by a disc herniation and relieves pain (127,128).

A complete exam should also include passive ranging of the shoulder and impingement maneuvers. Tinel’s testing at the supraclavicular fossa, elbow, and wrist and carpal tunnel stress maneuvers, such as Phalen’s testing and carpal compression, might reveal a proximal entrapment neuropathy or carpal tunnel syndrome. Observing for a pain response during palpation of the medial and lateral epicondyle can be helpful in ruling out a contributing focal tendinitis. Thoracic outlet maneuvers should also be considered to rule out a contributing lower-trunkplexopathy or dynamic vascular insufficiency, also referred to as “neurogenic” or “vascular” thoracic outlet syndromes. These regional exam techniques are included in the complete examination of a patient with suspected radiculopathy and might enlighten the examiner to a contributing musculoskeletal disorder or more distal neuropathic process.
Nonsurgical Treatment

Lees and Turner studied 57 patients with a diagnosis of cervical radiculopathy for a period of up to 19 years (35). Patients in this study pursued a variety of treatments, including rest, manipulation, therapeutic exercise, and soft-collar use. Of these 57 patients, 45% described one primary pain episode without a future symptomatic recurrence. Of the remaining patients, 30% described a continuation of mild symptoms, and 25% reported a persistence of significant symptoms or worsening. Of the 57 patients, 32 remained significantly improved, whereas 18 described their condition as static or symptomatic. In their series of patients with radiculopathy, none became myelopathic. Cervical collars appeared to correlate with initial symptomatic relief, but end results were similar among all patients studied.

A review of multiple retrospective and uncontrolled nonsurgical cervical radiculopathy studies revealed satisfactory outcomes in 80% to 90% of patients (73,129,130). Several randomized controlled studies in patients with lumbar syndromes have demonstrated an efficacy of nonsteroidal anti-inflammatory agents and muscle relaxants. Studies investigating the efficacy of oral corticosteroids have been inconclusive (131,132). In a prospective study of 26 consecutive patients treated nonsurgically, 20 achieved good to excellent results with an average of 2.3 years of follow-up. Oral analgesics, collars, traction, and physical therapy were employed in this treatment group. Twenty-two patients used oral steroids, and nine received epidural injections. Patients with extrusions were noted to fare better than those with contained protrusions (133). In a randomized prospective study, 81 patients were treated with anterior decompressive surgery, physical therapy, or collar use. The surgical group demonstrated significantly greater reductions in pain at 16 weeks, and both the surgical and physical therapy groups demonstrated improved functional abilities as compared with the collar group. No significant difference in outcome as measured by pain and functional profiles was observed between the groups at the final 12-month follow-up evaluation (134).

In a prospective study of 155 patients with cervical radiculopathy, one third were treated surgically, whereas two thirds were prescribed medical treatment that included oral analgesics and steroids, traction, and epidural injection. At 1-year follow-up, significant improvements in pain and functional status were observed in both groups, with a greater degree of improvement in the surgical group. Twenty-six percent of the surgical group continued to experience severe pain (135).

Cervical epidural injections can be employed in those individuals with severe pain or symptoms that have failed to respond to an initial trial of spine-specific physical therapy, oral anti-inflammatory agents, and analgesics. A recent prospective, randomized, double-blind, controlled study has demonstrated the efficacy of fluoroscopically guided lumbar transforaminal injections in the treatment of radiculopathy (136). The use of cervical epidural injections has arisen from the described success of therapeutic injections in the treatment of lumbar radicular syndromes. Although controlled trials are lacking, a review of the literature is inconclusive, suggesting good and excellent outcomes ranging from 0% to 69%. Outcomes are typically better in individuals with radicular as opposed to axial pain, and patients with normal radiographs likely fare worse (137).

In a prospective trial of 68 patients with cervical radiculopathy, transforaminal and translaminar epidural injections were employed with 76% of patients describing complete pain relief after 39 months of follow-up (138). In additional uncontrolled prospective studies evaluating the efficacy of transforaminal injections performed with fluoroscopic (139) and CT guidance (140) in patients with cervical radicular pain resistant to prior medical treatment, significant and lasting pain relief was realized without associated procedural complications (Fig. 32-12). In a retrospective study of 20 patients with radicular pain arising from degenerative foraminal stenosis, satisfactory outcomes were realized by 60% of patients with a mean follow-up of 21.2 months (141). A retrospective and uncontrolled study of patients with traumatically induced radiculitis and an absence of radiographic pathology suggests transforaminal injection procedures result in good or excellent outcomes in only 14% (142). In experienced hands, a very low incidence of complications arising from cervical transforaminal injections has been described (143,144). Dedicated reviews (145) have more extensively assessed the literature in terms of therapeutic outcomes and potential complications arising from cervical transforaminal injections. The more recent literature has highlighted rare but catastrophic complications arising from transforaminal injection, in some cases arising from more grossly misplaced needles or ill-advised technique. In other instances, concern has been raised regarding the particular nature of corticosteroid solutions and the potential for resultant arterial occlusion and infarct of neural tissue (146,147). Ultimately,
additional outcome studies assessing cervical transforaminal injection therapy, and in an injectate specific fashion (148), will be required to further clarify the balance between the potential therapeutic benefits and procedural risks to patients with cervical radicular pain. New injectable medications, beyond the more commonly used corticosteroids and anesthetics, which specifically target particular components of the inflammatory and pain-generating cascade (e.g., tumor necrosis factor or glutamate), might someday be introduced (149,150).

For those patients whose debilitating pain persists despite a full trial of nonsurgical treatments, surgical intervention should be considered. Surgery should be considered without delay in those patients with a pronounced and rapidly progressive neurologic deficit. Significant weakness and weakness, atrophy, and numbness that present without pain are also relative indicators for surgical intervention (106,151).

Surgical Treatment

Surgical procedures using a posterior approach for cervical spine decompression evolved during the 1940s. Decompression using an anterior approach was described in the late 1950s (152). Anterior cervical discectomy with fusion (ACDF) was initially described by Smith and Robinson (153) as an approach to provide both decompression of the neural elements and stability to the diseased segment. Indications for this approach include progressive or persistent symptoms arising from unilateral or bilateral lateral disc herniations or spondylotic neural foraminal compromise at one to two levels. ACDF is contraindicated in patients with congenital stenosis, stenosis arising predominantly from posterior structures, and disease at greater than three levels (154). The advantages of this approach include the ability to remove anteriorly offending structures without disturbing the cord, resultant neural foraminal distraction, distraction of the segment with an associated reduction in ligamentum flavum buckling, and segmental stabilization (30). Disadvantages of ACDF include graft donor site pain, the need for postprocedural immobilization, the possibility of graft dislodgement, and subsequent adjacent segment degeneration. When used for radiculopathy, ACDF has been described to result in excellent outcomes in 53% to 91% of patients. Multilevel surgery is associated with poorer outcomes (155–158).

Laminoforaminotomy is an alternative and posterior surgical approach to the patient with radiculopathy. This approach is particularly indicated in patients with soft lateral disc herniation or zygapophyseal joint arthrosis who present without significant neck or periscapular pain (30). The advantages with this approach include an avoidance of spinal destabilization, fusion-related complications, and the need for postoperative immobilization (78). Disadvantages include a greater level of postoperative neck pain compared with ACDF during the recovery period and an inability to decompress the cord. This approach is contraindicated in patients with cord compression or cervical kyphosis. Laminoforaminotomy has been associated with 75% to 96% excellent results (159,160). In a study of 846 patients with cervical radiculopathy arising from soft or hard disc material, good to excellent results were observed in 96% (159). An additional study of 170 patients studied postoperatively for 15 years demonstrated a satisfactory result in 86% of patients with a 6% symptomatic recurrence rate (161).

MYELOPATHY

Pathophysiology

Disease entities that can present in a similar fashion to cervical myelopathy include motor neuron disease, multiple sclerosis, and other demyelinating conditions, peripheral and entrapment neuropathies, intracranial pathology, spinal cord tumors or syrinxes, and other systemic causes of hyperreflexia (30). Neurologic symptoms of cervical myelopathy can arise during the cervical degenerative cascade from stenotic compromise of the central canal and spinal cord. Similar to findings in the lumbar spine, narrowing of the cervical central canal can arise from developmental abnormalities or acquired change. Cervical spondylosis remains the most common form of central canal compromise and resultant myelopathy. Disc space narrowing leads to a loss of the normal cervical lordosis and buckling of the posteriorly located ligamentum flavum. Combined with flaval buckling, osteophytes from the uncovertebral joints, zygapophyseal joints, and vertebral bodies contribute to the central canal compromise (162). Spinal cord compression arising from spondylosis is typically slowly progressive. Patients are often noted to have significant radiographic compression while remaining asymptomatic. The cord is apparently able to withstand significant chronic deformation without resultant dysfunction (152). In addition to direct compressive effects, dynamic stressors, vascular insufficiency, and ischemia likely contribute to the pathophysiology of myelopathy. In cervical extension, the central canal diameter decreases and the cord may be pinched between the anterior osteophytic vertebral bodies and discs and the hypertrophied posterior elements and ligamentum flavum. During flexion, although the central canal may widen, the cord can become tethered over spondylotic anterior elements (152,163). Instability may also contribute to spinal cord impingement. As spinal segments are stiffened by spondylotic change, adjacent motion segments might also develop relative hypermobility and even subluxation. Myelopathy can also evolve in the setting of ossification of the posterior longitudinal ligament. In ossification of the posterior longitudinal ligament, ligamentous ossification can occur at one or multiple levels. The ossified posterior longitudinal ligament can develop into a bulbous mass that contributes to anterior cord compression. Asian populations, such as the Japanese, have a particularly high incidence of ossification of the posterior longitudinal ligament as compared with other populations. The etiology of this condition remains less clear, but there is likely an associated genetic influence (152).

Pathologic cord findings from myelopathic patients include white and gray matter destruction with demyelinating changes ascending and descending from the site of injury (152,163). An examination of pathologic specimens and correlation with the degree of cord compression has revealed degeneration of the
lateral white matter tracts in mild to moderate compression and necrosis of the central gray matter in more severe compressive injuries. Interestingly, the anterior white columns remain relatively spared, even in more severe cord compression. Histologic changes observed include axonal demyelination followed by cell body necrosis and scarring or gliosis. Cystic cavitation can also be observed within the gray matter. This central destruction is believed to arise from ischemic changes arising from cord deformation (164). Bulging disc material and anteriorly located spurs might also compress the anterior spinal artery, contributing to a vascular myelopathy (152,163). The vascular supply to the gray matter has been shown to arise from arterioles branching from the anterior spinal artery. With compression of the cord in the anterior to posterior plane, these branches are subject to mechanical distortion, which can lead to relative ischemia of the gray and medial white matter (165). Atherosclerosis and intimal fibrosis have also been observed in the regional spinal arteries (152).

The corticospinal tracts are typically involved early in the evolution of myelopathy, resulting in lower-limb weakness. Subsequent posterior column dysfunction results in ataxia manifesting as a wide-based gait (30). A classification system has been proposed that includes five cervical myelopathic syndromes (166). According to this categorization, the most common presentation is the transverse lesion syndrome. In this scenario, the upper limbs are relatively spared with symptoms arising in the lower extremities from corticospinal tract, spinohalamic tract, and posterior column involvement. In the motor system syndrome, symptoms can resemble those of amyotrophic lateral sclerosis. In this scenario, the corticospinal tracts and anterior horn cells are affected by the compressive injury. The patient characteristically presents without sensory symptoms, and upper- and lower-limb weakness can present in combination with spasticity and gait disturbances. The third variant is a central cord syndrome, which is similar to the posttraumatic incomplete spinal cord injury. In these patients, upper-limb involvement predominates, and the prevailing hand weakness has a poor potential for recovery. The upper-limb syndrome of myelopathy has also been described as cervical spondylotic amyotrophy and myelopathy hand (167,168) (Fig. 32-13). In the fourth presentation, Brown-Séquard syndrome, unilateral cord compression and corticospinal tract involvement results in an ipsilateral hemiparesis, and spinohalamic tract compromise results in contralateral sensory disturbances. In the brachialgia cord syndrome, long tract signs present in combination with radicular pain arising from spinal nerve root compression (30,166).

**Radiologic Findings**

Although plain radiographs can demonstrate degenerative changes associated with aging and spondylosis, this imaging modality does not offer visualization of the neural elements. When neurologic symptoms are appreciated on exam, they might be localized to the site of pathology appreciated on plain films (169,170). In lateral plain films of patients with cervical spondylosis, anterior osteophytes are often larger than those appreciated posteriorly but typically do not cause symptoms. Posterior osteophytes are more clinically relevant because of the possibility of associated neural compression. The presence of such plain film findings should be regarded with caution, as degenerative change is commonly observed in asymptomatic subjects and increases in frequency with increasing age. In some cases, ossification of the posterior longitudinal ligament can be visualized on lateral images as a bar of bone extending along the posterior aspect of the vertebral bodies. Flexion-extension views are necessary to diagnose instability that may not be evident on a neutral lateral view (152). The typical cervical lordosis measures 21 degrees, ±13 degrees. Cervical lordosis decreases with age and degenerative change. Additionally, 9% of asymptomatic individuals can demonstrate a cervical kyphosis. The normal adult cervical canal demonstrates a midsagittal diameter of 17 to 18 mm (range from 13 to 20 mm). This measurement is taken from a point at the posterior vertebral body margin to a corresponding point on the laminar line. Cord compromise is believed to occur when the central canal is narrowed to between 10 and 13 mm. At this stage, stenosis is defined as relative. Absolute stenosis is defined by a central canal with an anteroposterior dimension of less than 10 mm. Such canal compromise is highly correlated with spinal cord compression. Cervical extension maneuvers further reduce the canal diameter, and the cord can be dynamically compressed with such maneuvers.

The most accurate means of determining spinal canal dimensions is by CT scan. Ratios have also been described that compare the central canal size to the width of the vertebral body. A ratio of less than 0.85 has been defined as consistent with stenosis, with a ratio of less than 0.8 predisposing the patient to myelopathy. It is believed that such comparisons of vertebral body and canal dimensions are not as reliable in larger individuals (30,37,171–173). MRI is the standard imaging modality in the assessment of patients with cervical myelopathy. The advantages of MRI imaging include superior visualization of the disc and neural elements.
as well as the ability to visualize intrinsic changes within the cord (30). In some instances, in the presence of more long-standing cord compression, cord atrophy might be observed. MRI also allows for identification of parenchymal changes, such as syrinx formation and abnormal signal within the cord, which may be representative of myelomalacia (Fig. 32-14).

Abnormal signal within the cord does not necessarily correlate with clinical outcome, but it does serve to identify pathologic cord change (152). In a recent study that attempted to correlate preoperative cord signal abnormalities with surgical outcome, it was determined that increased T2 cord signal likely represented a wide spectrum of compressive pathology and was therefore not a useful predictor of clinical outcome (174). It has also been suggested that increased T2 signal within the cord with well-defined borders might correlate with a worse outcome than when such T2 signal within the cord is less well-defined and more faint in appearance (175).

In a study of 52 patients with mild cervical myelopathy treated nonsurgically and followed for a mean of 3 years, increased T2 signal within the cord was not found to clearly correlate with worse clinical outcomes. In 18% of patients initially observed, increased cord signal was observed to naturally regress over time (176). In patients with significant upper-limb weakness and atrophy, also known as cervical spondylotic amyotrophy, high T2 signal intensity has been observed within the cord in the region of the anterior horn cells (Fig. 32-15). Such findings suggest that the lower motor neuron component, including upper-limb weakness and atrophy, in these individuals might arise from anterior horn cell damage induced by dynamic cord compression and circulatory compromise (167). The disadvantages of MRI include a more limited visualization of the neural foramen and a relative inability to distinguish between soft and hard disc materials and an ossified posterior longitudinal ligament. Therefore, some practitioners prefer CT myelography, which can offer greater differentiation of soft-tissue and osseous structures in the assessment of patients with cervical disorders (30).

History
The clinical presentation of myelopathy can be quite variable (177). Patients commonly describe paresthesias, which can occur in a fairly diffuse and nondermatomal distribution in the upper extremities. Many patients are not aware of their weakness but do describe subtle difficulties with balance and gait alteration. As cord compression and myelopathy become more severe, gait and balance difficulties become more pronounced. Fine motor coordination of the hands is adversely affected, with reports of difficulty manipulating buttons or handwriting abnormalities. Patients might report an inability to rapidly open and close the hand (178). Arm weakness may be unilateral or bilateral. Proximal lower-limb weakness usually prevails over a loss of more distal strength. For example,
patients might describe difficulty arising from a chair, but a foot drop would be a more rare presentation. In this regard, cervical myelopathy presents almost oppositely to the weakness associated with lumbar spinal stenosis. Bowel and bladder abnormalities may occur in more severe compressive syndromes, but these are rarer. Although patients will often present with neck pain, as many as 50% will not. As many as 15% of individuals with more advanced myelopathy do not report cervical discomfort, often leading to a delayed diagnosis (152,166,179). Individuals with associated radicular involvement may be more likely to report upper-extremity pain (30).

Physical Examination
In examining the patient with suspected myelopathy, the neck may be palpated for areas of tenderness. Cervical extension may be restricted and painful for patients with cord or nerve root compression (152). L’Hermitte’s sign, initially described in 1932 (180), is performed by passively flexing the neck with the patient in the seated position and causing an electric-type sensation that progresses along the trunk and to the upper and lower limbs. This may be appreciated in 10% of patients with spondylotic myelopathy. A positive response can also be observed in individuals with multiple sclerosis and spinal tumors (166,181).

Wasting of the shoulder girdle musculature might be appreciated in patients with spastic paraparesis at the C4-6 levels secondary to loss of local anterior horn cell function. Similarly, fasciculations might be noted in the upper-limb musculature. Wasting of the intrinsic hand musculature might be observed (152).

The patient can be asked to hold the fingers extended and adducted. If the two ulnar digits are observed to flex and abduct within 30 to 60 seconds, this is referred to as a finger escape sign, which may be indicative of a myelopathic process. Additionally, the patient should be able to rapidly open and close the hand repetitively up to 20 times in a 10-second interval. A more pronounced radiologic abnormalities, including cord atrophy, might result in a scapulohumeral reflex, marked by elevation of the scapula and abduction of the humerus (30,182). Gait should also be assessed for the ability to heel-and-toe walk, as well as heel-to-toe ambulation. Romberg testing, during which the patient stands with the eyes closed and the arms held forward, is a test of position sense. Loss of balance during this exam maneuver may be consistent with posterior column dysfunction (152). If indicated, electrodiagnostic studies might prove helpful in distinguishing a myelopathic presentation from a diagnosis of carpal tunnel syndrome, ulnar nerve entrapment, thoracic outlet syndrome, or a contributing peripheral neuropathy.

Natural History
Previous studies have described a stepwise and progressive deterioration in patients with cervical myelopathy left untreated (2,183). Lees and Turner (35) observed a natural history of myelopathy, which was rarely rapidly progressive in terms of symptoms and disability. Only 5% of patients demonstrated such deterioration. In this series, 75% reported a stepwise and episodic increase in symptoms. In 20% of patients, symptoms were more slowly progressive. Age greater than 60 years was a poor prognostic factor. Of 15 patients with severe disability from myelopathy, 14 remained severely disabled at long-term follow-up. Nurick (184) studied 56 patients treated nonsurgically for cervical myelopathy over 20 years. He similarly described episodes of symptomatic progression followed by intervals of more static symptoms. In this series, mild myelopathic symptoms carried the best prognosis for a more prolonged course with extended periods of nonprogressive disability. Epstein and Epstein (185) reviewed findings in 114 patients treated nonsurgically and described clinical improvement in 36%, no change in 38%, and worsening in 26%.

Patients with radiologic evidence of cord compression but no signs of myelopathy can be observed and monitored both clinically and radiologically. An exception might be made in a patient with such pronounced cord compression that even minor superimposed trauma could be particularly likely to result in cord injury. Although it might be less likely for such a patient to demonstrate no symptoms of myelopathy, these individuals do at least need to be instructed regarding an avoidance of activities that might lead to a hyperextension-type injury. For those individuals with milder myelopathic symptoms, such as hyperreflexia and more subtle gait abnormalities, the anticipated clinical course suggested by the literature should be reviewed with the patient. If the patient has not demonstrated a progressive deterioration, there may be a role for nonoperative management with neurologic reevaluation every 6 to 12 months (152).

Favorable prognostic indicators include symptoms of less than 1 year’s duration, unilateral motor deficits, the presence of L’Hermitte’s sign, and a younger age at initial presentation (83).

Surgical Treatment
Most authors recommend surgical intervention for patients with moderate to severe myelopathy. For patients with similar levels of myelopathic symptoms, operative intervention might be recommended earlier in those individuals with more pronounced radiologic abnormalities, including cord atrophy,
abnormal cord signal, or cervical kyphosis. Surgery is indicated in an attempt to halt further progression of neurologic deficits. The degree of ultimate recovery typically correlates with the severity of preoperative myelopathy. Other positive prognostic factors include larger transverse cord area, younger age, single rather than multiple levels of involvement, and shorter symptom duration (152,179,183).

Both anterior and posterior surgical approaches have been employed in the treatment of cervical myelopathy. Posterior options include multilevel laminectomy, laminectomy with fusion, and laminoplasty. Anterior approaches include single- or multiple-level anterior disectomy with fusion and corpectomy with strut graft fusion. Anterior approaches can be performed with or without anterior instrumentation. The surgical approach to the myelopathic patient is determined by the nature of the central canal compromise and the experience of the treating surgeon. Factors considered in choosing the most appropriate procedure include the number of pathologic levels, the sagittal alignment of the cervical spine, the location of the compressive lesion, and the presence of instability (152).

In the degenerative cervical spine, cord compression typically arises from anterior structures. The anterior surgical approach allows for direct decompression of the thecal sac and cord. If the cord compression arises from protruding disc material at one or two levels, an anterior disectomy with fusion may be performed. If spondylotic disease or ossification of the posterior longitudinal ligament results in compression above or below the disc space, partial or complete resection of the vertebral body may be necessary to achieve adequate anterior decompression.

Bohlman (186) described 17 patients undergoing one or more level ACDF for cervical myelopathy. Two thirds of patients were observed to demonstrate improved ambulation. Fourteen of seventeen had improved upper-limb strength and two thirds improved with regard to ambulation (30,152). A study following 87 patients after anterior decompression for myelopathy described complete motor recovery in 54 patients, partial improvement in 26, no change in 6, and worsening in 1 patient (179).

Patients with more widespread stenosis or cord compression arising from ligamentum flavum buckling may be ideal candidates for a posterior decompressive approach (152). Multiple-level laminectomy may also be indicated in patients with congenital stenosis; stenosis with predominantly posterior compression; or when technical problems, such as obesity, a short neck, or anterior soft-tissue cervical pathology, might interfere with an anterior approach (30). Posterior techniques are more indirect in their effect in that they require the cord to shift posteriorly within the thecal sac to diminish the compression arising from anterior structures. For this to transpire, there needs to be a preoperative straight or preferably lordotic cervical sagittal alignment. A cervical kyphosis will not allow adequate shift of the cord dorsally after posterior decompression. This is an essential determination in deciding between anterior and posterior approaches in the treatment of myelopathy. Multiple-level laminectomy is contraindicated in patients with cervical kyphosis and in children. Laminectomy is specifically contraindicated in children because of the high risk of postlaminectomy kyphosis deformity. Similarly, if instability is present, laminectomy alone will lead to further destabilization and therefore would need to be performed along with posterior fusion (30,152). When performed with a concomitant fusion, the incidence of postlaminectomy kyphosis may be reduced. The advantages of multiple-level laminectomy include its relative simplicity as opposed to anterior corpectomy as a means of addressing multiple levels of compressive pathology. Additionally, the increased pseudarthrosis rate associated with more extensive anterior fusion procedures can be avoided. Laminoforaminotomy can be added at any level, thereby addressing an associated radiculopathy. The disadvantages of the posterior approach include the associated destabilization of the spine and postoperative kyphosis, which is directly related to the extent of zygapophyseal joint resection performed (187–189). The results of laminectomy procedures for myelopathy do vary considerably, with a review of the literature suggesting a 68% to 85% good or excellent outcome (183,187,190–192).

Laminoplasty was developed by the Japanese to achieve posterior decompression and eliminate postoperative instability and kyphotic deformity by expanding the central canal with retention of the posterior elements (193,194). In this approach, the lamina is opened in a hingelike fashion with the patency of the open-door hinge maintained with either a suture technique or grafting (195,196). A unilateral hinge technique or a midline spinous process splitting with bilateral hinges can be performed (193,197,198). Indications for this approach include developmental stenosis, three or greater level degenerative stenosis, and ossification of the posterior longitudinal ligament (199,200). Contraindications to laminoplasty are similar to those described for multiple-level laminectomy. The advantages of this approach include reduced postoperative instability, zygapophyseal joint preservation, and the avoidance of more extensive anterior fusion procedures. The disadvantages include incomplete decompression, recurrent stenosis, and a higher rate of C5 nerve root palsy (30). A study of 25 patients with myelopathy who underwent laminoplasty demonstrated improved gait in 84%, reduced sensory symptoms affecting the hands in 87%, and improved bowel and bladder function in 77% 18 months following surgery (201). An additional study of patients followed for approximately 13 years after laminoplasty revealed a 75% rate of useful neurologic recovery (202).

RHEUMATOID ARTHRITIS

Although the cervical spine is commonly affected in patients with RA, the exact incidence of involvement is not clear (203). It has been reported that the radiographic abnormalities in the cervical spine can be observed in 19% to 88% of patients with RA (204,205). Persistent inflammation of the joints of the cervical spine predisposes the patient to eventual structural damage. The primary risk factor for cervical involvement in RA is the severity of the disease, including the presence of erosions, use of steroids, and a positive rheumatoid factor (206,207). The various subluxation syndromes, which may occur in isolation
or combination, arise following damage to cartilage, erosion of bone, and ligamentous laxity (203). Although fatal neurologic complications may occur in rare cases, it appears that cervical spine disease is indirectly associated with higher mortality in RA because it occurs with more aggressive disease (208).

The most common spinal abnormality in RA is anterior atlantoaxial subluxation (AAS), with a reported frequency of 19% to 70% (205). Inflammation of the synovial-lined space between the odontoid and the transverse ligament may result in laxity of the transverse ligament (203). This can result in anterior translation of the C1 vertebra. This displacement can become particularly pronounced during flexion. Using lateral radiographs, separation between the anterior aspect of the odontoid and the posterior inferior tubercle of C1 by greater than 3.0 mm in women and 2.5 mm in men is indicative of AAS (209). Patients with AAS might describe retroorbital, temporal, or occipital pain or an abnormal sensation during neck flexion (203). Myelopathy can arise from compression of the spinal cord between the odontoid and the displaced posterior arch of C1 (203). Vertebrobasilar insufficiency can also arise from AAS, as the vertebral arteries can become kinked around the displaced C1 vertebra with resultant vertigo or more rare instances of impaired consciousness (210). As opposed to anterior displacement, posterior C1-2 subluxation is a rare complication of RA (205). In this subluxation syndrome, erosion of the odontoid is so extensive that the arch of C1 translates posteriorly, potentially compressing the cord.

Vertical displacement of the odontoid into the foramen magnum has been described as cranial settling, vertical subluxation, and atlantoaxial impaction (203). Such displacement can be observed in 3% to 4% of the general RA population and likely occurs in approximately 22% of RA patients with cervical spine involvement (211,212). This translation results from bone erosion, cartilage loss, and eventual collapse of the articulation between the lateral masses of the atlas and occipital condyles and the atlantoaxial articulations. The odontoid migrates superiorly with respect to the Oc and enters the foramen magnum, potentially compressing the medulla (213,214). Projection of the odontoid tip greater than 4.5 mm superior to McGregor’s line (the line connecting the most caudal point of the occipital curvature to the dorsal border of the hard palate on lateral radiographs) is consistent with vertical subluxation (203). A myelopathy, or more rarely, life-threatening brainstorm compression can result (203,213,214). Lateral or rotary subluxation of C1 on C2 can result from asymmetric erosion and collapse of the C1-2 articulations. Such subluxation can be defined on anteroposterior plain radiographs if the lateral masses of C1 are displaced greater than 2 mm lateral to the C2 masses (212). Patients with this complication may present with limited head rotation or with the head tilted toward the side of the affected joint (205,215,216).

In 7% to 29% of patients with RA, subluxations occur below the C2 vertebral with C2-3 most commonly affected. Serial subluxation can result in a characteristic “staircase” appearance to the cervical spine (212). These subluxations result from laxity of the longitudinal ligaments and facet joint incompetence. Myelopathy is less common in these cases and only arises if there is more severe dislocation of a segment or infiltration of the cord by pannus originating from the uncovertebral joints (205,209). Spondylodiscitis is a painful condition that involves erosion of the end plates with loss of disc space height without associated osteophytic change. The pathogenesis of these lesions are unclear but is believed to arise from either primary inflammation arising in the uncovertebral joints (217) or a more traumatic origin in the setting of cervical instability (218,219). Spondylodiscitis may lead to neurologic compromise and can be difficult to distinguish from infection (205).

As neurologic deterioration is less common in milder subluxation syndromes, patients with mild AAS or vertical subluxations and stable neurologic deficits can be managed with careful observation. In more advanced subluxations or with progressive neurologic deficits, the threshold to operate should be reduced, and surgery may be lifesaving in cases of more rapidly progressive neurologic deterioration (206). There is also a role for surgical intervention in patients with advanced subluxation who are stable neurologically but have intractable neck pain (220). In children with juvenile RA, 60% to 70% will eventually demonstrate cervical spine involvement. Although the spinal pathology in these patients can be similar to that in adults, growth disturbances and zygapophyseal joint fusion can also result. Cervical fusion is most commonly observed in patients with systemic or polyarticular disease, and hypoplastic vertebrae can arise at the levels of fusion (221).

**MECHANICAL INTERVENTION, THERAPEUTIC EXERCISE, AND MODALITIES IN THE TREATMENT OF CERVICAL AXIAL AND RADICULAR PAIN**

Physical therapy and modalities are frequently prescribed for the patient with cervical pain and radicular syndromes. A wide range of therapeutic exercise programs have been employed in the treatment of this patient population. Clinical outcomes can be optimized by proceeding with a comprehensive diagnosis and symptom-specific approach and with the judicious use of modalities when indicated.

To date, there has only been one prospective, randomized, controlled study comparing the outcomes of nonsurgical and surgical treatments of cervical radiculopathy (134). Treatment consisted of an anterior cervical discectomy and fusion without postoperative rehabilitation, 15 individually adapted physical therapy sessions, or 3 months of cervical collar use. Physical therapy protocols were not uniform and could include any of the following: ultrasound, moist heat, massage, cold, manual traction, mobilization, neck and shoulder stretching, isometric strengthening, aerobic exercises, and ergonomic or postural training. This study concluded that long-term results of surgical and nonsurgical treatments for cervical radiculopathy may be no better than the natural disease history alone. Limited conclusions can be derived regarding the efficacy of physical therapy in the treatment of these patients, as treatment was not uniform and protocols were not particularly individualized.
Mechanical Therapy

The comprehensive mechanical assessment of the patient with a cervical disorder begins with having the patient actively perform repeated end-range movements and sustained positioning in both loaded and unloaded fashion. The performance of end-range movements might identify rapid responders to mechanical therapy and separate slow responders from nonresponders. The results of such movement testing will help both the patient and the treating therapist understand the patient’s tolerance of movement and likely positional bias (222). The intertester reliability in interpreting pain location and intensity change (223,224), the effects of repeated movements on symptoms (225,226), and the significance of these findings in predicting the patient's response to treatment (227,228) have received support in current literature. A recent study of patients with discogenic lumbar and extremity pain suggests that a comprehensive mechanical exam may be superior to MRI in identifying symptomatic discs and determining anular competence (229).

Robin A. McKenzie is responsible for developing the mechanical paradigm for the diagnosis and treatment of spinal disorders. This method enables the evaluating clinician to assign a directional preference of motion and instruct the patient in repetitive movements that can decrease spinal pain associated with soft-tissue dysfunctions, structural derangements, or postural syndromes (230). Treatment is based on the individual’s signs and symptoms during the movement assessment. Specific spinal motion can affect the patient’s baseline symptoms by increasing, decreasing, producing, abolishing, centralizing, or peripheralizing pain. It is also of particular importance to observe whether pain production is evident during a given movement or at end range. Information gathered from the initial mechanical assessment will enable the clinician to categorize patients into one of three different syndromes: postural, dysfunction, or derangement (230).

The postural syndrome is typically seen in younger patients or in individuals with a relatively sedentary occupation or lifestyle. Pain may be isolated to a given region, or cervical, thoracic, and lumbar spine pain can be simultaneously described. Postural syndromes are not associated with constant or referred pain, and repeated end-range movements will not have any effect on symptoms. Postural syndrome pain is reproduced only by sustained positions that are believed to produce tension on relatively healthy spinal tissues that are held at end range. Simply correcting faulty postures can remove such tissue stressors and abolish pain. This positive effect can often be seen on day 1 of treatment. Patients who present with pure postural syndromes should be made aware of the possibility of developing more persistent pain syndromes in the future if these positional issues are not addressed.

Patients presenting with a dysfunction syndrome will invariably have pain provocation when attempting end-range movement in a particular plane. Symptoms are believed to occur at end-range because of the motion restriction caused by mechanically stressing shortened and sensitive spinal tissues demonstrating reduced elasticity. The location of symptoms is always local, with the exception of dysfunctions that similarly involve the cervical nerve root. When the involved root lacks mobility or glide, a full dermatomal referral of symptoms may occur once the nerve root is similarly stressed. Similar to axial and local dysfunctions, pain produced in nerve root dysfunctions will abolish once the offended tissue is placed on slack. Treatment of this syndrome addresses remodeling and lengthening the adaptively shortened structures. Manual techniques consisting of end-range mobilizations can prove helpful when performed in conjunction with a daily stretching program (230).

The derangement syndrome is the most common and dynamic lesion encountered in the cervical spine. The term derangement refers to disrupted or displaced spinal tissue. Although the spinal disc is the most likely tissue to derange, the clinician should be aware that the facet joints, spinal ligaments, and other surrounding soft tissues have the potential to be similarly disrupted. Unique to this syndrome is the phenomenon known as centralization, first described by McKenzie in 1956 (230,231). Repeated end-range movements that cause symptoms to progressively decrease or change from a peripheral to a more central location are considered desirable and should be used as a basis for self-treatment (232). By contrast, movements that appear to increase symptoms or result in a peripheralization of pain should be discouraged (Fig. 32-16).

The ability to successfully centralize symptoms allows the therapist to treat the patient with the appropriate postural bias, advise the patient to avoid positions that aggravate the condition, and potentially predict a favorable outcome from rehabilitation (233,234). The first goal in the treatment of a derangement syndrome is to reduce the displaced or disrupted tissue. On occasion, specific manual techniques can be performed to expedite the reduction process. Subsequently, the patient is instructed in a specific home exercise program in order to maintain the reduction. Once the derangement is reduced and maintained for at least 3 days, the patient is introduced to progressive spinal loading strategies in order to ensure a proper restoration of motion and function. The final component of treatment includes prophylaxis exercises and providing information to prevent injury recurrence.

In summary, the McKenzie method is a dynamic and individually tailored approach that incorporates the patient's response to a given set of repeated movements or sustained positioning. The most powerful aspects of this paradigm are that it uses readily reproducible treatment strategies and centers the patient as an active participant in the ongoing rehabilitation process. This varies considerably from more passive therapy, which deemphasizes patient education and relies on costly rehabilitation equipment and modalities. The success of this approach is likely to correlate with the level of specialized postgraduate training completed by the treating therapist (235).

Mobilization
The goal of mobilization is to relieve pain and restore the normal range of motion of the involved soft tissues and joints (236). Mobilizations are often required to facilitate reduction of spinal derangements and can also expedite restoration of motion in spinal dysfunctions. Clinicians should be discouraged from performing spinal mobilization based on palpation techniques used to discern segmental movement. The poor interrater reliability of palpatory assessment techniques in detecting segmental spinal motion has been well documented (237–245). There is evidence to suggest that manual diagnosis techniques may be more reliable if the treating clinician relies on provoking the patient's typical symptoms rather than relying on positional or palpatory findings alone (240,245–249). The therapist should continuously communicate with the patient in order to correlate pain provocation with the direction of mobilizing force and resistance.

Geoffrey Maitland has described four grades of passive movement to restore spinal motion (250,251). Grade I mobilizations are performed using small oscillations at the beginning of range of motion. Grade II mobilizations are large oscillations that are performed during resistance-free range of motion. Grade III mobilizations are large oscillations performed at 50% of joint resistance during range of motion. Grade IV mobilizations are described as small oscillations similarly performed at 50% of joint resistance. A plus and minus system is also incorporated into this system to provide a specific grade of mobilization. Every plus or minus equals 25%; that is, a mobilization graded at 3+ is equivalent to a force of 75% of joint resistance while using large oscillations, and a grade of 3− is equivalent to a force of 25% of joint resistance while using large oscillations.

Mobilization with movement, natural apophyseal glides, and sustained natural apophyseal glides described by Brian Mulligan are other categories of mobilization that are quickly gaining popularity in field of manual rehabilitation therapy (252). Further clinical investigations will help to clarify the utility of manual therapy in the treatment of patients with cervical disorders.

Manipulation
Manipulation is a passive movement that is applied to a joint or structure beyond normal physiologic limits with the intent of improving range of motion or reducing a derangement (253). Some believe that manipulative procedures to the zygapophyseal joints enhance afferent signals from mechanoreceptors to the peripheral and central nervous system (254). This normalization of afferent impulses is hypothesized to reduce muscle guarding, improve muscle tone and local tissue metabolism, and lead to increased range of motion and pain reduction (235). Before performing spinal manipulation, the clinician must rule out cervical instability and vertebral artery insufficiency, as high-velocity movements could place the patient at risk for neurologic compromise or vascular insult (255). Manipulation should also likely be avoided in patients with radicular pathology or significant neural foraminal or central canal compromise. It is also recommended that spinal manipulation only be performed by skilled clinicians with comprehensive understanding of spinal disorders.

Although spinal manipulation appears to lead to short-term benefits and may be effective following an acute injury (256–258), several reviews of the manipulation literature have concluded that there are not enough positive studies of acceptable quality to support spinal manipulation as any more efficacious than other treatments (257,259,260). Similarly, there is no evidence that manipulation provides long-term relief, improves chronic conditions, or alters the natural course of the disorder being treated (253,261,262).

Collars
Soft cervical collars are often prescribed for the patient with acute cervical pain or radiculopathy. Soft collars may serve as a postural cue but do not restrict motion (263,264). If prescribed with the narrow end worn anteriorly, these collars might help to restrict cervical extension in the patient with radicular pain. By introducing relative cervical flexion, such use might prove helpful for the patient with radiculopathy while resting supine and during evening hours. More rigid cervical collars do offer true flexion and extension motion restriction and can be used in the patient with suspected instability (265). In a study of patients prescribed more rigid collars, patient compliance with their use was approximately 70% (266). Although soft cervical collars can improve patient comfort, long-term clinical outcomes
are likely not affected by collar use (267,268). In the absence of significant instability, it has been recommended that collars not be used continually for greater than 72 hours because their use can lead to soft tissue tightening and a delay in range-of-motion restoration (268–270).

**Modalities**

The clinician should limit the application of modalities to the acute injury phase with a primary objective of pain modification. Both superficial heat and ice have the potential to serve as an analgesic and reduce muscle spasm. Neural firing rates have been demonstrated to change in response to changes in temperature (271). Both an elevation and a reduction of muscle temperature resulting from the application of superficial heat and cold have been correlated with decreased gamma motor neuron activity (272,273). These changes in nerve firing rates are thought to contribute to a reduction in muscle spasm (274). Although superficial heat is often applied for comfort measures, application of ice is preferred in the acute phase of injury because of its greater depth of penetration and its analgesic and anti-inflammatory effects (235). Local regular icing can result in vasoconstriction and a reduction in the release of nociceptive and inflammatory mediators such as prostaglandins (275). The use of heat or icing modalities can be repeated several times daily. The decision to use heat or ice can be predominantly determined by which is perceived by the patient as most analgesic.

Ultrasound is probably one of the most commonly over-used physical therapy modalities. A review of 35 randomized controlled trials published between 1975 and 1999 studying the effectiveness of therapeutic ultrasound revealed that there was little evidence that active therapeutic ultrasound was more effective than placebo in treating people with pain arising from a wide range of musculoskeletal injuries or for promoting soft-tissue healing (276). Electric stimulation consisting of both low- and high-rate transcutaneous electrical nerve stimulation and interferential current is often similarly used in an effort to reduce pain and inflammation. Uncontrolled studies have suggested a benefit from electrical stimulation use in patients with cervical radicular pain (277). The use of modalities in the treatment of cervical and radicular disorders should be time limited and should not serve as a substitute for a more active rehabilitation approach.

**Traction**

Traction is often used in patients presenting with cervical radiculopathy. The literature is lacking a prospective, randomized, controlled study comparing traction with other rehabilitation approaches in the treatment of radiculopathy, and the efficacy of traction in improving treatment outcomes has not been demonstrated (278–282). A reduction in the signs and symptoms associated with spinal nerve root impingement has been reported with traction use, particularly if applied shortly after the onset of radicular pain (283). In a study of patients with whiplash symptoms, improved cervical spine range of motion was initially observed but with no change in eventual symptomatic outcome (284). It has been suggested that if a sufficient traction force is applied, the size of the neural foramen may temporarily be increased, reducing pressure on a spinal nerve root and potentially contributing to the healing process (285–287). The clinician is encouraged to apply manual traction to assess the likely therapeutic response to force before introducing mechanical devices. The application of traction might only be considered in those patients unable to reduce radicular symptoms through more independent mechanical exercises.

Maximum posterior elongation of the cervical spine is achieved during traction when the neck and angle of pull are in a position of 24 degrees of flexion (288). Over-the-door traction units are properly used with the patient facing the door with 20 to 30 degrees of neck flexion (288,289). Home traction devices have also been designed to apply cervical distractive forces in the recumbent position, and these might allow for more secure and reproducible patient positioning. Mechanical traction may be administered statically, with a uniform force throughout the treatment session, or intermittently, with the force varying throughout the session. It is generally recommended that static traction be performed if the area being treated is more acutely inflamed, if the patient's symptoms are easily aggravated by motion, or if the patient's symptoms are related to disc protrusion (290). The force of traction should always be kept low during the initial traction session in order to decrease the risk of reactive muscle guarding and to further confirm if mechanical traction is appropriate.

The recommended initial mechanical force for cervical traction is between 8 and 10 lb (271). Fifteen pounds of force will introduce muscle stretching, whereas 25 lb of force is the minimum force required to achieve vertebral separation. Cervical traction has been demonstrated to result in greater vertebral distraction with a 50-lb than with a 30-lb force. When used with a 24-degree angle of pull, no significant difference in vertebral separation is noted after 7, 30, or 60 seconds of force applied (289). With a 30-lb force applied for 7 seconds with alternating rest periods of 5 seconds, maximal vertebral body separation has been observed after 25 minutes of use with a resolution of vertebral body positioning after 20 minutes of traction termination (291). Nerve root function should be evaluated before and after the application of this modality. Traction use is contraindicated in patients with myelopathy, instability, and RA and should be introduced with caution in elderly patients who may be at greater risk for vascular injury (291,292).

**Stabilization**

Currently, there is no available literature regarding the efficacy of cervicothoracic stabilization programs in the treatment of axial pain or radicular disorders, but a recent study of lumbar stabilization prescribed in a diagnosis-specific fashion demonstrated a resultant reduction in pain and functional disability (293). O’Sullivan et al. have demonstrated functional gains from training the transversus abdominus muscle and the implementation of a lumbar spinal stabilization program.
in patients with degenerative lumbar disease. One of the innovators of the stabilization program used in the O’Sullivan study, Gwedon Jull of Australia, has recommended a similar treatment approach for cervical spine disorders consisting of strengthening exercises for the deep neck flexors and scapular stabilizers. To isolate the deep cervical flexors, investigators have suggested the use of a pressure-sensor device to provide visual biofeedback to the patient as controlled cervical flexion from neutral is performed (295).

The combined goal of stabilization is to introduce sound postural strategies, improve the strength of the regional musculature, and restore pain-free range of motion. Cervicothoracic stabilization incorporates strengthening of both the neck and the upper-quarter musculature. Special emphasis should be placed on training the scapular retractors and depressors. While performing stabilization exercises, the patient is instructed to maintain a neutral spine position. This can be reinforced through visual feedback with a mirror or verbal feedback from the treating therapist. Optimal stabilization training of the cervicothoracic region is achieved by also incorporating the lumbar spine and lower extremities (296). The lumbar spine and lower-extremity musculature provide the essential foundation for the cervicothoracic region and should not be ignored during a comprehensive stabilization program.

**Work Ergonomics**

Instruction in proper positioning at the work station represents an important component in the rehabilitation of patients with cervical disorders. Regardless of the therapeutic exercise program employed or the patient’s diligence with home exercise efforts, sustained spinal posturing out of the neutral zone can lead to further aggravation of the affected tissues and a perpetuation of symptoms. The sedentary worker spending most of the day seated at a desk is more likely to encounter proper positioning challenges. In general, the patient should be instructed to avoid, whenever possible, those positions that typically aggravate his or her cervical or limb pain. The recommended positioning at the desk or computer station is usually prescribed with approximately 90 degrees of hip, knee, and elbow flexion (295) (Fig. 32-17). Proper distance from the computer monitor is an arm’s length measurement, which is unique to each individual. The upper third of the computer screen should be situated at eye level. Prolonged phone use can also serve as a cervical stressor as ipsilateral neck flexion is typically performed while using the telephone. Headset use can help eliminate this provocative position (297). Individuals that require bifocal eyeglasses should be instructed to position reading materials in such a fashion as to avoid sustained cervical extension postures (298). In addition, the wrist or elbow, and preferably both, should be maintained in a supported position. This support contributes to total neutral positioning of the body and ultimately can lead to decreased stresses on peripheral nerves passing through the elbow and carpal tunnel.

**Massage**

In cases in which debilitating muscle guarding or soft-tissue pain persists despite the application of mechanical exercises or manual techniques, soft-tissue massage may prove helpful. A study evaluating the benefits from soft-tissue massage in patients with chronic tension headaches suggests a resultant reduction in pain scores and frequency of neck pain events (299). A more recent systematic review of massage employed for lumbar pain identified four randomized clinical trials in which massage was used as the sole treatment modality. This review concluded that massage appears to have some efficacy as a therapy for low-back pain. The author of this review emphasized the methodological limitations in the massage literature and the need for more well-designed clinical trials. The use of massage should likely be limited to the more acute injury phase, with an emphasis placed on graduating the patient to a more independent and mechanical approach to addressing the underlying symptom generator.

**The Physical Therapy Prescription**

The physical therapy prescription should mark the initiation of an ongoing dialogue between the physician and the therapist. Its purpose is to introduce the patient to the treating therapist and provide the therapist with pertinent physical and radiologic findings. As the relationship between the physician and the therapist matures, physical therapy prescriptions do not need to be extremely detailed. The prescribing physician should, whenever possible, avoid the assignment of more generic diagnoses, that is “sprain/strain,” and treatment recommendations. Although the astute spine therapist, through a comprehensive initial evaluation, will often be able
to appropriately design the rehabilitation program, information provided by the physician can be vital in selecting treatment strategies, maintaining precautions, and optimizing outcomes.

REFERENCES

3. Garfin SR. Cervical degenerative disorders: etiology, presentation, and imaging studies. \( ADOS \) Int.


CHAPTER 33

INTRODUCTION

The approach to low back pain (LBP) has changed dramatically in the past decades, while medicine has changed also (1–4). Formerly, according to a classical disease model, all LBP classifications considered a pathoanatomic basis and consequently proposed treatments; but the increasing burden on society in terms of costs and disability (5), presumably due to inappropriate treatment and medical approach, lead to a revolution in the understanding of the problem (4,6). Accordingly, LBP is now classified as secondary (<10% of cases) and primary, or idiopathic, or simply LBP; the latter is then divided according to the localization of symptoms (LBP and sciatica) and the duration of pain (acute, subacute, and chronic) (3,7–11). Generally speaking, in this model LBP is recognized as a biopsychosocial syndrome (4,12). Consequently, the importance of the discipline of Physical and Rehabilitation Medicine (PRM), as one with particular focus on biopsychosocial approaches according to the International Classification of Functioning (ICF) (13–20), has greatly increased in this field (21,22). As is always the case in medicine, these advancements require time to reach the clinical everyday world (10,23–27). Today the clinical reality is a mix of different approaches, usually driven by what each single specialty, doctor, and/or allied professional knows and offers in terms of treatments, more than by a coherent and systematic evidence-based approach (28,29). Considering the world of PRM and its inevitably multi-professional reality (14,30), combined with the almost complete outpatient reality of LBP treatment, the confusion seems even higher than in other specialties (31–35). Moreover, considering the widespread recognition of LBP as a biopsychosocial syndrome, and consequently the crucial role PRM should have in its treatment, this confusion must be cleared away. That is the main aim of this chapter.

This chapter is fully evidence-based, but it is also totally focused on what to do in the clinical everyday world once the evidence is known: in reality it is an evidence-based clinical practice tool. Therefore, we will start from some clinical-scientific premises, including assessment and outcome criteria, the actual scientific evidence on treatments of PRM interest, and also the diagnostic-therapeutic pathways (flowcharts) (9) to be followed. In fact, another unavoidable premise is that in LBP rehabilitation, as well as in all other PRM approaches to patients who are diagnosed with a specific disease, a correct disease diagnosis is the first step and must be achieved. We cannot rehabilitate the patient if we do not start with diagnosis, and only after that can we look for a mandatory biopsychosocial picture of the patient, possibly inside the mainframe of ICF (17,18,36–38). Finally, we will have an evidence-based description of the single clinical LBP pictures of PRM interest, concluding with the rehabilitation approach to each of them.

The last important premise is one of terminology: rehabilitation is neither a pain-orientated (PO) nor a conservative treatment. From our perspective, the term “conservative treatment” should be abandoned because it is based on the past, when the only need was to define whether a LBP was surgical or not (21,22). The term “conservative” is not the positive affirmation of what we have to be in regard to our patients. Instead it is only the negation of surgery, as if this were the gold standard. In the so-called field of conservative treatment there are abundant treatments that focus on pain (“PO treatments”); these have been traditionally used for years, and continue to be used, sometimes according to some evidence. However, with the increased scientific understanding of the limits of this PO approach (4,5,39), more complex rehabilitation approaches focusing on increasing function, on recovering activities and participation, and on rising quality of life have emerged and developed: we will call them functioning-orientated (FO) approaches (40–42). These FO approaches are based on WHO’s understanding of functioning (13), representing the comprehensive view taken by PRM (14–20,30). In most patients this is far more appropriate, and it should be the main concern of PRM specialists, who have the best conceptual tools with which to appropriately rehabilitate LBP patients (13,14,22,30,37,38). This does not lead us to ignore PO treatments, which in some or many cases must be applied, but the differentiation is crucial to achieving a better understanding of our everyday behaviors, including their advantages and limitations.
ASSESSMENT, DECISION-MAKING PROCESS, AND REHABILITATION TREATMENT TOOLS

Assessment and Outcome Criteria
An appropriate diagnostic workup is the cornerstone of an effective treatment. It is founded on taking the most accurate medical history and conducting a thorough physical examination (8–10), both of which are useful in forging a good doctor-patient relationship. The first objective of assessing the patient who is experiencing back pain is to identify the patient who presents serious spinal pathology. It is possible to establish a well-defined pathology in only about 15% of patients who present LBP (43). Among patients suffering from LBP and sciatica for a period of approximately 4 weeks, Deyo et al., as well as Udén, were able to identify the following: herniated disc (4% to 5%); lumbar spinal stenosis (LSS) (4% to 5%); spinal fractures (4%); metastatic tumor or osteomyelitis (1%); and visceral pathologies such as aortic aneurysms, kidney, or gynecological disorders (<1%).

According to Waddell (44), when dealing with a patient suffering from LBP, the first step is to identify any possible serious spinal pathologies (primitive or metastatic neoplasia, infections or inflammatory diseases such as ankylosing spondylitis, osteoporotic fractures, cauda equina syndrome, or any of various neurologic diseases). An assessment of the pain is then conducted in order to determine whether it originates in the nerve root (sciatica from a prolapsed disc, spinal stenosis, or surgical scarring):

• On one side of one of the lower limbs, generally radiating below the knee, down to the ankle and the foot
• Very often more intense than LBP
• Hypoesthesia or paraesthesia with the same distribution
• Positive signs of radicular irritation
• Motor, sensory, or reflex changes related to a single nerve root

Assessing LBP means evaluating pain symptoms, with all the ensuing uncertainties that this entails. It is especially difficult to objectively assess the severity of the pain due to the complex neurophysiologic sensory mechanisms and the psychosomatic assimilation that occurs in the experience of pain, and which cannot be differentiated by the person experiencing it (45). Regarding duration of the pain, the updated universally accepted classification is (3,5,8,11,46):

• Acute LBP (ALBP): up to 4 weeks
• Chronic LBP (CLBP): over 6 months
• Subacute LBP (SALBP): from 2 to 6 months

A recent study has shown the presumable transition between ALBP and SALBP in 14 days (47), while others propose 90 days as the starting point of CLBP (9). As the symptoms of pain persist, there is a shift from a pathology of a physiological nature with biological ramifications to one characterized by psychosocial implications. The assessment at this stage no longer simply concerns a disorder but also disability and diminished quality of life.

Scientific Evidence on Assessment
Generally in the medical field, the basic requirements for any diagnostic test are its accuracy, safety, and reproducibility (48–50). Accuracy comprises the twin parameters of specificity and sensitivity: specificity refers to the ability of the test to be positive only when a pathology is present, while sensitivity is the characteristic of being positive whenever the pathology is present. Moreover, in order to ascertain the accuracy of a diagnostic test it is necessary to have a gold standard against which to measure the sensitivity and specificity. For spinal pathologies, in some cases the gold standard is the presence or absence of defined tissue pathology confirmed through a surgical procedure, while in other instances it is the confirmation or relief of pain. In all these cases, given the difficulty in obtaining an objective assessment of pain, the true accuracy of any test that provokes or alleviates pain carries an intrinsic uncertainty (51).

Various clinical guidelines (CGs) (8,9,52) stress the importance of taking a detailed medical history and a comprehensive physical examination of the patient, and they emphasize that during the first visit in particular, the physicians should dedicate their full attention to the patient and spend the appropriate amount of time needed to earn the individual’s trust. Indeed, evidence shows that an appropriate clinical approach can only have a beneficial impact on the pain and on allaying the patient’s fears (53).

Due to its biopsychosocial characteristics, an important future advancement in LBP assessment could be represented in the near future by the application of the ICF (15–20) and mainly its core set (36–38,54–56).

Red and Yellow Flags
To improve the accuracy of excluding serious spinal pathologies, as well as to reassure the patient, almost all existing CGs (8,9,52) have recommended that a series of risk factors called “red flags” (Table 33-1) be assessed as the patient’s medical history is taken down. The presence of one or more of these flags would determine whether there is a need for a targeted diagnostic confirmation and subsequent consultation with a specialist. Deyo has shown that red flags have a high degree of sensitivity for excluding a cancerous pathology among 2,000 patients with LBP (57).

Other CGs (58) have identified other risk factors, being indicators of the possible evolution in chronicity of pain, that have been called “yellow flags” (58,59), whose presence, if found in ALBP or SALBP, indicate the need to take into account psychosocial factors and apply early therapeutic strategies in a cognitive behavioral approach (see Table 33-1).

Medical History and Physical Examination
The meta-analysis of van den Hoogen et al. (60) does not allow the drawing of definitive conclusions on the accuracy of medical history-taking, the physical examination, and the assessment of a Visual Analogue Scale (VAS), notably due to inadequacies in the methodologies of the primary studies, particularly regarding referral bias. However, there is agreement
### TABLE 33.1 In the Diagnostic Process of LBP Patients It Is Important Since the Beginning to Search for Red Flags (Risk Factors for Secondary LBP Due to Important Pathologies): In Subacute Cases It Is Crucial to Search for Yellow Flags (Risk Factors for Chronicization of LBP) (9)

**Red flags (Risk Factors for Secondary LBP Due to Important Pathologies)**
- Back pain in children <18 y with considerable pain or onset >55 y
- History of violent trauma
- Mild trauma in an aged patient
- Constant progressive pain at night
- History of cancer
- Systemic steroids
- Drug abuse, human immunodeficiency virus infection
- Weight loss
- Systemic illness
- Persisting severe restriction of motion
- Intense pain or minimal motion
- Structural deformity
- Difficulty with micturition
- Loss of anal sphincter tone or fecal incontinence; saddle anesthesia
- Widespread progressive motor weakness or gait disturbance
- Inflammatory disorders (ankylosing spondylitis) suspected
  - Gradual onset <40 y
  - Marked morning stiffness
  - Persisting limitation of motion
  - Peripheral joint involvement
  - Iritis, skin rushes, colitis, urethral discharge
- Family history

**Yellow flags (Risk Factors for Chronicization of Acute and SALBP)**

<table>
<thead>
<tr>
<th>Category</th>
<th>Factors</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Personal</strong></td>
<td>Age (U form correlation)</td>
</tr>
<tr>
<td></td>
<td>Female gender</td>
</tr>
<tr>
<td></td>
<td>Minor ethnicity</td>
</tr>
<tr>
<td></td>
<td>Low income</td>
</tr>
<tr>
<td></td>
<td>Low education</td>
</tr>
<tr>
<td><strong>Medical</strong></td>
<td>High BMI</td>
</tr>
<tr>
<td></td>
<td>Previous surgery</td>
</tr>
<tr>
<td></td>
<td>Impairment</td>
</tr>
<tr>
<td></td>
<td>Neurological deficit</td>
</tr>
<tr>
<td></td>
<td>Radicular impingement (SLR, Wassermann tests)</td>
</tr>
<tr>
<td><strong>Pain related</strong></td>
<td>Duration</td>
</tr>
<tr>
<td></td>
<td>Intensity</td>
</tr>
<tr>
<td></td>
<td>Leg pain</td>
</tr>
<tr>
<td></td>
<td>Pain in lateral flexion and/or in flexion-extension</td>
</tr>
<tr>
<td></td>
<td>Difficulties in sitting</td>
</tr>
<tr>
<td><strong>Impairment disability related</strong></td>
<td>High referred impairment</td>
</tr>
<tr>
<td></td>
<td>High functional limitation at 4 wk</td>
</tr>
<tr>
<td></td>
<td>High disability (Roland-Morris, Oswestry, Sickness Impact Profile)</td>
</tr>
<tr>
<td></td>
<td>Perceived risk of not recovering</td>
</tr>
<tr>
<td><strong>Psychosocial</strong></td>
<td>Not appropriate signs and symptoms</td>
</tr>
<tr>
<td></td>
<td>Avoidance behavior</td>
</tr>
<tr>
<td></td>
<td>Psychological burden</td>
</tr>
<tr>
<td></td>
<td>Vital energy reduction</td>
</tr>
<tr>
<td></td>
<td>Reduced emotional confronting capacity</td>
</tr>
<tr>
<td></td>
<td>Social isolation</td>
</tr>
<tr>
<td></td>
<td>Depression (SCL-90, Zung, Back Depression Inventory)</td>
</tr>
<tr>
<td></td>
<td>Somatization (SCL-90)</td>
</tr>
<tr>
<td></td>
<td>Reduced coping strategies (CPCT)</td>
</tr>
</tbody>
</table>
TABLE 33.1

| Red Flags (Risk Factors for Secondary LBP Due to Important Pathologies): In Subacute Cases It Is Crucial to Search for Yellow Flags (Risk Factors for Chronicization of LBP) (9) (Continued) |
|---|---|
| **Work related** | High requests
Reduced control on own work
Monotony
Low satisfaction
Treatment before retiring from work
Disability compensation
Heat and cold therapies
Physiotherapy
Back school |

on the validity of medical history and physical examination to gather the red flags to detect conditions such as cancer, cauda equina syndrome, spinal stenosis, ankylosing spondylitis, or spinal deterioration due to osteoporosis.

Several studies have focused on how a physical examination is performed in regard to the range of motion of the lumbar spine, muscle strength, and provoking or relieving pain symptoms with specific movements. The McKenzie technique for determining which movements can cause the centralization of pain and what courses to take has not found unanimity in the literature for indicating prognosis. While positive findings have been reported in Long and Donelson (61,62), other authors have yet to establish a clear clinical benefit (63).

A study by Simmonds et al. (64) has established the reliability, validity, and good clinical use of nine physical evaluations comparing asymptomatic subjects and LBP patients. Also, a modest correlation with the degree of disability was detected through a questionnaire. No evidence of clinical benefit has been obtained in the analysis of muscle strength carried out by various machines (65). In fact, a good correlation of testing and retesting has shown that the Biering-Sörensen test is simpler to conduct and is of greater benefit in clinical practice (66). Regarding muscle fatigue, Taimela et al. (67) have shown that LBP sufferers first get tired, whereupon the fatigue reduces the quality of movement, and consequently the muscle fatigue lessens the kinesthesia. Muscle fatigue, therefore, is a risk factor.

Clinical trials, as well as various national CGs, have taught us that a fundamental point of the evaluation is physical contact with the patient: palpation of the vertebral structure provoking pain and the direct assessment of regional and segment motion are an integral part of a complete physical examination. The analysis of literature in this area by Seffinger et al. (68) shows that the most reliable tests are those that provoke pain, while less reliable diagnostic testing is that which is done by palpating the soft tissues. Regional motion testing is also more reliable than segmental motion testing, and intra-examiner reliability is greater than inter-examiner reliability.

If, on the one hand, there is a lack of definition of the specific structural causes in a great number of LBP sufferers, on the other hand the practice tends to identify specific subgroups of subjects affected by LBP. Many consider lumbar instability to be one such subgroup. In this regard, a recent study (69) has examined predictive clinical indicators of instability, the gold standard being x-rays in the flexed and extended positions. A positive correlation was observed in the concurrent presence of hypermobility in the intervertebral motion test and the range of the lumbar flexion greater than 53 degrees (LR 12.8).

**LBP and Leg Pain**

Regarding the pain that radiates to a lower limb below the knee and the probable correlation with the presence of a herniated disc, scientific evidence (70,71) generally refers to secondary or tertiary health care settings, the gold standard being the presence of a herniated disc confirmed by surgery. Therefore, it is most likely that only the more serious cases are being reported instead of cases of patients seeking primary care. Based on these studies, the distribution of pain seems to be the only sensitive symptom, while paresis and sensory deficits seem neither sensitive nor specific. Among the indicators, the depression of reflexes is not sensitive and straight-leg raising (SLR) seems to be an indicator of sensitivity (but of low specificity), while the crossed SLR has less sensitivity but a higher specificity. The positive presence of both suggests a better clinical indication of a herniated disc.

Concerning a prospective study of 105 patients (72), the strongest indicators of a herniated disc among the 25 clinical variables (the gold standard being the computer tomography [CT] and/or magnetic resonance imaging [MRI] scan) were the range of lumbar motion in forward flexion and the distribution of pain in the lumbar extension while in the standing position. Patients with herniated discs had significantly less mobility during forward flexion but a higher distribution of pain to the leg during a lumbar extension (Table 33-2).

**Diagnostic Imaging**

The choice of imaging examination for diagnostic purposes should only be based on an evaluation of the symptoms, medical history-taking, and physical examination. The purpose of diagnostic imaging is to verify the suspicion of a serious
pathology due to red flags, or a herniated disc due to symptoms ascribed to nerve-root pain. In the case of LBP or simple sciatica, and in the absence of red flags, diagnostic imaging is not required within 30 days of the onset of symptoms (3,8,9).

Regarding x-ray examinations, all the current CGs agree on refuting their diagnostic or therapeutic value in the absence of red flags and discourage any routine use. van Tulder et al. (73) have noted that the only x-ray anomaly that could be linked to LBP is disc deterioration (odds ratio 1.2,3.39). However, due to inadequate methodologies in primary studies this conclusion cannot be considered definitive. Consequently, the authors claim that “no firm evidence exists for the presence or the absence of an association between x-ray findings and nonspecific LBP” (Table 33.3).

The noninvasiveness and characteristics unique to MRI in satisfactorily evaluating both soft and bone tissues allow it to be a more comprehensive tool for the diagnostic imaging of an LBP patient. That being said, there is no evidence that its use provides any real advantage in the treatment of simple LBP (74,75). Many studies have shown that the presence of disc anomalies (bulges, swelling, hernia) in the vertebral canal, the foramen, and the vertebral structures are found in asymptomatic subjects. Consequently, Roland and van Tulder (76) recommended, somewhat provocatively, that radiologists add the following to the reports along with their referrals: “This finding may be unrelated to patient symptoms because it is often seen in asymptomatic subjects.” The MRI is a highly sensitive exam, and hence there is a risk of obtaining many false positives. We must bear this in mind in clinical practice so that costs can be kept down and unnecessary surgical procedures avoided.

### Psychosocial Factors, Disability, and Quality of Life

The assessment of a patient with LBP must not be limited to a physical examination but should also include an analysis of psychosocial factors that play a crucial role in the chronicity of pain, delays in returning to work, and the success of the treatment. Evidence shows that some conditions—such as low level of job satisfaction, poor work motivation, disability compensation, dissatisfaction with previous treatment—are risk factors for chronicity and/or the relapse of LBP. Regarding psychological factors as chronicity/disability predictors, a recent review (77) has examined the role of psychological discomfort and depression (Zung scale), and somatization, while less evidence was attributed to cognitive factors such as catastrophism. However, although the symptoms of mental stress in asymptomatic subjects can be predictors (78), the main indicator for LBP is a positive medical history (79).

Persistent pain always entails a certain degree of disability. The measurement for disability in LBP can be gauged through means of dedicated and scientifically validated questionnaires such as the Oswestry LBP Disability Questionnaire and the Roland-Morris Questionnaire (80–83). A recent study (84) has correlated pain (VAS), disability (Oswestry Questionnaire), and the quality of life (EuroQoL) in LBP patients. The data have shown a weak but significant correlation among the three elements. Thus, it appears that the persistence of pain and disability worsen the quality of life to an appreciable extent. Clearly, the impact of pain and disability on decreasing the quality of life depends more on their duration than their intensity. Moreover, even a clinically significant variation in the degree of pain can bring about nearly imperceptible changes in disability and quality of life. We must, therefore, adopt the concept of the minimal clinically important difference (85,86), that is, the slightest change in an evaluation scale that will be perceived by the patient as an improvement in his or her condition. It is important to stress the benefit of such a concept as a better outcome criterion because it is aimed at the patient from a functional and psychosocial perspective. However, its a well-known fact that today there is also the need

### Table 33.2

<table>
<thead>
<tr>
<th>Finding</th>
<th>Prediction Strength</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crossed SLR reproducing pain in the symptomatic leg</td>
<td>+++</td>
</tr>
<tr>
<td>SLR &lt;60 degrees (SLR reproduces leg pain)</td>
<td>+</td>
</tr>
<tr>
<td>Ankle dorsiflexion weakness</td>
<td>+</td>
</tr>
<tr>
<td>Great toe weakness</td>
<td>+</td>
</tr>
<tr>
<td>Impaired ankle reflex</td>
<td>+</td>
</tr>
<tr>
<td>Sensory loss, pins and needles, paresthesia</td>
<td>+</td>
</tr>
<tr>
<td>Patellar reflex weakness</td>
<td>+</td>
</tr>
<tr>
<td>Ankle reflex weakness</td>
<td>+</td>
</tr>
<tr>
<td>Severe radicular pain</td>
<td>++</td>
</tr>
<tr>
<td>Pain causing awakening at night</td>
<td>++</td>
</tr>
<tr>
<td>Severe lumbar motion restriction</td>
<td>++</td>
</tr>
<tr>
<td>Loss of lordosis and/or sciatic scoliosis</td>
<td>++</td>
</tr>
<tr>
<td>Unilateral leg pain worse than back pain</td>
<td>++</td>
</tr>
<tr>
<td>Radiation into foot</td>
<td>+</td>
</tr>
<tr>
<td>Pain drawing (exact dermatome depicted)</td>
<td>(+)</td>
</tr>
</tbody>
</table>

### Table 33.3

<table>
<thead>
<tr>
<th>Radiographic Finding</th>
<th>No. of Studies</th>
<th>Odds Ratio</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disc degeneration</td>
<td>12</td>
<td>1.2–3.3</td>
<td>Moderately positive</td>
</tr>
<tr>
<td>Spondylosis</td>
<td>3</td>
<td>1.2–2.0</td>
<td>Negative</td>
</tr>
<tr>
<td>Spondylolisthesis and</td>
<td>6</td>
<td>0.82–1</td>
<td>Negative</td>
</tr>
<tr>
<td>Spina bifida</td>
<td>2</td>
<td>0.5–0.6</td>
<td>Negative</td>
</tr>
<tr>
<td>Transitional vertebrae</td>
<td>3</td>
<td>0.5–0.8</td>
<td>Negative</td>
</tr>
<tr>
<td>Scheuermann disease</td>
<td>2</td>
<td>0.8–3.6</td>
<td>Unclear</td>
</tr>
</tbody>
</table>
to evaluate a treatment’s efficiency in terms of managing the costs of health care.

This section concludes with a word about younger patients: The findings of more recent studies show there is an increased incidence of LBP among children and adolescents. Among the defined risk factors in a school population of 10,000 subjects are the excessive weight of backpacks, the sitting level of students compared to that of the teacher, and the heights of chairs and desks (87). Starting with the premise that a better predictor of LBP is precisely a positive medical history and the awareness that children and adolescents exhibit significant differences regarding medical history, physical examination, and diagnosis compared to the adult population (88), it would appear altogether appropriate that new studies be conducted so as to bring this subgroup to the fore and thereby formulate better treatment and avoid the relapse of LBP in adulthood.

Diagnostic Therapeutic Flowcharts

Subsequent to the pioneering experience, not only for LBP but also for medicine in general, of the Quebec Task Force down in 1987 (89), which represented the first systematic effort to produce a CG in terms of a summary of the existing evidence on LBP treatment, this field has been one of the most prolific for CGs production (10,90). In this context, with the objective of painting in a simple way the everyday clinical behaviors, those published in 2006 in *Europa Medicophysica* (9) (now the European Journal of Physical and Rehabilitation Medicine) (91) appear very interesting. The “Diagnostic Therapeutic Flowcharts for LBP Patients” (DTP) have been developed as real flowcharts (Figs. 33-1, 33-2, 33-5, 33-7, 33-9, 33-11, and 33-13), in order to provide a complete idea of what should be done, and to cover, through expert multidisciplinary consensus, the multiple gray areas in which the actual evidence does not offer answers. Moreover, these DTP have other unique characteristics in terms of PRM (21) that are reported below.

In Figures 33-2, 33-5, 33-7, 33-9, 33-11, and 33-13, the most interesting flowcharts of the DTP are given to the reader, while in the open access version of the European Journal of Physical and Rehabilitation Medicine (EJPRM) (www.ejprm.org) they can be viewed for further details.

The DTP have been developed through a systematic search of the literature and a multidisciplinary consensus of all Italian scientific societies engaged in this field. The classical distinction between ALBP, SALBP, and CLBP is at the base of the DTP. Even though a recurrent form of LBP was hypothesized at the start of the DTP project, the absence of literature and the lack of any evidence on the possibility of differentiating a recurrent episode from a single first episode of LBP made it necessary to eliminate this classification item (9).

One of the most interesting points is represented in the flowchart in Figure 33-2. In fact, while preparing the DTP it was discussed how to grade the importance of the different rehabilitation approaches possible for SALBP and CLBP patients. It was agreed that, from the viewpoint of rehabilitation and also prevention, the most important patients were the SALBP ones, where it was still possible to try avoiding chronicization, which is the most awful end of the story, because it is extremely rare that a patient will exit from CLBP. In the meantime, which CLBP patient would deserve the highest attention so to reduce the burden of the biopsychosocial situation for them and for society? According to the literature, such a choice has been defined using the level of disability as a way to discriminate among patients (9,21). This is a totally new conceptual development, but it is consistent with the actual knowledge of CLBP (22). Moreover, the actual disease-specific disability questionnaires have already defined cut-off levels in order to distinguish high-disability and low-disability patients (48,81). This conceptual step forward provides a means by which to classify the various possible treatments (see Fig. 33-2).

Another important point of the DTP has been to recognize the importance not only of the therapy proposed but also

![Diagnostic flow-chart for low back pain and sciatica](image-url)
Evidence

<table>
<thead>
<tr>
<th>Evidence</th>
<th>Flow-chart</th>
</tr>
</thead>
<tbody>
<tr>
<td>Over 4 weeks of pain</td>
<td></td>
</tr>
<tr>
<td>A* Has been performed any evaluation during the acute period?</td>
<td>Yes</td>
</tr>
<tr>
<td>A Yes</td>
<td></td>
</tr>
<tr>
<td>A* Is there any red flag?</td>
<td>Yes</td>
</tr>
<tr>
<td>A No</td>
<td></td>
</tr>
<tr>
<td>A* Perform lumbar X-rays in standing</td>
<td></td>
</tr>
<tr>
<td>A Arthrosis, congenital lumbo-sacral defect or Schmorl herniation? (1)</td>
<td>Yes</td>
</tr>
<tr>
<td>A No</td>
<td></td>
</tr>
<tr>
<td>A* Is there a spondylolisthesis?</td>
<td>No</td>
</tr>
<tr>
<td>A* Yes</td>
<td></td>
</tr>
<tr>
<td>C Are there symptoms or signs suggesting possible instability? (2)</td>
<td>No</td>
</tr>
<tr>
<td>C Yes</td>
<td></td>
</tr>
<tr>
<td>B Is there a scoliosis over 30° Cobb? (3)</td>
<td>No</td>
</tr>
<tr>
<td>B Yes</td>
<td></td>
</tr>
<tr>
<td>A* Is there any other specific pathology?</td>
<td>No</td>
</tr>
<tr>
<td>A* Yes</td>
<td></td>
</tr>
<tr>
<td>A* Is there any leg pain?</td>
<td></td>
</tr>
<tr>
<td>B No</td>
<td>Over 3 months of pain?</td>
</tr>
<tr>
<td>B Yes</td>
<td></td>
</tr>
<tr>
<td>A* Has been performed a sub-acute evaluation?</td>
<td>No</td>
</tr>
<tr>
<td>A* Yes</td>
<td></td>
</tr>
<tr>
<td>A* Are there neurological signs or both leg symptoms?</td>
<td>No</td>
</tr>
<tr>
<td>A* Yes</td>
<td></td>
</tr>
<tr>
<td>A* Over 3 months of pain?</td>
<td>No</td>
</tr>
<tr>
<td>A* Yes</td>
<td></td>
</tr>
<tr>
<td>A Is there any yellow flag? (4)</td>
<td>No</td>
</tr>
<tr>
<td>A Yes</td>
<td></td>
</tr>
<tr>
<td>B Disability in ADL and/or work (high: Roland Morris &gt; 14 - Oswestry &gt; 25) (low: Roland Morris&lt;=14; Oswestry&lt;=25)</td>
<td></td>
</tr>
<tr>
<td>B High</td>
<td>Go to chronic high-disability LBP therapy</td>
</tr>
<tr>
<td>B Low</td>
<td>Go to chronic low-disability LBP therapy</td>
</tr>
</tbody>
</table>

Notes
1. Arthrosis: discopathy, osteophitosis, reduction of discal space, and/or vertebral endplates thickening. Usually these are radiological diagnosis without any relevance.
2. Neuro-muscular spinal instability (different from the osteo-ligamentous one) has not been defined in the literature. It’s possible to consider the following criteria:
   - sharp and brief acute pain following sudden position changes and/or efforts
   - pain during stabilization tests: e.g. sudden release after isometric contraction of hip muscles (flexion, adduction, abduction), trunk destabilization
3. Scoliosis over 30° can progress in adulthood and needs specialist control.
4. See Table 1.

FIGURE 33-2. Subacute and CLBP diagnostic flowchart (9).
the other elements constituting a complete approach to LBP patients. This starts with the biopsychosocial theory, which is the basis of a modern approach to LBP patients (3,4). It clearly contradicts the usual disease-oriented approach, where only therapy could be considered enough. This is correct and typical of CGs for specific diseases, but it is incorrectly maintained for analogy in the field of LBP. Consequently, in the DTP all therapeutic schemes (Tables 33-4 through 33-8), including those on secondary LBP, consider counseling, work and activities of daily living (ADL) interventions, physical activity counseling, PO therapy, and specific FO rehabilitation. Again it emphasizes the importance of all these elements in approaching what is not a “simple” disease but is instead a biopsychosocial syndrome.

Secondary LBP is usually not included in CGs (10), eventually with the exception of disk herniation (DH): In the DTP there are also specific flowcharts for spinal stenosis, spondylolisthesis and spinal instability, adult scoliosis (AS), and paediatric scoliosis. Excluding the last one, where a specific rheumatologic approach is proposed, in the others the main ideas were as follows:

- Identify specific cut-offs for the eventual proposal of surgery;
- A part from rare medical urgencies, postpone in any case surgery after an appropriate complete rehabilitation process;
- Verify the absence of contraindications to surgery;
- Fully inform the patient regarding the advantages and disadvantages so as to allow an individual, informed choice.

For each secondary LBP clinical picture, specific rehabilitation is listed according to the actual literature. For example, the best types of exercises are proposed even if there is no definitive proof in the literature today, but this is obviously discussed according to the strength of evidence.

The different flowcharts (modified) will be reported throughout this text when appropriate (see Figs. 33-2, 33-5, 33-7, 33-9, 33-11, and 33-13).

Rehabilitation Approach: Treatment Tools

Actual Evidence in the Literature

Many therapies are proposed in LBP management in everyday clinical practice. For some therapies there are no data at all or limited evidence at best. Occasionally, the only rationale of prescription is traditional, while for others we can rely on more consistent findings. We distinguish the treatment approach as PO treatments (oral drugs and injections, physical therapies, manipulations), FO approaches, and educational interventions (9). This distinction could now appear to some readers less than totally justified: in fact, most of the PO approaches propose themselves as “normalizers” of physiology and/or “aetiological” (34,61,92–96), but to date there are no scientific proofs of these hypotheses. Moreover, according to the actual literature we will consider together ALBP and SALBP because there are not enough papers to make a real distinction between the two.

Acute and Subacute LBP

Pain Management

Nonsteroidal anti-inflammatory drugs (NSAIDs) have been shown to be more effective than placebo concerning a general effect on ALBP. However, no single NSAID molecule overcame the others in terms of effectiveness. There is moderate evidence that NSAIDs are not more effective than paracetamol for ALBP, and paracetamol had fewer side effects. Traditional NSAIDs seem to have more gastric side-effects compared to COX-2, but there are no differences in terms of efficacy. In any case, the effect ranges are small. Benzodiazepines, as well as central muscle relaxants other than benzodiazepines, were shown to be more effective than placebo on ALBP (97).

Manipulations were shown to be more effective than placebo on pain in the short-term (within 6 weeks) but were not more effective on function (98). There is moderate evidence that lumbar supports are not more effective than no intervention or training in preventing LBP, and there is conflicting evidence about whether they are effective as supplements to other preventive interventions. It remains unclear whether lumbar supports are more effective than no or other interventions for the treatment of LBP (99).

No relevant data are available concerning steroid injections, oral colchicine, acupuncture, electromiographic biofeedback, lumbar supports, massage, transcutaneous electrical nerve stimulation (TENS), traction, thermic-effect based therapies (ultrasound, ice, heat), insoles (there is strong evidence that insoles are not effective for the prevention of LBP) as LBP treatments (98,100–103). Some herbal medicines (harpagophytum procumbens, salix alba, and capsicum frutescens) seem to reduce pain more than placebo (104).

Educational Interventions

Advice to stay active is effective and sufficient for the long-term improvement of function in ALBP (105,106). The indication to rest in bed is less effective than the indication to stay active (107). There is strong evidence that an individual 2.5-hour session of oral education is more effective on short- and long-term return-to-work than no intervention. Educational interventions that were less intensive were not more effective than no intervention (108). There are no randomized controlled trials (RCTs) on ALBP using back schools (109).

Function-Oriented Rehabilitation Approaches

There are no RCTs on ALBP considering the multidisciplinary rehabilitation treatments. Exercises showed no effect in the first 2 weeks of LBP, but they were shown to be effective in SALBP in the occupational setting (110). There is some evidence that the McKenzie method is more effective than passive therapy for ALBP. However, the magnitude of the difference suggests the absence of clinically significant effects (63).
Chronic LBP

Function-Oriented Rehabilitation Approaches

Strong evidence exists to the effect that intensive multidisciplinary biopsychosocial rehabilitation with a functional restoration approach improves function as compared to inpatient or outpatient non-multidisciplinary treatments. There is reasonable evidence that this approach improves pain as compared to outpatient non-multidisciplinary rehabilitation or usual care. There is contradictory evidence regarding vocational outcomes of intensive multidisciplinary biopsychosocial intervention (42).

Combined respondent-cognitive therapy and progressive relaxation therapy are more effective than waiting-list controls on short-term pain relief. However, it is unknown whether these results can be sustained in the long term. No significant difference has been detected between behavioral treatment and exercise therapy (111). A short-period of cognitive behavioral treatment and exercise therapy has a modest effect on pain and disability at one year. This effect can be influenced by preference toward treatment (112).

Exercise therapy appears to be slightly effective at decreasing pain and improving function in adults with CLBP, particularly in health care populations (110). There is limited evidence regarding the use of McKenzie method in CLBP. The effectiveness of classification-based McKenzie is yet to be established (63).

Educational Interventions

The effectiveness of individual education is still unclear (108). However, there is moderate evidence suggesting that back schools, in an occupational setting, reduce pain and improve function and return-to-work status in the short term and intermediate term as compared to exercises, manipulation, myofascial therapy, advice, placebo, or waiting-list controls for patients with chronic and recurrent LBP (109).

Regarding CLBP, there is strong evidence to support the use of advice to remain active, in addition to specific advice relating to the most appropriate exercise and/or functional activities by which to promote active self-management (106).

Pain Management

No clear evidence exists to the effect that antidepressants are more effective than placebo in the management of patients with CLBP. These findings do not imply that severely depressed patients with LBP should not be treated with antidepressants (113). The benefits of opioids in clinical practice for the long-term management of CLBP remain questionable (114).

Both TENS and traction, as a single treatment for LBP, are probably not effective (101,115). Low-level laser therapy (LLLT), in contrast to a sham treatment, may be beneficial for pain relief and improved disability in patients with SALBP or CLBP, although the treatment effects are small. However, when LLLT is added to exercise and compared to exercise therapy, either with or without sham treatments, there appears to be little or no difference between the groups in terms of pain and disability (116).

The evidence is conflicting in regard to the efficacy of prolotherapy injections for patients with CLBP. When used alone, prolotherapy is not an effective treatment for CLBP, but combined with spinal manipulation, exercise, and other interventions it may improve CLBP and decrease disability (117).

Rehabilitation Tools

The rehabilitation treatment tools used in the field of LBP have three different backgrounds: mainly physical (exercises); mainly psychological (cognitive-behavioral approaches); finally social (education approaches). These are only tools that can be combined in a thorough rehabilitation of the individual patient so as to achieve the best results. They should not be considered separately, but only as ingredients in the correct mix designed to achieve the best individual rehabilitation. In fact, there is strong evidence that intensive multidisciplinary biopsychosocial rehabilitation with a functional restoration approach is useful for CLBP (41,42); in this case, the complete rehab package includes exercises (functional restoration) (118,119) together with psychological (cognitive-behavioral) (111) and social approaches to the patient (40,42) that allows to achieve the best results.

Exercises

The use of exercises as a therapeutic tool for LBP is quite common in most countries of the West (120), despite scientific evidence that shows its efficacy to be quite inconsistent. In a 2000 Cochrane review, the authors highlight the modest effectiveness of exercises in treatment for patients with CLBP while noting its effectiveness in reducing work absenteeism in patients with SALBP. In ALBP, the benefits of exercises are considered to be as effective as conservative treatment or non-treatment (110).

The low effectiveness of exercises, as reported in a great number of studies published in scientific journals, is in contrast to the benefits perceived by patients and experts in the treatment of LBP. This discrepancy may be due to the fact that in a great number of these studies with high scientific impact the subjects are randomly chosen and placed in various treatment groups without being first classified into subgroups based on the criteria of pain characteristics (121–124). Such a method, being conceptually flawed from the outset, can skew the outcome of clinical trial results. Because classification based on a proposed relevant pathology is possible in at least 10% of cases (125), one of the objectives of a diagnostic procedure would be to allow the collection of data useful in placing subjects in homogeneous groups in order to prescribe the most appropriate treatment (121). If it is not possible to classify according to etiology, presumably it will be possible for functional characteristics, or others that are now under scientific exploration. Until then, we have to wait before being able to state clearly what the evidence on exercises can be.

Moreover, this may explain why various types of exercise—albeit with very different physiological origins—have all shown benefits in treating LBP, particularly CLBP (110,120,126).
A few years ago, an open question in the treatment of patients with LBP was to explore whether trunk-flexion exercises would be more efficient than trunk-extension exercises. Studies did not completely resolve this question because the choice of prescribing a preferred course of exercises could not be entrusted to randomness but instead through a careful examination of subjects based on the characteristics of the symptom and its modification (62). Recently, an in-depth meta-analytical study examined precisely the types of exercises that were potentially the most effective in the treatment of CLBP (110,120,126). While promising results emerged from the study, both in the area of muscle stretching and strengthening, its main conclusion was that much more research would be needed before a better definition of protocols could be reached, and that the best course is most likely that of subclassifying patients.

The objectives of exercise (Fig. 33-3) could include:

• Reducing symptoms;
• Regaining function, reducing disability and fears related to movement, and encouraging regular physical activity in SALBP and/or CLBP;
• Preventing relapse.

Reducing the sensation of pain could be linked to biological changes in tissues, thanks to increased blood circulation, improved mechanics of the joints in question due to the stimulation of joint capsules and their ligaments, better function of stabilizer muscles, and a neurological de-sensitization of tissues, thanks to the repetition of movement (127).

In SALBP and CLBP, along with pain the patient must also face various other problems. Many studies have shown that “pain-related fear” has a negative impact on some basic spinal functions, such as elasticity and strength (128). This condition increases the risk of a progressive worsening of disability and the development of a deconditioning syndrome, which could be an additional obstacle to improving conditions. Another typical behavior is that a CLBP patient may exhibit “avoidance behavior”: The patient fears incurring back damage and adopts motor behavior marked by excessive caution (129), which further worsens motor neurological quality. An intensive, targeted exercise program could reduce the risk of kinesiophobia and have a positive impact on the related disability (130,131). In this respect, there are proofs that a graded increased functional training has good results in SALBP (132–134), but this methodology can quite easily be extended to CLBP patients (110,120) allowing to achieve cognitive-behavioral goals as well (111).

There are many studies that have shown how the characteristics of the muscles that maintain the stability of the spine change after an episode of LBP. The lumbar multifidus muscle exhibits a delayed activation (135) and a decrease in the transversal size on the side of the pain (136). It has been observed notably that in the absence of specific treatment these deficits remain, even when the LBP disappears. The high incidence of recurrence in LBP the year after the first episode could be due to precisely this functional weakness in the stabilizer muscles. Some studies have supported this hypothesis by showing how a program of functional reactivation of these muscles could work as a safeguard against future episodes. A year after an episode, a group of subjects who had participated in a specific program for strengthening the lumbar multifidus muscle had a 30% rate of relapse compared to 84% in the group of patients who had not participated. Even a follow-up three years later revealed a significant difference in the results, with a 35% rate of relapse in the group tested compared to 75% in the control group (137).
The same results were observed for adolescent LBP (138,139). The choice of favoring a program of stabilizing exercises to prevent the relapse of LBP is not clearly supported by scientific literature. In the case of recurrent nonspecific LBP, a program of general exercise reduces disability more effectively than a program centered on stabilizing the spine, which should only be an option for cases having obvious signs of instability.

Finally, the so-called functional rehabilitation approach is discussed (41,118,140,141). This gained popularity in the recent past and has been proposed as an inpatient intensive training of 3 to 5 weeks (142,143) or as a long-term training program, mainly machine-guided outpatient training (67,119,144). The main theory at the basis of this approach is derived from sports medicine (141) and considers function more than pain: in fact results are best in the functional-disability domains than on the pain itself that is in any case decreased (41). Even without considering that any kind of rehabilitation by definition must be “functional” (22), this approach has shown good results (41) and must be adopted as a concept, even if other settings (i.e., outpatient without specific machines) can be considered apart from the ones presented in the literature.

Cognitive Behavioral Approach

Cognitive-behavioral intervention is commonly used in the treatment of disabling CLBP, and it originates from a new viewpoint in regard to chronic pain. The traditional medical approach considers pain as a cause of illness and consequent disability according to an established illness model (4,12). This provides a circle in which physical damage causes pain that will eventually cause impairment, and the impairment will ultimately induce a disability. Nevertheless, while acute pain has a biologic means of alarm to signal tissue damage, chronic pain lacks this characteristic; it is not only influenced by somatic pathology but also by psychological and social factors (44). Moving from this consideration, Waddel (Fig. 33-4) theorized a new model of illness for LBP, known as biopsychosocial model, in which various aspects can determine and explain chronic pain. They are:

- **Physical dysfunction**: It depends on an imbalance between the demands of physical ability and real body capacities that are not ready to provide the required performances.
- **Belief and coping**: Human thought and the way of perceiving pain play a crucial role in determining how the patient manages his health problem. Frequently, patients with CLBP are persuaded that they suffer from a serious pathology and are therefore hardly considered curable. They have incorrect assumptions about the possibility of recovery, often due to previous failed treatment. This leads them to adopt a discouraged approach to the problem and to new proposed treatments. The different strategy for coping with pain can explain why certain patients overcome the acute phase, while others come to suffer from CLBP. There are two means of response to pain: actively face it (copers), or undergo it (noncopers). Pain-related fear, catastrophizing beliefs, and lack of psychological and cognitive instruments with which to oppose painful symptoms lead the patient to assume an avoidance behavior from the same pain. This means the reduction of physical activity, work, and social relationships until one arrives at a physical and psychological deconditioning.
  - **Distress**: Increased pain perception, emotional stimulus, psychological factors are deeply linked, and they can give rise to a vicious circle. Feelings of fear, anxiety, anger, and depression are common in patients of this kind.
  - **Illness behavior**: It is heavily conditioned from prejudices about the pathology, future treatments, and the ability of medical care to resolve the pain.
  - **Social relationship**: Social networks such as family, friends, and colleagues can influence the emotional status, development of illness beliefs, and coping strategy. A favorable family activity can help to face and overcome pain, while an accommodating ground to illness will increase it. This model provides a multidisciplinary approach to the problem that requires, above all, a multidisciplinary treatment through the use of different techniques appropriate for the individual subject, such as his or her psychological state.

Two systematic Cochrane reviews (42,111) and additional trials (132,133,145), all of which are considered high quality studies, showed there is strong evidence that behavioral treatment is more effective for pain, functional status, and behavioral outcomes than placebo, no treatment, and waiting-list control, most of all when it is intensive, and that a graded activity program using a behavioral approach is more effective.
than traditional care for returning to work. One low-quality trial (146) found there is no difference between the effects of behavioral therapy and exercise therapy in terms of pain, functional status, or depression for as long as a year after treatment.

The goal of the cognitive-behavioral approach to nonspecific CLBP is the ability to modify wrong beliefs about health status and changing the perception of health. Weisenberg (147) and Meichenbaum (148), who in 1977 first introduced this model, support the importance of change in the health pattern by a cognitive incentive, seeking to modify the patient's relationship with the chronic pain by offering him/her the possibility of reacting to pain through an awareness of the real problem. This approach must be presented to the patient like a process of correct learning (14,30), in the passage from an illness behavior to a wellness behavior. Generally, three behavioral treatment approaches can be distinguished (149,150):

- Operant treatment, which is based on the operant conditioning principles of Skinner (151) and applied to pain by Fordyce (152) consists of the positive reinforcement of healthy behaviors;
- Cognitive treatment, which aims to identify and modify the patient’s cognitions regarding his/her pain and disability;
- Respondent treatment, which aims to modify the physiological response system directly.

The first therapeutic aim is to forecast the positive effects of treatment results by acting on external events. This approach allows the patient to move from a control pain model, typical of the acute phase while improving behavioral and functional ability through communication, education, and motivation, which are methodological instruments peculiar to the cognitive-behavioral model. Communication must be efficacious and bidirectional in order to educate the patient in viewing his health condition from the correct perspective. Physicians and therapists should ensure that the patient understands his/her problem and that they are ready to help. Only in this way they can gain the patient’s confidence, which is necessary to ensure good compliance with the treatment. The advantage of the educational program is to offer the patient explanations about the true extent of the problem through easy lessons about anatomy and physiology until the patient can be clearly informed in regard to CLBP. It is useful to encourage the patient to manage his/her pain instead of simply suffering, and giving him/her simple means to be applied in his/her everyday life. This is done by explaining how this active approach to LBP influences, in a crucial way, the perception of pain and the disability correlated to it. This does not signify the minimization of the problem, but instead it helps the patient face it, rid himself from incorrect beliefs and the behaviors of pain avoidance that only serve to strengthen the pain.

The methodology applied is not simply “learn to change” but also “test the change.” To follow this model it is necessary to establish, before treatment begins, certain realistic aims, and to document the improvements through self-evaluation techniques for involving the patient so that he/she will be responsible for the change. The operating setup of this theoretical warning is to test during daily living what the patient learns and to make him/her aware of improvements. Central to diagnostic evaluation is an understanding of the deep interactions between the physical and psychosocial factors, and how they can support themselves according to a vicious circle, since disability in this kind of patient also means chronic pain, physical dysfunction, and illness behavior.

**Educational Tools**

A primary objective of LBP management is to provide the patient with accurate information. LBP treatment strategies have changed through the past few decades, thanks to the results of various clinical trials conducted during this period. An emblematic example of this is the general consensus regarding the recommendation of remaining active as much as possible during cases of ALBP (153). However, while it is an important element of LBP primary care, this consensus is not a widely held belief and many people continue to believe their patients need rest.

Accordingly, an updated collection of data to adequately address the issue is an important tool in modifying popular beliefs. In the case of ALBP, it is crucial to reassure the patient and inform him or her of the appropriate methods for managing symptoms (106). In the case of CLBP, the information must be particularly targeted toward the prevention of chronicity by providing the patient with useful advice on how to identify any behavior that could delay healing (106). As for CLBP, the information should include advice on how to manage pain, control catastrophism, and decrease avoidance behavior (106,111).

Certainly, an informational brochure is the most common educational tool (154,155). Many experiments have been carried out to quantify the brochure’s tangible effectiveness in changing common perceptions among patients, modifying behavior, and having an impact on pain and disability. For ALBP patients, the information contained in a brochure seems to be quite effective in reducing pain and the likelihood of relapse (156), as well as in decreasing fear-avoidance behavior (154). The brochure has also been used among the institutionalized elderly, and has shown a positive impact on improving disability in the 6 months following its use (157).

Back schools have been proposed in the past as a possible important tool for LBP treatment (106,109), and for a period of time they were quite popular. However, they have been widely criticized (158) because the original proposals (159–162) were mainly based on ergonomic assumptions and a disease-oriented model of illness that was clearly overcome by the actual biopsychosocial one. Nevertheless, the back school can be considered a therapeutic tool that can be filled with the most actual contents and/or according to the individual needs and/or clinical realities, more than a schematically uniform treatment. In everyday clinics this is the reality, because there are as many back schools as there are therapists applying them. Today there are some proofs of
efficacy, mainly in specific professional settings (109). As a therapeutic tool, if it is used as a cognitive behavioral exercise-based group approach to nonspecific SALBP and/or CLBP, the back school can be important as a low-cost approach to large numbers of patients (9).

Media is another interesting information tool for educational purposes. With a multimedia information campaign, a significant change can be observed in popular beliefs, with a considerable number of people abandoning the notion that they need rest when experiencing pain and instead embracing the correct idea of remaining active (163). This change in behavior is cost-effective and continues for several years after the end of the information campaign (25).

Lastly, various studies have shown that a crucial element in reducing the cost of LBP primary care management is the awareness of getting family physicians to adopt the right behavior (164), in which a multimedia campaign could play a part.

Main Pain Management Tools of PRM Interest

These approaches have different backgrounds, according to the pathoanatomic hypotheses proposed by the different authors and schools who have developed or use each of the tools. Some have a certain efficacy, but not in all patients: In an evidence-based clinical practice view they should be regarded not for their theory but only for their efficacy. Moreover, to date there is no consensus on how to choose each single treatment for each single patient (9). Presumably, and as far as we know today, the best way is through trial and error, starting with what is preferred by the patient and the treating physician. What should be avoided is absolute approach, coming from the idea of superiority of the PO treatment proposed versus any other, and based on specific diagnoses (diagnostic labels) (165,166) not scientifically sound: in fact it has been shown that such a medical style highly increases the probability of chronicization (96,165,167), constituting an important, iatrogenic yellow flag.

Manual Approaches

Hands are probably the most ancient tool man has employed in bringing relief for LBP. Both written evidence and artistic representations show that manual methods were highly valued even when very simple techniques were used, because there was little knowledge of anatomy and the mechanisms of joint physiology. Over the centuries, the progress of knowledge and wider experience have led to the development of various manual techniques for the treatment of LBP.

The term “manual therapy” includes techniques aimed primarily at the treatment of soft tissues, such as massage, mobilization techniques carried out to increase the range of motion, and techniques based on the application of small-amplitude high-velocity thrusts, such as manipulation (168). Theories proposed on the way of action of manual therapies are almost as many as the different schools (34); terminologies such as “ostheopatic lesion,” “minor vertebral derangement,” all supposing little injuries of the mobile segment mainly at the zygapophyseal joints, have been proposed; nonetheless, until any of these possible lesions will be proved this treatment is definitely a PO one.

Massage is undoubtedly the most ancient and widely used form of manual therapy. It is practiced in every region of the world and through a variety of techniques. Massage has also spurred considerable scientific interest, and this is shown in the large number of systematic reviews conducted by Cochrane to explore the effectiveness of its use. The results of these reviews on massage and LBP (169) have led to the conclusion that massage is effective in SALBP and CLBP, particularly when combined with exercise and education. The authors recommend further studies that would corroborate these conclusions and properly assess the cost-benefit ratio of manual therapy, its long-term impact, and the resumption of work.

Osteopathy and manipulation are two other manual therapy techniques widely used to fight LBP. While there is not much literature on the benefits of the former for treating LBP, there is considerably more in regard to the latter. One such review by Cochrane shows that in the treatment for ALBP and CLBP, manipulations are only more effective than sham therapies. There is no evidence that they are more beneficial than traditional treatments such as exercise or painkillers (170). Considering the low benefits and high cost, the effectiveness of these treatments is also low. However, it should be noted that when therapeutic techniques are characterized by a close relationship between the patient and the therapist—as in the case of manual therapy—even in those cases of treatment failure only 50% of the patients require alternative therapies, most likely because the ties forged with the therapist enable the patient to better face the symptoms and disability (171).

Modalities of Physical Therapy

Modalities of physical therapy include a wide variety of devices that apply physical principles in treating many problems in rehabilitative medicine. Actually, it is frequently proposed in LBP management in everyday clinical practice, sometimes alone but sometimes as part of a more complex approach. Despite the specific principle behind the action of each device, it can be considered a PO therapy because they act mainly on pain perception and transmission, even if they are supposed to possibly act on cell membrane or on inflammation.

Electricity is probably the main physical principle applied in this field, and within this group TENS is probably the form most commonly used. The development and application of TENS was based on the Gate Control Theory conceptualized by Melzack and Wall (172). According to this theory, the stimulation of large-diameter (A-β) primary sensory afferents activates inhibitory interneurons in the substantia gelatinosa of the spinal cord dorsal horn, thereby accentuating the transmission of nociceptive signals from small-diameter A-δ and C-fibers (172,173). Supraspinal mechanisms involving the endogenous opioid system have also been described (174–177). In summary, the postulated effect of TENS is to “close the gate” and dampen the perception of pain (172). Despite its rationale and well-documented biological effects, TENS as
a single treatment for LBP is probably ineffective (101,115). Other forms of electrical therapy have been proposed, and some papers have been published (178,179), even if we still lack a substantial basis for them.

Ultrasound has for many years been used in the treatment of musculoskeletal disorders. Laboratory research has demonstrated the application of ultrasound results in the promotion of cellular metabolic rate and increased viscoelastic properties of collagen (180). In animal studies, an exposure to 1 MHz ultrasound at 50 J/cm² is reported to be sufficient for the increase of tissue temperature (181). This rise in temperature is assumed to be the mediating mechanism for tissue repair, enhancement of soft-tissue extensibility, promotion of muscle relaxation, augmentation of blood flow, and alleviation of inflammatory reactions of soft tissue (180,182–184). Despite the theoretical benefits and widespread use, conclusive evidence on the effectiveness of ultrasound therapy in CLBP is not yet available.

LLLT is currently used by some as a therapeutic intervention for musculoskeletal disorders such as LBP (185,186). LLLT is a light-source treatment that generates light of a specific wavelength. It emits no heat, sound, or vibration. Instead of producing a thermal effect, LLLT may act through nonthermal or photochemical reactions in cells. It is also referred to as photobiology or biostimulation (187,188). LLLT is thought to affect fibroblast function and accelerate the repair of connective tissue (189). It has also been reported that LLLT has anti-inflammatory effects due to its action in reducing prostaglandin synthesis (190). Some studies suggest that LLLT has a beneficial anti-inflammatory and pain-attenuation effect in humans (191). A possible mechanism of the effect of LLLT on pain relief is its anti-inflammatory and connective-tissue repair process, which have been shown in some in-vitro and in-vivo studies (190,192). The effectiveness of laser therapy in painful disorders is still unclear and should therefore be examined more rigorously (185,193).

Pulsed magnetic field therapy (PMFT) is a simple, noninvasive technique used extensively for the treatment of muscle pain (194). This technique is based on changes in the cell membrane induced by magnetic fields (MF) and, to a limited extent, the electric field. Exposure to pulsed MF has been shown to have a therapeutic benefit in both animals and humans. MF exposure does not affect basic human perception but can increase pain thresholds in a manner indicative of an analgesic response (194). Nevertheless, no data is available regarding the effect of MF on CLBP.

All the aforementioned physical therapies have some basic research studies and have been widely applied, but conclusive research on their efficacy is lacking. These approaches can be considered in the PO phase of an integrated rehabilitative process, but until there is further proof they should be avoided as stand-alone LBP treatments.

Drugs
Pharmacological treatment is the first way to control pain in patients with LBP. There are various kinds of medications, and each one has a unique balance of risks and benefits. A systematic Cochrane review (97) of NSAIDs—which are widely used in LBP—demonstrated that there is a moderate evidence that NSAIDs are not more effective than other drugs for ALBP, and there is a strong evidence that various type of NSAIDs (including COX-2 NSAIDs) are equally effective. COX-2 NSAIDs had statistically significantly fewer side-effects than traditional NSAIDs, but recent studies have shown that COX-2 inhibitors are associated with increased cardiovascular risk in specific patient populations.

The use of muscle relaxants (195) in the management of nonspecific LBP is controversial even if they seem to be effective. However, the adverse effects—particularly, the effects on the central nervous system—require caution in the use of muscle relaxants. The potential for tolerance and withdrawal, combined with the risk of misuse and dependency, lead clinicians to restrict the prescription of opioids, even if they are commonly used for CLBP and may be efficacious for short-term pain relief. The long-term efficacy (>16 weeks) is unclear (114,196). A Cochrane review (113) reported that there is no clear evidence that antidepressants are more effective than placebo, even if such findings do not imply that severely depressed patients with LBP should not be treated with antidepressants.

A pharmacological strategy for the treatment of LBP, particularly in regard to leg pain, is the epidural injection of corticosteroids, for which (92) there does not appear to be any evidence to support the current common practice of a series of injections. Abdil et al. (197) reported that there is limited evidence regarding the lumbar spine for long-term relief for interlaminar epidurals and moderate long-term improvement in managing nerve-root pain for transforaminal epidural steroid injection.

A review of evidence for an American Pain Society/American College of Physicians (198) recently summarized this knowledge in terms of good short-term evidence with moderate effectiveness for NSAIDs, acetaminophen, and skeletal muscle relaxant (for ALBP), while the minimal one proposed on tricyclic antidepressant for CLBP has been overcome, according to the last Cochrane review (113). They also found a reasonable amount of evidence that opioids, tramadol, benzodiazepines, and gabapentin (for radiculopathy) are effective for pain relief as opposed to systemic corticosteroids, which the evidence strongly suggested was ineffective. The authors found that evidence is insufficient to identify one medication as offering a clear overall net advantage due to the complex tradeoffs between benefits and objectives. This leads to the choice of one or more medications after a specific analysis for each patient, taking into account the comorbidity and others drugs just assumed.

Low Back Pain
We definitively recognize, according to the actual knowledge, the term “LBP” as a diagnosis instead of being merely a symptom. LBP diagnosis is made by exclusion through a triage (5): in such cases other terms such as “sprain,” “injury,” “trouble,” or “lumbago” have been used for years, in most cases supposing and supporting the idea of different specific pathoanatomic background. Today the scientific community has abandoned this approach: the most up-to-date classification
of LBP is based on the localization and on the duration of pain (3,8–10). Even if this could appear as an epidemiologic framework unduly applied to the clinical field, actually it is the only real way to differentiate patients with different prognoses and pathologies. Hopefully in the future it will be possible to deepen this broad classification, but it is already a good way to arrive at clinical and rehabilitative everyday choices.

Research thus continues its gradual progress toward the subclassification of different syndromes, mainly with the objective of increasing the quality of approach (62,122,124). However, this process will take years to reach a satisfactory conclusion.

The incidence, prevalence, and costs of LBP are incredibly high (59,199–205). ALBP life prevalence is more than 80% of the population, year prevalence counts up to 30%, and costs in terms of absenteeism rank second in importance, following only cold and flu. However, over 90% of ALBP resolve in less than 30 days; on the other side, CLBP is an everyday experience for 4% to 7% of the population. It consumes 75% to 80% of the entire enormous costs of LBP and less than 5% of patients achieve a complete resolution of pain. SALBP has scarcely been studied and today we have no reliable epidemiological data, even if this stage of pathology should deserve the highest attention.

**Acute LBP**

**Definition and Pathogenesis**

ALBP is defined as pain or discomfort in the lumbar region, on one or both sides, eventually irradiating to the buttocks and lasting no more than a month (3,8,9,52). What really characterizes ALBP is that we do not know exactly where it comes from (96,165): Presumably, we can exclude the involvement of bones and main nerves. Instead, soft tissues are the main actors (muscles, ligaments, apophyseal joints), together with discs eventually and/or vessels, and/or the sympathetic nervous system. Nevertheless, the hypotheses are as numerous as the authors who have studied them.

**Clinical Presentation**

ALBP is clinically multiform: it can appear suddenly when one encounters a significant physical effort, or it can appear without an apparent cause. It can start insidiously, or it can arrive at the complete lumbar blockage. It can start on one side and move to the other, or remain fixed in time. Independently from the quantity of pain, it can last from several minutes to several weeks. Apparently the level of pain has no correlation with the anatomical lesion and its recovery.

**Diagnosis, Classification, and Prognosis**

Diagnosis is made by exclusion, and it must be done only through clinical examination (3,8,9,52). Red flags must be carefully excluded (see Table 33-1), and if there are no warning signs the health risk and the costs of imaging, and other exams will not be justified (3,8,9,52). Today we have no useful accepted subclassification, even if scientific efforts are underway in this respect (121,123). Prognosis is very good: ALBP is auto-resolving in most of the cases, usually in less than a month (59,206). Most of the time, ALBP is recurrent (206–208). Even if there is a clinical widespread suspicion, there is no proof that it implies a progressive worsening of the clinical situation (9).

**Treatment Approaches**

ALBP should be treated mainly by general practitioners (GPs), both to reduce costs and increase clinical long-term results (9,23,25,209–211). The main objective and therapy in ALBP is, in fact, to reassure the patient and provide accurate preventive information (3,8,9,52). PO treatment is useful but secondary, even in the eyes of patients (212). Specialists should reinforce the role of the GPs and eventually offer them an educational support (23,210).

**Rehabilitation Approach**

There are very few issues on which the international scientific community has reached a general consensus. One of these issues is the right behavior to adopt in cases of ALBP. The key recommendation is to remain as active as possible and to avoid bed rest. This advice stems from consistent results (207,213) from a large number of scientific studies that have prompted the publication of a Cochrane review (169) whose findings recommend not to invest any further effort in examining this issue, as it has been abundantly proven.

With regard to LBP, the term “rehabilitation approach” is a bit of a misnomer, as a patient suffering from LBP for less than a month does not present any disability at this stage, and education appears far more appropriate. This patient merely needs reassurance and support until symptoms disappear.

Caring for a patient suffering from ALBP should, however, take into account two important factors related to prevention simultaneously: avoiding chronicity and regaining any fitness lost during the period of acute pain which may open the way to frequent relapses.

The real issue with LBP entails the progression from an isolated episode into chronic pain. This transformation depends on a complex series of not always identifiable psychosocial factors, “yellow flags” (58,59,214) that must be carefully taken into consideration and which could enable the health care provider to distinguish between a single episode of LBP and the transformation into a chronic disorder (215).

Many studies reveal how LBP brings about a rapid change in the function of the paravertebral and stabilizer muscles (136,216,217). It was noted, in particular, how rehabilitation of the musculature is not spontaneous with the disappearance of the symptoms and it is believed that this deficit may be one of the reasons for the frequent relapses reported after an episode of LBP (137). Studies have shown that the rate of relapse drops drastically when the patient participates in a rehabilitation program focused specifically on the stabilizer muscles. A short-term follow-up (1 year) shows that the relapse rate drops from 87% in the group that did not participate in a specific program to 30% in the group that did follow the program. This same gap also holds for
the medium term. The probability of experiencing LBP 3 years following the first episode is 35% in subjects who participated in a specific exercise program, compared to 75% in subjects who returned to their daily routine without exercise (137,218).

Subacute LBP

This is the less known type of LBP, and research in this field is really lacking, even if in recent years it has progressively flourished due to the understanding of the key role of this stage of LBP.

Definition and Pathogenesis

SALBP is defined as pain or discomfort in the lumbar region, on one or both sides, eventually irradiating to the buttocks and lasting more than 1 month but less than 6 months (3,8,9,52). Some authors distinguish a subacute phase (1 to 3 months) and a subchronic phase (4 to 6 months), but there is neither epidemiological proof nor a real clinical decision based on that distinction. The pathogenesis of SALBP is complex. As far as we know, it is a real intermediate stage, being the passage between ALBP and CLBP. In this period, apparently there is something that impedes the spontaneous resolution of ALBP, maintains the previous situation, makes it more complex and gradually drives the patient, through a series of vicious cycles, to CLBP (4,12,219). These factors have been studied in these years and therefore recognized as the chronicization risk factors (see Table 33-1) (58,77,214,220–224). According to the definition of biopsychosocial syndrome, they can be divided into biological, psychological, and social factors. Moreover, certain other distinctions, as shown in the table, can be made. What can be supposed, according to the actual knowledge, is that an organic and definite cause of pain, which is manifested in the acute phase, cannot reach a resolution because of pre-existing factors and/or newly developed that appears in predisposed personal/social contextual factors. In this way the resolution becomes more difficult. The patient usually starts a pilgrimage between different specialists and professionals, each one giving his/her own etiological opinion, proposing a treatment, and using his/her own diagnostic labels (165,166). All this makes the situation progressively worse because it encourages the patient’s feeling of having a bad pathology that is impossible to understand and cannot be treated in a definitive way.

Clinical Presentation

Apparently, the SALBP patient is not very different from the ALBP patient, even if a careful history-taking and a good clinical examination can usually let the expert physician discover the stigmata of chronicization. A key factor is to look for the time passed since the inception of LBP, and another one is to let the patient speak freely about his or her problem. This way, if there are psychological and/or social disturbances, they usually come out quite easily, and the presence of relatives and their interrelations can offer other clues. What must be clearly pointed out is that these patients are usually quite the same on the surface, but with due time and careful evaluation it is possible to slowly discover a multiform, individual manifestation of SALBP. Ultimately, this constitutes the real clinical presentation.

Diagnosis, Classification, and Prognosis

Diagnosis, as is the case in all LBP situations, is made by exclusion. Nevertheless, it is widely recognized that further evaluation is needed at this stage, because true spinal pathologies must be excluded (3,8,9,52). Standing (and eventually dynamic) x-rays in the case of LBP, or MRI (and/or TC) in the case of sciatica, and eventually other ematological and/or neurophysiological exams according to each individual evaluation can be proposed (Figs. 33-2 and 33-5) (3,8,9,51,52). Nevertheless, these examinations should not be repeated continuously once the supposed pathology has been ruled out (51,225). As soon as a definite disease can be excluded, the patient should be evaluated carefully in a biopsychosocial way so as to understand which contributing factors are maintaining the pain, so that the patient can be treated accordingly. There are no subclassifications for these patients, nor is the prognosis as good as one might expect, even if treatment has a key role in this process (220).

Treatment Approaches

SALBP is mainly of PRM interest (Table 33-4) (9,21). In fact, it is the situation in which it is still possible to avoid CLBP. However, this necessitates a definite rehabilitative approach. In this situation, maintaining a PO attitude that is usual in the disease-oriented specialties and in some “interventional” approaches, without being able to solve the problem definitively but instead using diagnostic labels, is a recognized risk factor of chronicization and should be avoided. Only a true biopsychosocial approach can lead to a possible resolution.

Rehabilitation Approach

The approach to the SALBP patient is not easy because of the already painted multiform characteristics of these patients. Moreover, there is clearly a lack of research and the evidence is quite weak for all types of treatments (226). Nevertheless, a clear-cut clinical behavior emerges from research and the knowledge we have regarding ALBP and CLBP (see Table 33-4).

The first choice is to decide whether there is a need for PO treatment. There is evidence regarding its efficacy (226), even if drugs do not significantly decrease the disability (227), but this solution should be considered carefully because most often it is what the patient has always received, which usually means short-term results with no final resolution. Some patients will consider it a nonsolution, while others already run the risk of finishing in the “Saint Graal” vicious cycle (the infinite but sometimes impossible research for the “final solution” of their pain, possibly passive and without any individual effort), that physician behaviors can start and maintain. Consequently, a PO treatment can be proposed but should avoid any diagnostic label. Instead one should carefully explain that this is not the final solution but a treatment with possible short-term benefits that can be added to the compulsory FO rehabilitation.
Rehabilitation is the most important approach, combining educational, cognitive-behavioral, and physical exercise treatments according to the individual needs (40,228–231). Generally, patients in the first subacute stage are more in need of educational approaches than individual cognitive-behavioral ones, while the opposite appears more appropriate in later stages (subchronic ones), where the painful situation can be more established with many psychological and social vicious cycles already in effect (9). Nevertheless, these choices must be totally individual, as it happens for the physical-exercise approach whereby the evaluation leads to the exercise.

The approach to SALBP must definitely comprise a team. This is because a multiprofessional contribution facilitates a better understanding of the different risk factors involved, while maintaining coherent messages from different professionals has a crucial reinforcement role that allows the best results to be obtained. The multiprofessional evaluation must focus on the different risk factors of chronicity (see Table 33-1), so as to individually face these factors and prevent the establishment of the CLBP syndrome. Consequently, the physical, psychological, and social approach will change on an individual basis. This gives SALBP treatment the considerable adaptability required by the situation itself.

The efficacy of a psychological approach (232) to SALBP, even in the hands of specialists other than psychologists, has clearly been shown. Moreover, this is the only way to treat these patients who perceive themselves as “not mad” but as psychologically healthy. The psychological approach must be part of the treatment. However, it should be in the background instead of being the main goal, because most often the patient will not accept a direct psychological intervention and/or a treatment that is totally psychological. This is particularly so in SALBP, when the psychological implication of pain has not yet been fully understood by the patient and his or her relatives.

### Chronic LBP

#### Definition and Pathogenesis

CLBP is a pain with or without functional limitation in the posterior region, including the area between the inferior limit of the costal arch and the inferior buttock fold that lasts more than 6 months. Chronic sciatica is a pain in one or both legs that lasts more than 6 months. Given the absence of real differences between the two symptom presentations in terms of pathogenesis and clinical approach, they can be considered together (9). Even if the only difference between acute and chronic pain definitions seems to be the persistence of pain, their pathogeneses and clinical pictures are quite different.
In the literature three frequently used models regarding the development and maintenance of CLBP functional limitations are described (140):

- The physical deconditioning model, which assumes that loss of muscle strength, endurance, and aerobic capacity, is responsible for reduced activity levels and hence functional limitations (94).
- The cognitive behavioral model, which postulates that functional limitations result from the maladaptive beliefs and avoidance behaviors that are maintained by learning processes (150,233).
- The biopsychosocial model, which assumes that the loss of functional abilities results from the deconditioning and the cognitive behavioral model (234).

CLBP, as far as we know today, can be defined as a biopsychosocial syndrome (4,12,39) in which all the aspects of the syndrome have developed over time: The patient is enveloped in a series of vicious cycles that emerge from SALBP through a full development of the chronicity risk factors for LBP (see Table 33-1). A simplification of all these processes is represented in Figure 33-6: When pain is not resolved, altered behavioral answers inevitably develop.
The patient develops a fear-avoidance behavior (244,245). It also happens that limitations due to the fear of pain are much more disabling than the real intensity of the pain. CLBP patients are more sensitive to the risk of feeling pain, so they frequently limit their activities in order to prevent pain. This leads to a “nonuse syndrome,” sometimes called deconditioning syndrome (142,246), which is a progressive loss of physical abilities that can maintain and contribute to CLBP (94) even if there is an absence of current data to support this concept (93).

**Diagnosis, Classification, and Prognosis**

Again, diagnosis is made by exclusion (see Fig. 33-2). Usually, CLBP patient has already undergone a number of Rx, CT, MRI procedures and many others, so that no further exams are really necessary in order to define a diagnosis. If the chronic situation is quite new and there has been no SALBP evaluation, then this must be done (9). Quite often one will notice that the patient has been to many different MDs, therapists, and health professionals in search of a definitive answer to the problem, and usually such a patient has received as many diagnoses, exams prescriptions, and therapies as there were consultations. Almost every time they are disappointed by the outcome, which has been provisional at best or nothing at worst. Actually, the main reasons for these scant outcomes are inadequate treatments that were merely PO instead of FO therapies.

A major topic in LBP treatment is the attempt to subgroup patients (62,121–124). It is generally agreed that LBP is a common final pathway of many different clinical pictures, of different problems of the back that, complicated by chronicity factors, have led to a more complex clinical situation with a disability sometimes quite important (22). The possibility of subgrouping patients in a reliable way could be the opportunity to provide for each one a more personalized and specific treatment. To date, no subgrouping shared by LBP experts has been made.

A subgroup of CLBP patients has been described using the deconditioning syndrome term, meaning patients in whom the physical aspects of dysfunction and deconditioning prevail (93,94,118,127,144,242,243). Nevertheless, this syndrome is not completely defined. Apparently the more useful approach is to subgroup the patients on the basis of disability (9,21). This is done using specific questionnaires, that is, the Roland Morris (247) (cut-off 14 points (81)) and the “Oswestry Disability Questionnaire” (83) (cut-off 25 points (81)) (see Fig. 33-2). They were designed for LBP patients and are now disseminated worldwide, and have been translated and validated in many languages. The usefulness of this classification relies on the ability and opportunity to choose patients requiring a multidisciplinary approach and those who can be initially treated with a less-expensive approach (9,21).

CLBP resolves completely in very few patients (207). Thus, the patient should be informed about the difficulties in reaching a complete restitution to wellness, and should be provided with tools to manage the problem. There is scientific consensus that the predictors of prolonged disability are more
TABLE 33.5  CLBP Treatment (9)

<table>
<thead>
<tr>
<th>Evidence</th>
<th>Contents</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A</strong></td>
<td><strong>Aim of Treatment</strong>&lt;br&gt;Chronic pain resolution occurs in &lt;5% of patients. In case of low-disability, aim of treatment is reducing actual disability and avoiding its progression through instruments to manage the problem (active approach by the patient) and control pain</td>
</tr>
<tr>
<td><strong>A</strong></td>
<td><strong>Counseling</strong>&lt;br&gt;There is no significant pathology&lt;br&gt;It is difficult to eliminate pain completely&lt;br&gt;Pain can be reduced&lt;br&gt;It is possible to improve quality of life and reduce disability&lt;br&gt;Learn pain management&lt;br&gt;Reduce stress&lt;br&gt;Be fit&lt;br&gt;Work is not enemy&lt;br&gt;Physical exercises are important and useful</td>
</tr>
<tr>
<td><strong>A</strong></td>
<td><strong>Work and ADL Interventions</strong>&lt;br&gt;Continue/resume gradually&lt;br&gt;Eventually change/reduce work activities&lt;br&gt;Control posture&lt;br&gt;Reduce stress</td>
</tr>
<tr>
<td><strong>A</strong></td>
<td><strong>Physical Activities Counseling</strong>&lt;br&gt;Start gradually, preferred physical activity&lt;br&gt;Practice regularly, at least twice a week</td>
</tr>
<tr>
<td><strong>A</strong></td>
<td><strong>Expert Physician Evaluation</strong>&lt;br&gt;Complete diagnostic reevaluation&lt;br&gt;Physical fitness evaluation (biological)&lt;br&gt;Behavioral evaluation (psychological)&lt;br&gt;Disability evaluation (biopsychosocial)</td>
</tr>
<tr>
<td><strong>A</strong></td>
<td><strong>Expert Multidisciplinary Team Intervention</strong>&lt;br&gt;See Table 33-6</td>
</tr>
<tr>
<td><strong>B</strong></td>
<td>Chronic LBP therapy changes according to patient disability level (low or high)</td>
</tr>
<tr>
<td><strong>C</strong></td>
<td>In case of low-disability, a not-expert approach is possible</td>
</tr>
<tr>
<td><strong>C</strong></td>
<td>There is no evidence of efficacy in the literature of a not-expert approach, but it could be preferred in terms of cost/benefit ratio</td>
</tr>
<tr>
<td><strong>C</strong></td>
<td>A multidisciplinary approach is complex, nevertheless it is preferable in case of:&lt;br&gt;High-disability&lt;br&gt;Low-disability but early chronicization (it is still possible to solve the problem)&lt;br&gt;Low-disability, no previous trial of this approach and highly motivated patient</td>
</tr>
<tr>
<td><strong>C</strong></td>
<td>Multidisciplinary approach is not recommended in case of low-disability and:&lt;br&gt;Complex treatment difficult because of cognitive, psychological, or motivational factors&lt;br&gt;The patient does not believe in a possible solution</td>
</tr>
</tbody>
</table>

psychosocial than biomedical in nature (77,220,248–250). It has been shown that fear-avoidance beliefs about work and physical activity, catastrophizing the lack of belief in one’s own ability to manage pain, cope and function, and self-efficacy beliefs are all significantly related to disability in patients with chronic pain.

**Treatment Approaches**

A CLBP patient is a typical rehabilitation patient because of the disability deriving from long-lasting pain (9,21,22). At this stage the attention should not be centered only on pain, because pain itself is only part of the picture (Table 33-5). It also consists of loss of function, physical deconditioning, and fear-avoidance behaviors. For many years it has been clearly demonstrated that the attention should be focused on function more than on pain (96,251). A PO treatment alone is destined to fail, and usually many different approaches have been proposed and performed without lasting results in each CLBP patient. This point should be clearly explained to patients, since the call for the quickest possible solution to pain...
CHAPTER 33 | REHABILITATION OF LUMBAR SPINE DISORDERS: AN EVIDENCE-BASED CLINICAL PRACTICE APPROACH

is common, and pills are the symbol of every quick solution. A PO therapy can be a useful adjunct to FO rehabilitation, especially in high-disability LBP, but it cannot grant by itself a lasting improvement of pain and disability. Some patients use NSAIDs, muscle relaxants, or sometimes opioids or steroids every day, believing this is the only way to survive pain. Usually, when they try to reduce the amount of drugs they discover their pain is nearly the same. The literature data clearly shows that drugs are effective in CLBP only in the short term, but a psychological addiction can lead to abuse (97,198).

The Rehabilitative Approach

Overview
The rehabilitation tools that are effective in CLBP include biopsychosocial inpatient approaches (42), cognitive behavioral approaches (111), exercises of different kinds (126), and educational interventions (109). All these differ in their cost/benefit ratios, even if they have generally been shown to be effective in selected patients. They can be considered as different rehabilitation tools in the hands of the treating physician: the choice of the more adequate depends on individual features of the single patient. There is quite a range of clinical presentations among CLBP patients, since physical deconditioning and dysfunction, fear-avoidance beliefs, social and psychological factors, and functional limitations can combine in many ways. Therefore, we should choose the less “invasive” and less costly treatment as well as the most effective treatment possible on an individual basis, even if there is no scientific way to distinguish groups of patients according to disability (9).

Regarding high-disability CLBP (9), one should propose as the most effective rehabilitation treatment a cognitive behavioral approach with specific exercises, aimed at recovering the physical limitations of the back, improving function, and giving the patient the right coping strategies. In low-disability CLBP, the suggested intervention is a back school in-group with exercises and education (9); if that does not work, other steps forward include an individual approach with specific exercises, individual cognitive behavioral therapy, and a cognitive behavioral approach with specific exercises (Table 33-6). According to the DTP (9), a higher level of disability justifies the choice of higher cost and individual effort for therapy, while a lower one could benefit from an approach that is generally considered less effective.

General Guidelines
Despite the foregoing, there are some basic points coming from the literature to be maintained in all CLBP patients, whichever

---

**TABLE 33.6** High-Disability and Chronic Low-Disability LBP Treatment (9)

<table>
<thead>
<tr>
<th>High Disability</th>
<th>Low Disability</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Evidence</strong></td>
<td><strong>Contents</strong></td>
</tr>
<tr>
<td>A</td>
<td>Multidisciplinary Rehabilitation (1,2)</td>
</tr>
<tr>
<td>A</td>
<td>Functional recovery therapy with cognitive-behavioral approach</td>
</tr>
<tr>
<td>A</td>
<td>Individual cognitive-behavioral therapy</td>
</tr>
<tr>
<td>B</td>
<td>Specific individual exercises</td>
</tr>
<tr>
<td>A</td>
<td>Back school in group (education + exercises)</td>
</tr>
<tr>
<td>C</td>
<td>Painkiller Therapy (3–6)</td>
</tr>
<tr>
<td>A</td>
<td>Paracetamol with or without opioid</td>
</tr>
<tr>
<td>A</td>
<td>NSAIDs</td>
</tr>
<tr>
<td>A</td>
<td>Antidepressants</td>
</tr>
<tr>
<td>B</td>
<td>Muscle relaxants</td>
</tr>
<tr>
<td>A</td>
<td>Pain-kill exercises</td>
</tr>
<tr>
<td></td>
<td>Manipulation/mobilization</td>
</tr>
<tr>
<td></td>
<td>Massage</td>
</tr>
<tr>
<td>C</td>
<td>Surgery (7)</td>
</tr>
<tr>
<td>C</td>
<td>Spinal fusion</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Evidence</strong></th>
<th><strong>Contents</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Multidisciplinary Rehabilitation (1,2)</td>
</tr>
<tr>
<td>A</td>
<td>Back school in group (education + exercises)</td>
</tr>
<tr>
<td>A</td>
<td>Individual specific exercises</td>
</tr>
<tr>
<td>A</td>
<td>Individual cognitive-behavioral therapy</td>
</tr>
<tr>
<td>A</td>
<td>Functional recovery therapy with cognitive-behavioral approach</td>
</tr>
<tr>
<td>C</td>
<td>Painkiller Therapy (3–6)</td>
</tr>
<tr>
<td>C</td>
<td>Paracetamol with or without opioid</td>
</tr>
<tr>
<td>A</td>
<td>NSAIDs</td>
</tr>
<tr>
<td>A</td>
<td>Antidepressants</td>
</tr>
<tr>
<td>A</td>
<td>Muscle relaxants</td>
</tr>
<tr>
<td>A</td>
<td>Manipulation/mobilization</td>
</tr>
<tr>
<td>A</td>
<td>Massage</td>
</tr>
<tr>
<td>B</td>
<td>Pain-kill exercises</td>
</tr>
</tbody>
</table>

**Notes**
2. Choice recommended according to cost-effectiveness, patient preferences, availability, and previous results.
3. Painkiller therapy should be proposed only when necessary.
5. Follow the specific indications of each treatment.
6. Painkiller therapy kills only pain, but is not therapeutic.
7. Only after 2 years of unsuccessful appropriate rehabilitation and according to patient’s choice.

Note: For definition of high and low disability, look at the text and Figure 33-2.
treatment is proposed. The main claim of a LBP patient is to somehow solve his or her problem, which is nearly constant pain, and hopefully to do it quickly. In the meantime, the patient is not so convinced that a solution is possible, because he/she has already tried many different approaches without achieving a final solution. Because a complete recovery is quite unlikely, even after a long time, the first step of any rehabilitation approach to CLBP should be informing the patient of this prognosis. It should be made clear that the goal of treatment is not to miraculously eliminate the pain but to improve quality of life and reduce the disability by giving the patient the right tools with which to manage CLBP. Obviously, hope must not be erased, but realistic goals must be set at the beginning in order to avoid another disillusion. Moreover, the patient will totally understand that these are realistic goals and the patient himself, due to previous experience, would not believe otherwise.

It has been demonstrated that coping patients have a different response to pain with respect to noncoping patients, who abnormally activate brain areas linked to the emotional processing of pain (237,252–255). Findlay, in a paper presented at the ISSLS Meeting in 2007, has shown that non-copers are “passive patients,” while copers are able to actively manage pain more adequately. It follows that each patient should be the primary actor in his own rehabilitation instead of being part of the audience. Patients are not used to being active during treatments in general, in any disease but in LBP particularly, and sometimes they react by showing disappointment: They want to be taken care of using a pill or a drug, by a physician, or through passive maneuvers by a physiotherapist, a chiropractor, or an osteopath. They want to receive some passive treatments and avoid doing anything independently. Indeed, changing their minds on becoming protagonists of the treatment leads to better results.

Good counseling of the patient can achieve many preliminary good results. The patient should be reassured that he or she has no significant malignant pathology and that even if we cannot completely eliminate pain we can reduce its intensity and decrease the number of acute episodes. Moreover, they must know that managing the situation by becoming copers, reducing physical dysfunction, and deconditioning will lead to a reduced disability and ultimately improve their quality of life. This will allow patient to regain many activities he or she has already experimented. It is necessary to emphasize that this approach is quite expensive, both physically and psychologically, and cannot be used in all patients. It is not recommended when cognitive, behavioral, or motivational factors are present or when the patient does not believe in the possibility of a solution (see Table 33-5) (9).

Cognitive Behavioral Approach to Physical Exercise

According to the literature, we envision a cognitive behavioral approach based on specific exercises, since it deals with the psychosocial and physical aspects of CLBP. Moreover, exercises are a very good and accepted way to reach a psychological recovery (110,111,132,134). The fear-avoidance behavior must be faced through a reassuring and courageous challenge to what is considered the main problem: movement. This approach can be used when the patient shows relevant psychosocial and physical limitations without a clear predominance of one aspect over another. The physiotherapist inside the rehabilitation team will choose where to concentrate the efforts according to the individual needs and reactions to treatment, both on the physical (exercises) and psychosocial aspects (counseling). The most typical case in which this approach should be the gold standard is the high-disability CLBP patient (see Table 33-6) (9). Usually in these patients, physical impairment is strictly connected to psychological and social limitations, so there is not a clear predominance of one single aspect. A clear causal relationship between these aspects is generally missing, because what we see is the final outcome of a vicious cycle in which pain limits physical activity, thus rendering movements more difficult and painful with each occurrence. Fear-avoidance beliefs are a crucial factor in generating and aggravating this cycle. Social and other elements are always part of the picture, but this sequence is so deeply bound that one cannot find the beginning.

Usually, a long period of outpatient rehabilitation (months) is needed in order to reach a satisfying result (118,119,256); alternative can be inpatient rehabilitation, that nevertheless must be adequately prepared in advance to reach the best results (142,143,243,257). Patients are not very trustful of therapists at the beginning, generally because they have already tried many therapies without significant results. It is very difficult for a patient to clearly understand the difference between this treatment and a general physiotherapy he or she has already experimented. It is necessary to emphasize that this approach is much more complex, team-driven, since it also deals with fear-avoidance beliefs and wrong behaviors, which are secondary to long-lasting pain. The patient should feel that the team trusts him and does not think he is crazy or a hypochondriac. Therefore, every professional must balance words and explanations in the proper way. It is highly relevant to this purpose that the cognitive behavioral treatment is performed by a physiotherapist. This makes the patient feel more comfortable, since he is generally not prepared for a real psychologist to handle this part of the treatment.

This approach is quite expensive, both physically and psychologically, and cannot be used in all patients. It is not recommended when cognitive, behavioral, or motivational factors are present or when the patient does not believe in the possibility of a solution (see Table 33-5) (9).

Cognitive Behavioral Approach

When psychosocial aspects are predominant and the general physical condition and function are almost preserved, we can
choose a simple cognitive behavioral treatment. The difference from the previous one is the lack of a specific physical approach with exercises. Usually, the patient who can benefit from this approach is quite young and has good overall physical function but relevant fear-avoidance beliefs, and is quite upset about LBP and the prospect for the future. Poor information from the media or a physician, who is not fully prepared, along with a general health-related fear or even real hypochondria, can lead this patient to chronicity, despite a physical limitation that, by itself, is not so relevant. The patient needs to be reassured about his/her future and health, and should be encouraged to return to sports without thinking too much about the previous experience of LBP.

**Physical-Exercise Approach**

On the other hand, a simple exercise approach that is specifically developed for the patient and is gradually increased (110,120,132,134) can be proposed when the physical aspects are relevant and the psychosocial aspects are less important. Even if a specific psychological treatment is not provided through this approach, the therapist should provide appropriate general counseling about back-related problems. This approach should be proposed to the patient who lacks psychological and social chronicity factors but is physically deconditioned and limited by dysfunction. A typical example is the older patient who does not have particular chronic stigmata but has undergone a relative loss of function consequent upon the progressive decrease in his or her level of physical activity. It is quite difficult for such a patient to start moving again, and often he or she will feel embarrassed about not being used to performing exercises, sports, or physical activity in a gym. Usually these patients achieve good results and recover to a good functional level, but they still need a maintenance protocol that will preserve the benefits.

**Cognitive Behavioral Back School**

A back school in-group with exercises and education is the least burdensome rehabilitative approach for CLBP. As with all group treatments, it has the advantage of being less expensive, and it gives the patient a point of comparison by being in contact with other persons with the same problem. The main limitation is that it cannot be as specific as an individual treatment. Nevertheless, it can provide good information and a valid psychosocial support for such patients. Its main indication is low-disability CLBP (see Table 33-6) (9), when the quality of life is not so impaired that a more aggressive treatment would be required. In the case of a high-disability CLBP, this approach can be used too, particularly, when the possibility or opportunity for a stronger treatment is absent.

**Conclusion**

Generally, all these rehabilitation tools should last several months in an outpatient setting, or must be intensive in an inpatient one for several weeks. This is the minimum needed to achieve a real change, to stop the progression of chronicity and revert it to some extent, and thereby return to a better quality of life. If specific causes of LBP are not known, a maintenance protocol would not be necessary; however, it is important to stop the treatment by reducing the medication of the patient, since he must have acquired effective tools with which to manage his problem. The best maintenance protocol is a physical activity, whether general or through specific machines (144), performed regularly a couple of times each week. Loss of maintenance in CLBP can drive to relapses of the problem (144).

Rehabilitative treatments have demonstrated their effectiveness in reducing disability and pain intensity. A precise evaluation of the patient’s features can allow the physician to choose the best rehabilitation tool in terms of cost/benefit ratio. Even if the most effective rehabilitation strategy is a multidisciplinary approach with cognitive behavioral treatment including specific exercises, there are cases in which other strategies can be more effective, or perhaps equally effective but less expensive.

**Secondary LBP**

This section considers the spinal diseases that are recognized as causes of LBP. It must be clearly stated that there is a big difference between imaging diagnosis and clinical diagnosis. According to the actual literature, the clinical diagnosis of secondary LBP should be defined when there are symptoms and signs confirming that the imaging is meaningful and strictly related to the clinical picture (3,8,9,46,52). If this is not the case, the best interpretation of the imaging is the presence of specific risk factors for the problems. In contrast, the worst interpretation of the imaging finding is giving to the patient a “diagnostic label” (165,166) that can propel him toward chronicization.

**Sciatica and DH**

As with LBP, the approach to sciatica has also changed with time: if at the clinical evaluation there are no neurologic signs (weakness, ipo/anesthesia, loss of reflexes, radicular impingement), only subacute sciatica is managed differently from LBP (9) (see Fig. 33-2). Consequently, this paragraph will consider subacute sciatica and DH (Figs. 33-6 and 33-7).

**Definition and Pathogenesis**

Sciatica is generally defined as a pain in the lower back and hip, radiating in the distribution of the sciatic nerve (258). It affects many people, and the annual prevalence of disc-related sciatica in the general population is estimated at 2.2% (259).

Sciatica, in the majority of cases, is caused by irritation of the sensory root or dorsal root ganglion of a spinal nerve. A conflict between the intervertebral disk and the nearby neural structures is the origin of irritation, but lumbar stenosis and tumors are other possible causes (46). When irritation arises, it causes ectopic nerve impulses, which are perceived as pain in the distribution of the axon.

There are two distinctive but not mutually exclusive pathomechanisms for neural irritation. First, the nerve roots, which are subject to sustained compression for long periods, may become sensitized to mechanical stimulation and show pathological changes such as focal demyelination, intraneuronal edema, Wallerian degeneration, and axonal damage.
Clinical Presentation
The typical clinical picture of sciatica is radicular pain: a deep, severe pain that starts low on one side of the back and then shoots down the buttock and leg when certain movements are attempted. Radicular pain is perceived in the territory innervated by the affected nerve root in the lower extremity when L4/5/S1 nerve roots are involved, or in the anterior thigh when L2/3 are involved. Pain does not always follow the corresponding dermatome, so it is sensory loss that indicates the affected segment.

Usually, lumbar radicular pain travels through the lower limb along a narrow band. It is perceived as sharp, shooting, or lancinating. It can be experienced superficially and deeply, worsens with coughing and other maneuvers that increase the pressure around the nerve, and radiates to the foot.

The pain is usually worse with prolonged sitting and standing. Bending backward can also increase the pain.

The lumbar spine may present an altered range of motion with limited forward flexion and spasm of the paraspinal muscles. Occasionally, the somatic pain can exist as a dull, aching pain.

If axonal damage is severe, weakness in the leg or the foot may occur, leading to neurogenic claudication. When there is bladder dysfunction with urinary retention or overflow incontinence, saddle anesthesia, and unilateral or bilateral leg pain and weakness, a cauda equina syndrome is present.

Diagnosis, Classification, and Prognosis
Sciatica is mainly diagnosed through history-taking and physical examination (see Figs. 33-2, 33-5, and 33-7). By definition, the patient mentions radiating pain in the leg and may also report sensory symptoms. Occasionally, the patient will also have LBP, but this is usually less severe than the leg pain.

The first task of medical history and the physical examination is to rule out the suspicions of tumor, infection, significant trauma, or dangerous nonspinal pathology mimicking a back problem. The red flags are a list of relevant risk factors and signs by which an underlying severe disease (262) can be suspected.

A severe neurologic compromise, limping or coordination problems, severe guarding of lumbar motion on all planes, vertebral point tenderness to palpation, or spinal-cord dysfunction are the most relevant findings in the physical examination, which may suggest relevant underlying diseases.

The neurologic examination can focus on a few tests that seek evidence of nerve-root impairment, peripheral neuropathy, or spinal-cord dysfunction.

More than 90% of all clinically significant lower extremity radiculopathy is due to DH and involves the L5 or S1 nerve root at the L4-5 or L5-S1 disc level.

The neurologic examination includes sensory examination, testing for muscle strength and trophism, reflexes, and clinical tests for sciatic tension.

Sensory examination is performed by testing with a light touch or pressure in the medial (L4), dorsal (L5), and lateral (S1) aspects of the foot. Muscle strength is tested through the patient’s ability to toe-walk (calf muscles, mostly the S1 nerve root), heel-walk (ankle and toe dorsiflexor muscles, L5 and some L4 nerve roots), or perform a single squat-and-rise (quadriceps muscles, mostly L4 nerve root). Muscle atrophy can be detected through circumferential measurements of the calf and thigh bilaterally, and the difference may be significant when there is a difference of more than 2 cm in the measurements of the two limbs at the same level.

Reflexes may test the S1 nerve root (ankle jerk reflex) or the L4 nerve root (knee jerk reflex). Up-going toes, in response to stroking the plantar footpad (Babinski or plantar response), may indicate upper motor neuron abnormalities (such as myelopathy or demyelinating disease).

The SLR test can detect tension in the L5 and/or S1 nerve root. SLR may reproduce leg pain by stretching the nerve roots. Pain below the knee at less than 70 degrees of SLR, aggravated by dorsiflexion of the ankle and relieved by ankle plantar flexion or external limb rotation, is suggestive of tension in the L5 or S1 nerve root related to DH.

Because the sensitivity of the SLR test is estimated as 91% with a corresponding specificity of 26%, if a patient reports the typical radiating pain in one leg combined with a positive result on one or more neurologic tests indicating nerve root tension or neurologic deficit, the diagnosis of sciatica is justified (46).

Regarding acute sciatica, imaging may be indicated if the results influence further management or infer the suspicion of an underlying disease (infections, malignancies) rather than...
DH. Diagnostic imaging may also be indicated in patients with severe symptoms who fail to respond to conservative care after 6 to 8 weeks (46).

The clinical course of acute sciatica is favorable and most pain and related disability resolves within few weeks. Therefore, in most cases the prognosis is good, but at the same time up to 30% of patients continue to have pain for at least a year (46).

Surgical Treatment: Indication and Limits
The broad consensus is that a cauda equina syndrome is an absolute indication for immediate surgery, but the surgical treatment of sciatica is still a matter of debate. Its goal is to remove DH and eventually part of the disc or foraminal stenosis, thereby eliminating the suspected cause of the sciatica. Surgery may relieve leg pain, but it has no effect on LBP.

Recently, an extensive trial including 1,244 patients compared surgery to medical treatment: 5/01 patients participated in a RCT and 743 in a prospective cohort study in which patients chose either surgery or medical treatment (263,264).

Unfortunately, in the randomized controlled trial, adherence to the assigned treatment was limited. Within the first 3 months 30% of patients in the medical group had undergone surgery, whereas 50% of patients in the surgical group had improved to such an extent that surgery was not performed. Consequently, no definitive conclusion can be drawn from the study, since it merely suggests that surgery may effectively relieve the pain without affecting the overall outcome (263).

In the observational cohort study, patients who chose surgery had better outcomes at 3 and 12 months than those in the medical-treatment group for all the evaluation criteria, despite the fact that they had more factors of adverse prognostic significance than the medical-treatment group (264). Even in this case a definitive conclusion cannot be drawn due to the limitations of this uncontrolled study.

Finally, a recent updated Cochrane review on surgical interventions for lumbar disc prolapse has stated that surgical discectomy for carefully selected patients with sciatica due to lumbar disc prolapse provides more rapid relief from the acute attack than conservative management, although any positive or negative effect on the lifetime natural history of the underlying disc disease remain unclear (265).

Surgery, in conclusion, appears more suitable for the small percentage of patients with uncontrolled pain or deteriorating neurologic symptoms.

Rehabilitative Approach
DH was, for years, thought of as a disease of surgical interest. Only recently has proof been accumulated to show that spontaneous recovery is common, many times with improvement of the DH itself (266–269). This has also been proved in the context of specific approaches to rehabilitation (270). In the meantime, proof has accumulated that DH is also quite common in asymptomatic people (266,269), leading to a different understanding of the sciatica symptom: Nowadays an imaging technique is not considered useful in the first month of symptoms (acute sciatica) if the neurologic symptoms are mild or absent, because the spontaneous evolution can be rapid and in most cases auto-resolving (8,9,46). This means it is necessary to understand which kind of rehabilitation program could be most useful in such patients, and now some studies are underway regarding specific conservative and rehabilitative treatments (62,270,271) (Table 33-7).

PO Approaches
PO treatment is certainly a first choice in these cases, where usually we confront important neurologic pain. According to the actual evidence, we have some proof of efficacy for NSAIDs (97) and short periods of bed rest (107). On the other hand, even without the existence today of any real evidence, manual therapies and manipulation, physical therapies, bed rest for

<table>
<thead>
<tr>
<th>Table 33.7 DH Treatment (9)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Evidence</strong></td>
</tr>
<tr>
<td>A Counseling</td>
</tr>
<tr>
<td>A Counseling</td>
</tr>
<tr>
<td>A Counseling</td>
</tr>
<tr>
<td>A Counseling</td>
</tr>
<tr>
<td>A Counseling</td>
</tr>
<tr>
<td>A Work and ADL Interventions</td>
</tr>
<tr>
<td>A Work and ADL Interventions</td>
</tr>
<tr>
<td>A Work and ADL Interventions</td>
</tr>
<tr>
<td>A Work and ADL Interventions</td>
</tr>
<tr>
<td>A Work and ADL Interventions</td>
</tr>
<tr>
<td>B Physical Activities Counseling</td>
</tr>
<tr>
<td>B Physical Activities Counseling</td>
</tr>
<tr>
<td>B Physical Activities Counseling</td>
</tr>
<tr>
<td>B Physical Activities Counseling</td>
</tr>
<tr>
<td>B Physical Activities Counseling</td>
</tr>
<tr>
<td>B Physical Activities Counseling</td>
</tr>
<tr>
<td>B Physical Activities Counseling</td>
</tr>
<tr>
<td>B Physical Activities Counseling</td>
</tr>
<tr>
<td>B Physical Activities Counseling</td>
</tr>
<tr>
<td>B Physical Activities Counseling</td>
</tr>
</tbody>
</table>

Notes
1. Professional risk: load mobilization, trunk movements, vibrations.
2. Only one short-time treatment, not repeated.
4. Follow the specific indications of each treatment.
6. Choice recommended according to cost-effectiveness, patient preferences, disponibility, and previous results.
long periods of time, and steroids are commonly employed. The choice of the type of PO treatment in these cases should follow a progression to be chosen according to the individual risk/benefit ratio and the patient’s preferences.

**Anti-inflammatory Treatment**

Anti-inflammatory treatment could have a crucial role in DH. This role could go beyond what is usually thought, even with the future discovery of major point that could raise some caution issues. In the past the main pathoanatomical hypothesis for sciatica in the context of a DH was compression, but in the last decade the inflammatory hypothesis has gradually become more significant and today appears to prevail (272,273). Consequently, the research on new drugs is flourishing and could in the near future lead to new advancements. Also, we know the resorption of DH is highest in the worst situations, such as in big DHs and sequestered fragments, both in terms of natural history (268) and as a consequence of treatment (270). Therefore, even if these are usually the most painful situations and have the most frequent neurologic involvement, if the latter is not so important that it motivates surgery, the prognosis is the best one. What must be considered in this context is that DH resorption is presumably led by inflammation through the activation of macrophages (273–277). If this be the case, what is the true role of an anti-inflammatory treatment? On one hand, it could help reduce the immediate neurologic damage, but on the other it could also delay or even impede the spontaneous process of DH resorption (278–280). Obviously, these points will be considered in future research, and today the choice of such a treatment should be based on what we know, namely the pain-killing function of NSAIDS. Other more specific anti-inflammatory drugs are still being researched (e.g., infliximab (274,281)) or lack proof of efficacy, such as steroids both in general administration and locally through infiltration (92,198,282,283).

**Rehabilitation and/or PO/ Anti-inflammatory Exercises**

Considering the various types of rehabilitation strategies, again we must think about what they really do. In fact, there are some exercise treatments that are proposed mainly in terms of a PO approach (63), while other strategies such as education and cognitive behavioral support have broader objectives. In any case, confronting with a DH and an acute sciatica is a disease-oriented treatment more than a biopsychosocial rehabilitative approach. Accordingly, a conservative treatment spine expert could be the specialist to consider, and these many times are PRM doctors. Therefore, it relates to our field, but we have to understand what we are really doing, meaning whether it is a FO rehabilitation or a PO treatment.

Research on rehabilitation is in the earliest stage, and recently a systematic review was concluded with the absence of any real evidence (282). When compared to surgery, rehabilitation has almost always been proposed in terms of “standard care” (263–265), an expression that means nothing more than “whatever a more or less expert physiotherapist wanted/ was able to offer.” The same occurred in a RCT comparing physiotherapy to bed rest and the continuation of ADL, which showed that neither of the two interventions had any effect (284). We cannot rely on such methodologies as the means to understand something real. In a retrospective review of patients with good results, mobilization and general exercises have shown better efficacy than spasm-reduction interventions (285). In the literature we cannot find much data, due to the fact that this field of research is very young.

The McKenzie approach (61–63,286) is a physiotherapy method that theoretically focused on DH or, better, on disk behavior as the possible source of all LBP problems. The main merits of this method include the discovery of the importance of exercise choice (62), and today it is used by others in their own development of subclassification methods (121,123). This technique has not yet provided definitive answers on its efficacy (63), but it must be looked carefully as a possible pain-killing tool with the advantage of being administered by a physiotherapist who could also offer other rehabilitative advice. Obviously cost is the main limitation, as compared to other PO treatments.

According to the DH pathophysiology and natural history, the first objective of rehabilitation—when facing an acute sciatica and/or a DH trying to avoid surgery—should be to help spontaneous resorption, and if possible to increase the speed of this process. On the other hand, one must avoid increasing pain and the problem itself. In this respect movement is known as a means to increase bloodflow: consequently, whichever movement that does not increase pain or tissue damage can be useful to accelerate the wash-out of the inflammatory and proteolytic enzymes while increasing oxygenation, nutrient caption, and the arrival of macrophages, thereby reducing inflammation and facilitating a rapid DH resorption. Accordingly, the McKenzie approach, more than being a way to mechanically drive the disk back into position, could be a way to mobilize the spine and symptomatic disk without causing an increase in pain. Therefore, it could facilitate an immediate wash-out of inflammatory agents and a progressive reduction of the pathological situation. If this is true, it could explain why the physiotherapy approach that showed the best results in a retrospective analysis of sciatica treatment was joint mobilization (285).

Setting aside any other consideration, central to a rehabilitation approach to DH is education and a cognitive behavioral counseling with the goal of encouraging DH resorption and reducing the risk of increased symptoms, but also of chronicization and ultimately increasing the patient’s self-care.
Lumbar Spinal Stenosis

Definition and Pathogenesis

LSS is defined as any type of narrowing of the lumbar spinal canal, causing compression of its content (Fig. 33-8). This narrowing causes direct mechanical compression on the neural elements or on their blood supply, which may lead to symptoms (289,290). The symptoms can decrease the patient’s quality of life and cause them to seek treatment.

LSS may occur at different places in the spinal canal, sometimes in more than one location at the same time. In central canal stenosis, the nerve roots in the cauda equina may be compressed. Lateral recess stenosis and foraminal stenosis may cause compression of the nerve roots leaving the spine (291,292).

The presence of a narrow canal in radiographic imaging does not define the syndrome. Instead, a diagnosis of LSS is defined by symptoms and clinical findings that must be supported by radiographic evidence.

LSS results from degenerative, developmental, or congenital disorders. The degenerative type is often due to arthritic changes: Disc degeneration, facet degeneration and hypertrophy, degenerative spondylolisthesis (SL) and ligamentum flavum hypertrophy, and calcification can cause LSS alone or in combination, or can further compromise the space of canal that is already small (293,294). The degenerative type occurs most often, particularly in patients over 50 years of age (290,291,294,295). The congenital type may occur earlier in life, and LSS is a result of congenitally anatomic changes or malformations, for example, an excessive scoliosis or excessive lordotic curvature (289,290,295). Developmental LSS is a condition in which the narrow spinal canal is caused by a growth disturbance of the posterior elements (296).

Clinical Presentation

LSS patients frequently present with few objective physical findings. Approximately 65% of patients have decreased walking ability, but up to 95% of patients treated surgically have only subjective symptoms, chiefly pain (297–299).

Patients with LSS report three types of complaints: LBP, leg symptoms, and neurogenic claudication. LBP is indistinguishable from nonspecific LBP. Leg symptoms are due to a radiculopathy with an aching or sharp pain and/or symptoms of pain, burning, numbness, and paresthesia following a specific dermatomal distribution in one lower extremity. The fifth lumbar nerve root associated with L5 stenosis is most often involved in this type of LSS (300).

Regarding neurogenic claudication, patients complain of cramping, weakness, pain, numbness, and tingling. These do not follow a specific dermatomal distribution but are most often experienced in both lower extremities. Occasionally, the patient will complain of bladder dysfunction and sexual difficulties. In LSS with less dermatomally specific neurogenic claudication, nerve roots below L5 are most commonly involved (300). In several patients, radiculopathy and neurogenic claudication coexist in a mixed type of LSS (300,301). Full-blown cauda equina syndrome only occurs in rare instances (140).

Because both the central canal and foraminal dimensions increase in flexion and diminish in extension, patients with LSS experience the exacerbation of symptoms with extension and improvement with flexion, meaning the symptoms are provoked by the act of standing but relieved when the patient is seated.

Diagnosis, Classification, and Prognosis

The tendency in diagnosing LSS is to focus on imaging studies (Fig. 33-9). However, it has been reported that in approximately 30% of asymptomatic subjects lumbar spinal abnormalities can be seen on imaging studies (302–304). In asymptomatic individuals, 60-years-old or more, LSS was detected on MRI in 21% of cases by Boden et al. (303). In a study by Jensen et al the percentages of spinal abnormalities over 30% were reported (269).

Due to the discrepancies between clinical symptoms and imaging findings (296,305), there is a difference between the
clinical and radiologic diagnosis of LSS. Therefore, clinical and radiologic findings should be considered together when diagnosing this disease.

Regarding patient history, the key factors that seem most strongly associated with the diagnosis of LSS are a higher age, severe pain in the lower extremities, and the absence of pain when seated (306). The physical findings most strongly associated with the diagnosis were a wide-based gait, an abnormal Romberg test, thigh pain following 30 seconds of lumbar extension, and neuromuscular deficit (306). Treadmill walking is also useful for diagnosis: a combination of time to onset of symptoms and recovery time after treadmill walking, or a longer walking time during inclined treadmill walking are suggestive of LSS (307). A combination of these key diagnostic factors has been proposed as a diagnostic algorithm for clinical practice to assess the likelihood of LSS, to guide and facilitate the clinical decision-making process (306,308–310).

Once the diagnosis of LSS is suspected due to history and clinical findings, the anatomical structures can be evaluated through CT or MRI, which, in a recent meta-analysis, seem to have comparable accuracy. Myelography should be avoided due to its invasiveness and lack of superior accuracy versus CT or MRI (307).

Both CT and MRI allow a precise measurement of canal diameter. An anterior-posterior (AP) canal diameter of less than 10 mm constitutes absolute stenosis, whereas an upper limit of less than 13 mm denotes relative stenosis (311–313).

According to a recent study, patients with smaller canals report greater perceived disability, but AP spinal canal diameter was not significantly associated with other clinical symptoms (314).

LSS, in radiological terms, has been classified by the location of the stenosis into central, lateral recess, and foraminal stenosis. Theoretically, central stenosis gives rise to cauda equina compression and neurogenic claudication, whereas lateral recess (or foraminal) stenosis is associated with radiculopathy.

Electromyography and nerve conduction studies may be useful in the diagnosis of LSS because fibrillation potentials could discriminate clinical LSS from other disorders with excellent specificity, and could detect alternative or comorbid disorders (315).

The few studies of the natural history of LSS generally seem to indicate a relatively benign course (316–320). Anatomy is seldom related to the severity of symptoms and probably other factors come into play. Segmental hypermobility, local neurovascular compromise, venous congestion, foraminal stenosis, or facet joint effusion may explain pain exacerbation and disability. Generally, LSS is a fluctuating syndrome with an overall improvement, and it is in a continuum with LBP and the absence of symptoms.

**Surgical Treatment: Indication and Limits**

LSS is the most common reason requiring lumbar spine surgery in adults older than 65 years (321). The patient’s subjective experience of disablement must be an important part of the indication for surgery, since anatomic and neurologic deficits do not predict future function (see Fig. 33-9).

The aim of the operation is to improve the quality of life, and surgical candidates include patients who have persistent severe leg symptoms and functional limitations.

Different studies have suggested the benefit of surgery in selected candidates. Atlas et al reported a greater improvement in surgically treated patients compared to nonsurgically treated patients in both the 1- and 4-year evaluations (322,323). Similar results were reported in a prospective 10-year study, in which considerably better treatment results were found in a group of patients randomized for surgical treatment compared to a group randomized for conservative treatment (324).
Additionally, a meta-analysis of the literature showed that an average of 64% of patients treated surgically for LSS were reported to have good to excellent outcomes. Nevertheless, it is important to note that the reoperation rate and failures are relatively high (325); moreover, surgery is only valuable for leg pain while it does not necessarily relieve LBP (283,325,326). Depression, cardiovascular comorbidity, disorder influencing walking ability, and scoliosis are predictors of poorer subjective outcome after surgery, and they may be considered in the decision-making process for surgical indication (327).

Less aggressive surgical techniques that provide for adequate decompression have recently been reported. These procedures have been described as fenestration, laminotomy, selective decompression, and laminar resection, and are purported to improve postoperative morbidity, provide early mobility, and reduce the hospital stay. However, their long-term impact on disability has not been determined.

Rehabilitative Approach
According to the earlier definition, when a patient has an imaging diagnosis of LSS without any clinical sign, he/she should be considered as carrying a specific risk factor for LBP; but nothing more. In the literature there are studies comparing surgery to the conservative approach, but they are mainly observational, and cross-over between groups due to patient’s choice has been quite common. Usually, in the literature the tested rehabilitation programs are not well described (317,322–324). Generally, it is possible to find some “expert opinions” that, in any case, are not in full agreement with one another (322,328–330).

We can consider two main possible exercise approaches to LSS (Table 33-8): One is aimed at directly facing the pathology and trying to achieve improvements with treatment; the other focuses on allowing better management of an already stable situation, even if the possible future increase of problems can be the long-term result. The first protocol should be pursued whenever possible, while the second can be considered when the first one fails.

LSS is partly due to new bone formation and partly due to the thickening of soft tissues (292,311). Only in the latter is it possible to try to allow better bloodflow while walking and avoid (or delay) the development of symptoms. In this respect the elasticity of soft tissues is the main objective, along with guaranteeing the best physiology through appropriate movement. On one hand, this means manual therapy, mobilization techniques (328), and exercises to increase the range of motion, while on the other hand it means neuromotor control, proprioception, strength, and endurance training. In the meantime, achieving a good capability of counteracting gravity force so as to slightly elongate the lumbar spine and provide more space for neural tissues can be another goal of exercises. This can be achieved through postural education and an increased endurance of deep stabilizing muscles, as well as with various activities in elongation. In this regard, specific muscular impairments have been shown in LSS patients (315,331), and these could definitively contribute to the symptoms. In this exercise approach, it is also important to educate the patient in maintaining a good erect posture while walking, because we are trying to renormalize the pathoanatomic situation as much as possible and avoid a progressively forward-bent posture, which is quite typical of long-standing LSS patients (316).

The other possible program is the one most discussed in the literature (322,328,330), and should be proposed in the worst cases and/or if the previous one fails. By contrast to the previous program, this one aims to increase the gait autonomy by teaching the patient to move in flexion, because flexion increases the space for bloodflow and can allow one to move better than extension in LSS (296,332). This program has been evaluated in the literature and short-term results have been presented (333). Nevertheless, in this case we can ultimately facilitate the progressive flexion of the spine, or we can with time reduce autonomy and require that the patient use a cane to move from one place to the other.

Apart from these specific programs, aerobic training can increase the peripheral oxygen uptake, thereby reducing symptoms and increasing the patient’s maximum walking distance (330,334). Another alternative procedure, or eventually a first step toward achieving the programs previously described, can include body-weight support treadmill training (329). Additionally, water-based exercises and traction have been proposed, as well as neural mobilization techniques (328) despite the absence of evidence in the literature. Orthosis and lumbar supports have been proposed, but again no data exist in the literature (328).

As is always the case with LBP patients, counseling is a crucial point (9). Key messages should include knowledge of the pathology, its main symptoms, and its prognosis in terms of pain and the ADL. Particularly, the possible evolution in a flexed posture (FP) must be explained with the aim of avoiding it. Moreover, pain control and prevention are crucial, as are strategies for pain management. ADL can be adapted according to the need, and the use of a bicycle instead of walking is suggested. Or, when needed, the use of a cane could be proposed.

Lumbar Spine Instability and Spondylolisthesis
Definition and Pathogenesis
Instability of the lumbar spine is a controversial topic. Despite several efforts, an accepted definition of lumbar spinal instability (LSI) is not yet available.

A reasonable definition has been proposed by Pope and Panjabi (335) and Frymoyer and Selby (336). Instability can be defined as an abnormal response to applied loads, characterized kinematically by abnormal movement in the motion segment beyond normal constraints (336). This abnormal movement can be explained in terms of damage to the restraining structures (i.e., facet joints, disks, ligaments, and muscles) that, if damaged or lax, will encourage altered equilibrium and thus instability (335).

LSI is considered to represent one of the potential conditions causing nonspecific LBP (335).

Traditionally, the occurrence of SL in subjects with CLBP has been considered one of the most obvious manifestations
### TABLE 33.8 Secondary LBP (Spinal Stenosis, Spinal Instability, and AS) Treatment (9)

<table>
<thead>
<tr>
<th>Spinal Stenosis</th>
<th>Spinal Instability</th>
<th>Adult Scoliosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evidence Content</td>
<td>Evidence Content</td>
<td>Evidence Content</td>
</tr>
<tr>
<td>C Counseling</td>
<td>C Counseling</td>
<td>C Counseling</td>
</tr>
<tr>
<td>Difficult resolution of spinal stenosis symptoms</td>
<td>Distinction between structural and neuromotor vertebral instability</td>
<td>A scoliosis over 30 degrees can progress even during adulthood</td>
</tr>
<tr>
<td>Spinal stenosis natural history is not well known</td>
<td>Improvement of stabilization capacity can reduce pain</td>
<td>If the scoliosis already progressed, likely will keep on progressing</td>
</tr>
<tr>
<td>A progressive flexion of the spine is possible with time</td>
<td>During long time reactive artrosic rigidity gives positive prognosis</td>
<td>In the long term it is possible a forward flexion of scoliosis with difficulties in maintaining a normal posture in elder age</td>
</tr>
<tr>
<td>Control progressive flexion of the spine</td>
<td>Learn how to control and prevent pain</td>
<td>Aesthetics worsen with progression of scoliosis</td>
</tr>
<tr>
<td>Learn pain control and prevention</td>
<td>It is necessary to face not to undergo pain</td>
<td>Respiratory capacity must be regularly checked and cardiopulmonary apparatus should be constantly trained</td>
</tr>
<tr>
<td>Learn pain management</td>
<td></td>
<td>Exercises can help for pain and provide short-term improvements, but there is no evidence that they can stop progression in the long term</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Exercises must be continuous in time</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Learn pain control and prevention</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Learn pain management</td>
</tr>
<tr>
<td><strong>Work and ADL Interventions</strong></td>
<td><strong>Work and ADL Interventions</strong></td>
<td><strong>Work and ADL Interventions</strong></td>
</tr>
<tr>
<td>A Avoid long walks, use a bicycle</td>
<td>A Avoid excessive loads and repeated end of ROM reaching</td>
<td>B Avoid excessive loads</td>
</tr>
<tr>
<td><strong>Physical Activity Counseling</strong></td>
<td><strong>Physical Activity Counseling</strong></td>
<td><strong>Physical Activity Counseling</strong></td>
</tr>
<tr>
<td>B Mild aerobic activity without impact</td>
<td>B Mild aerobic activity without high-ROM and impact</td>
<td>Aerobic activity</td>
</tr>
<tr>
<td><strong>Painkiller Therapy</strong></td>
<td><strong>Painkiller Therapy</strong></td>
<td><strong>Painkiller Therapy</strong></td>
</tr>
<tr>
<td>A See chronic LBP</td>
<td>B See chronic LBP (Table 33-5), but avoid spinal mobilization and manipulation (1)</td>
<td>See chronic LBP (Table 33-5), but avoiding spinal mobilization and manipulation (1)</td>
</tr>
<tr>
<td><strong>Rehabilitation</strong></td>
<td><strong>Rehabilitation</strong></td>
<td><strong>Rehabilitation</strong></td>
</tr>
<tr>
<td>C ROM increasing exercises and progressively increased gait</td>
<td>C Regular and continuous stabilizing exercises (22)</td>
<td>Regular and continuous stabilizing exercises (1)</td>
</tr>
<tr>
<td>C Lumbar supports</td>
<td>C Lumbar supports</td>
<td>Lumbar supports</td>
</tr>
<tr>
<td>B Rigid orthosis</td>
<td>B Rigid orthosis, eventually</td>
<td>B Rigid orthosis</td>
</tr>
</tbody>
</table>

**Notes**
1. Mobilization means performing repeated manoeuvres till the end of ROM that imply an increase of ROM during time.
2. Stabilizing exercises increase spinal neuromotor control ability and are based on the improvement of: proprioception, kinesthesia, spinal coordination precise neuromotor control of movement, strengthening of stabilizing muscles (particularly multifidus and transversus).
of LSI (Fig. 33-10) (337,338). Furthermore, several studies (339–341) have reported increased and abnormal intersegmental motion in subjects with CLBP and in the absence of other radiological findings.

Clinical Presentation
Nonspecific LBP and sciatica are the most frequent clinical presentations of LSI and SL. Higher incidences of SL might be seen in people involved in repetitious alternate loading activities such as gymnastics, weightlifting, and football.

Numerous anamnestic and clinical findings are proposed to suggest instability: frequent recurrent episodes of LBP, short-term relief from manipulation, a history of trauma, improvement of symptoms with a rigid brace or external fixation, palpation for the presence of a “step-off” between adjacent spinous processes, aberrant motions such as an “instability catch,” or increased mobility with passive intervertebral motion testing. Generally, the validity of these findings has not been reported.

Recently, a questionnaire completed by patients diagnosed with LSI described back pain symptoms as “recurrent” (70%), “constant” (55%), “catching” (45%), “locking” (20%), “giving way” (20%), and/or “accompanied by a feeling of instability” (35%) (342).

Data by subjects diagnosed with LSI involved in recent clinical trials revealed that half of the subjects developed their LBP condition secondary to a single event injury while the other half developed LBP gradually in relation to multiple minor traumatic incidents (343).

In another questionnaire completed by physical therapists, identification of muscle dysfunction, motor control abnormalities, and loss of strength were the greatest indicators of clinical lumbar instability. The descriptors of poor lumbo-pelvic control include segmental hinging or pivoting with movement, muscle guarding/spasm, poor coordination/neuromuscular control, decreased strength and endurance of local muscles at the level of segmental instability were the top three component scores (342).

In another study the two most predictive factors of radiographic instability from the history and physical examination were lumbar flexion ROM and a lack of hypomobility during lumbar intervertebral motion testing. The presence of both findings increased the probability of instability from 50% to 93% (69).

In a recent study about the correlation between LSI and clinical symptoms, the translation of the lumbar segment had a greater influence than angulation on lumbar symptoms, and the combination of translation and angulation was associated with worse symptoms and the persistence of pain (344).

Diagnosis, Classification, and Prognosis
A systematic clinical history and physical examination should be performed to characterize the LBP, distinguish referred from radicular symptoms, document sagittal alignment and spinal mobility, and establish the presence of any neurologic deficit, paying particular attention to the function of the nerve roots exiting at the level of theolisthetic vertebrae (Fig. 33-11).

Spine instability may be classified into two categories: radiologic appreciable instability and clinical instability. Radiologic instability reflects marked disruption of passive osseoligamentous anatomical constraints, as in the case of spondilolysthesis. Clinical instability is more challenging to diagnose and may involve discrepancies in radiographic findings.

The imaging analysis begins with conventional radiology, with anteroposterior, lateral, and flexion-extension radiographs. A number of radiographic indices can be made from the standing lateral radiograph, including the degree of slippage, slip angle, sacral inclination, sacrohorizontal angle, and lumbar index. CT and MRI are useful advanced imaging methods, particularly in the preoperative planning stage, for better defining both the bony and soft tissue anatomy, respectively.
Knutsson originally described a method for diagnosing segmental instability using lateral radiographs with the patient performing maximum lumbar flexion and extension (345). The amount of sagittal plane translation and rotation occurring at individual spinal motion segments are calculated (346–348).

White and Panjabi defined criteria for diagnosing instability from flexion-extension radiographs as sagittal plane translation greater than 4.5 mm or greater than 15% of the vertebral body width, or sagittal plane rotation greater than 15 degrees at L1/L2, L2/L3 or L3/L4, greater than 20 degrees at L4/L5, or greater than 25 degrees at L5/S1 (349). Although concerns about the validity of flexion-extension radiographs have been raised (350,351), this method has become the standard by which lumbar instability is diagnosed.

At the lumbosacral junction, stability is dependent on the spatial orientation of L5 to the sacrum, lumbosacral angle, sacral slope, and pelvic incidence as well as an intact osteo-discal-ligamentous complex. Because spinal parameters are dependent on pelvic parameters, alterations in local spatial orientation, such as SL, can produce global spinal imbalance. Some spondylolistheses progress to severe deformity, while others progress very little. The pelvic incidence seems to be one of the most important factors linked with a higher degree of slippage.

Given the aging population, degenerative SL is getting more and more common. It occurs in individuals older than 50 years, it occurs four times more frequently in women, and is most commonly seen at L4-5 (352–354). Rarely, however, does the slip exceed 25% to 30% of the width of the subjacent vertebra (355).

### Surgical Treatment: Indication and Limits

Management decisions for adults with LSI need to take into account the severity and duration of symptoms and patient comorbidities. Those patients with intolerable back and/or leg pain recalcitrant to a prolonged conservative treatment program could be candidates for surgery (see Fig. 33-11).

In patients with SL, progression of the slippage occurred only in about 30% of cases. Over 70% of the patients who were initially neurologically intact did not deteriorate over time and these patients may be treated conservatively. Conversely, most patients with a history of neurogenic claudication or vesicorectal symptoms deteriorated with poor final outcome and these patients should preferably have surgical treatment (353).

In case of SL, indications for surgical treatment are persistent or recurrent back and/or leg pain or neurogenic claudication, with significant reduction of quality of life, despite at least 3 months of nonoperative treatment, progressive neurologic deficit, or bladder or bowel symptoms (356).

Because there is no consensus as to what constitutes the optimal surgical treatment or what constitutes the more appropriate nonoperative program, the decision to recommend surgical treatment to an adult patient with LSI must be carefully individualized.

Spine fusion is the selected procedure for LSI. It was initially performed to stabilize vertebral fractures, spinal tuberculosis, and deformities such as scoliosis. In recent years, however, a majority has been performed for degenerative spinal conditions. The rationale for fusion is that it should reduce abnormal motion and, therefore, reduce pain. The efficacy of fusion in this context remains uncertain (357–359).

---

### FIGURE 33-11. Spinal instability and spondylolisthesis diagnostic flowchart [9].

<table>
<thead>
<tr>
<th>Evidence</th>
<th>Flow-chart</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Perform standing standard and dynamic x-rays</td>
</tr>
<tr>
<td></td>
<td>Is there a spondylolisthesis over 60%?</td>
</tr>
<tr>
<td>B</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Is there radiographic instability? (1)</td>
</tr>
<tr>
<td>B</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Conservative treatment</td>
</tr>
<tr>
<td>C</td>
<td>Back to previous flow-chart</td>
</tr>
<tr>
<td></td>
<td>Resolution of symptoms within 3 months?</td>
</tr>
<tr>
<td></td>
<td>Yes → Conservative treatment</td>
</tr>
<tr>
<td></td>
<td>No</td>
</tr>
</tbody>
</table>

| B        | Surgical evaluation |
|          | Does the patient prefer surgery? |
|          | Yes |
|          | No |
| A        | Is surgery possible? |
|          | Yes → Spinal fusion |
|          | No |

**Note**

1. Over 3 mm of mobility or over 10° of intervertebral angle.
Regarding SL, the mainstay of surgical treatment is decompression and there is no general agreement about the indications for fusion and instrumentation. The goals for decompression are to relieve radicular symptoms and neurogenic claudication, the goals for fusion are to relieve LBP by elimination of instability, and the goals for instrumentation are to promote fusion and to correct l isthesis or kyphotic deformity.

Two considerations rise doubts on the efficiency of surgery. First, the long-term outcome of surgery is uncertain: most studies are predominantly retrospective and most certainly describe a wide variety of surgical indications and surgical techniques. Second, the reoperation rate: in a recent study among over 24,000 patients after lumbar surgery the 11-year cumulative incidence of reoperation was 19.0% with a higher incidence among patients whose initial operation involved a fusion (360).

Rehabilitative Approach

An important contribution to understand clinical instability was provided by Panjabi (361), who proposed the existence of three separate but interrelated subsystem that act to control intersegmental spine stability: the passive subsystem (e.g., ligaments, vertebral disks), the active subsystem (musculature), and the neural subsystem (e.g., sensory receptors, cortical, and subcortical controls). Instability can result from a deficit in one of these components and, according to Panjabi’s (362) definition it consists in a significant decrease in the capacity of the stabilizing system to maintain the intervertebral neutral zones within the physiological limits. The intervertebral neural zone (362) is the zone, within the “range of motion,” of high flexibility, laxity within which the spinal motion is produced with minimal internal resistance. The components of the stabilizing system are functionally (363) interdependent so that compensation for system dysfunction may occur. Instability could be a result of tissue damage, making the segment more difficult to stabilize, insufficient muscular strength or endurance, or poor muscular control; instability is usually a combination of all three. This is the ratio for rehabilitation program in the patient with spinal instability due whether to a functional, neuromuscular instability, or SL (see Table 33-8). The major source of spinal stability and the component that can be enhanced from a rehabilitation approach is the second one (364), meaning the musculature, which appears particularly effective in the neutral zone, when passive structures assume only a minor role. Trunk stabilizing muscles are also defined as a core stabilizing system that is described (364) like a box with the diaphragm on top, pelvic floor on bottom, abdominals in front, and paraspinal and gluteal muscles in back. Core musculature is required for the spine (365) to move freely throughout its entire range of motion, and it also serves as a functional center of the kinetic chain by connecting the upper and lower extremities. This musculature is divided in the local deep unisegmental muscle (366) (transversus abdominis and multifidus) whose function consists mainly in stabilization, and global, superficial, multisegmental muscles (external oblique, erector spinae, rectus abdominis, psoas) intended mainly to produce control movement of the trunk. Moreover, in healthy individuals, fibers of the multifidus and transversus abdominis are the first fibers to become active when a limb is moved in response to visual stimulus (367), and the timing of superficial fiber activation depends also on the direction the limb is moved to assist with control of spinal orientation. Nevertheless, the neuromuscular subsystem coordinates muscle activity to respond to both expected and unexpected forces. This system must activate the correct muscle at the right time (368) by the right amount to protect the spine from injury and also allow the desired movement. Poor neuromuscular control may explain the onset of ALBP (363,369), even though there is no external load. Pain may be due to an adjustment mistake by one or more muscle in terms of magnitude and timing of responses caused by faulty afferent information.

In the acute phase, the first aim consists in providing a good control of pain (370) by pharmacological treatment or physical modalities and some easy exercises to minimize pain, avoiding spinal manipulation or traction; according to cognitive behavioral approach, physician will explain what is happening to the spine (distinguishing between a mechanical instability, due for example to a SL or a neuromuscular one; informing patient that the degenerative SL natural history shows an improvement as the disc space collapses and progression of the slippage occurred only in 34% of the cases), what patient can do to manage pain instead of undergo it. It is also useful teaching to avoid heavy and sudden loads, and extreme flexion and extension movement of the spine; patients with SL must know the risk factors of worsening (like hyperextension movements, high level physical activity), and main problems that may present like nerve-root pain, neurologic claudication. An aerobic conditioning will be encouraged, such as swimming, walking, or stationary bicycling that promotes spinal flexion and avoid wear and tear associated with impact aerobic exercises such as running (370). Lumbar support may be used in the acute phase only a few hours per day in order to help spine to support itself, but with the beginning of stabilizing exercises it will gradually be removed. A rigid brace (371), in selected cases of SL, can provide a way to control worsening of listhesis and sometimes also a reduction.

In the chronic phase, treatment consists of an intensive continuous stabilizing program (Fig. 33-12) with a deep cognitive behavioral intervention. Lumbar stabilization exercises aim (363) at a sensorimotor reprogramming of the spine stabilizer muscle intended to improve their motor control skills and delay of response, and consequently to compensate for weakness of the passive stabilization system. Stiffness is achieved with specific patterns of muscle activity, which differ depending on the position of the joint and the load on the spine. Core muscle endurance seems to be more important than total strength (67). This is due to the fact that in normal circumstances only a small amount (approximately 10% of maximal contraction) is needed to provide segmental stability. The first phase of the exercise program previews specific
isometric transversus abdominis-multifidus co-contractions (364), while maintaining the spine in a static neutral position. Exercises are based on the drawing-in action, called abdominal hollowing, presented by Richardson (364) who recommends performing exercises for 10 minutes ten times a day, while maintaining a normal breath. It can be performed by abdominal hollowing in a prone position, eventually with diagonal elevation of the arm and leg, or in a four-point kneeling position, which facilitates maintenance of neutral lordosis and a sensation of transversus abdominis muscle contraction. Therapist will use techniques of facilitation such as muscular palpation or visual feedback. The second stage requires co-contraction in situation where patients feel “unstable,” experience or anticipate pain. The third stage of training includes functional demands of daily requiring a low degree of attention for adequate contraction. However, exercises can then progress from training isolated muscle to training the core as an integrated unit to facilitate functional activity. The neutral spine, considered like a pain-free position that is touted as the position of power and balance, has been advocated as a safe place to begin exercises. Neuromuscular control can be enhanced through a combination of joint stability (co-contraction) exercises, balance training, perturbation (proprioceptive) training, polymeric (jump) and also via wobble boards, roller boards, and physio-ball. O’Sullivan et al. (372) found that individuals with CLBP and a radiological diagnosis of SL, who underwent a 10-week specific exercise program, showed a statistically significant reduction in pain intensity and functional disability levels, which was maintained at 30-month follow-up.

**Vertebral Deformities in Adults**

**Definition and Pathogenesis**

The ageing spine might develop two characteristic deformities: scoliosis and FP. Scoliosis in the adult is a disorder that involves a convergence of deformity and degenerative disease in the spine. It is defined as a spinal deformity in a skeletally mature patient with a Cobb angle of more than 10 degrees in the coronal plain (373). Independently from this definition, scoliosis can worsen in adulthood when it exceeds 30 degrees Cobb, and this evolution is almost certainly more than 50 degrees (374–378). The pathomechanism in these cases of AS is quite predictable: The primum movens is an asymmetric load or degeneration. Asymmetric degeneration leads to increased asymmetric load and therefore to a progression of the degeneration and deformity. The progression of a curve might be further supported by osteoporosis, particularly in postmenopausal women. The destruction of facet joints, joint capsules, discs, and ligaments may create monosegmental or multisegmental instability and finally LSS. There is also a de novo scoliosis (373) directly caused by asymmetric degeneration in a previously straight spine. Even if quite painful, this situation usually does not lead to a wide increase in Cobb angle over time (379).

FP is defined by thoracic kyphosis, protrusion of the head, and in more severe cases, knee flexion. The pathophysiology of FP in the elderly is most likely multifactorial: low bone mineral density, vertebral fractures, intervertebral disk degeneration, and back extensor strength, which decreases with age, are the most frequently encountered factors (380).

**Clinical Presentation**

The patient with significant AS may be relatively asymptomatic or severely disabled by his or her deformity.

The most frequent clinical problem of AS is LBP (373,381,382). LBP at the site of the curve can be localized either at the apex, on the prominence/hump, or in its concavity. It can be combined with radicular leg pain. It can be the expression of a muscular fatigue or of a genuine mechanical instability.

When the lumbar curve is accompanied by the loss of lumbar lordosis, the overloaded and stressed paravertebral back muscles may become a source of diffuse, permanent muscular pain. The pain, however, is generally present when the patient is upright, especially when standing and sitting, presenting as a so-called axial LBP, and the patients often indicate that they can control their pain well when lying down flat or on their side, and when the spine is relieved of its axial load.

The second important symptom of AS is radicular pain and claudication symptoms when standing or walking (383). The radicular pain may be due to a localized compression or root traction whereas claudication may be due to single- or multiple-level LSS.

The third important clinical presentation is a real neurologic deficit, including individual roots, several roots, or the whole cauda equina with apparent bladder and rectal sphincter problems. However, an objective neurologic deficit is rare, and when present it is due to a significantly compromised
space in the spinal canal with relatively acute aggravation and decompensation.

The fourth relevant symptom or sign is a progression of curvature. It usually becomes relevant when the curve has reached a certain degree and/or when osteoporotic asymmetric collapse could be relevant as a contributor to such curvature. Once the curvature has reached a certain degree, the progression will automatically follow due to the axial mechanical overload of individual facet joints and/or osteoporotic vertebral bodies. In case of evolution, usually AS progressively drives the spine laterally and ultimately in anterior flexion in the elderly, causing the need for a cane when walking.

Concerning FP, the kyphotic changes of the spine may cause local pain for vertebral overloading, inappropriate stretching of the ligaments and muscles, as well as overstress on the paravertebral back muscles (380,384). Compensatory hyperlordosis in the lumbar spine can cause LBP and a forward bending posture has been suggested to be associated with sacral and pelvic pain. FP also alters the load distribution on vertebral joint surfaces, leading to an increased spondylosis and vertebral bodies deformities when osteoporosis is present. Patients with severe FP show greater depression, reduced motivation, muscle impairment with a weaker spine extensor and ankle plantar flexor, lower scores in the balance and gait tests, a slower and wider base of support, and diminished ability in the ADL (380).

**Diagnosis, Classification, and Prognosis**

The diagnosis of AS is made on conventional x-rays (Fig. 33-13). It is classified into four major groups (373):

*Type 1:* Primary degenerative scoliosis, mostly on the basis of a disc and/or facet joint arthritis, affecting those structures asymmetrically with predominantly LBP symptoms, often accompanied with signs of LSS or without such signs. These curvatures are often classified as *de novo* scoliosis.

*Type 2:* Idiopathic adolescent scoliosis of the thoracic and/or lumbar spine, which progresses in adult life and is usually combined with secondary degeneration and/or imbalance. Such patients may develop secondary degeneration and progression of the adjacent curve.

*Type 3a:* Secondary adult curvature in the context of an oblique pelvis, for example, due to a leg length discrepancy or hip pathology, or as a secondary curve in idiopathic; as neuromuscular and congenital scoliosis or as asymmetrical anomalies at the lumbosacral junction.

*Type 3b:* Secondary adult curvature in the context of a metabolic bone disease (mostly osteoporosis) combined with asymmetric arthritic disease and/or vertebral fractures.

In AS, sagittal balance seems to be the most important and reliable radiographic predictor of clinical health status, as patients with positive sagittal imbalance report worse self-assessment in pain, function, and self-image domains (384). Age also influences the natural history of AS. Thoracic kyphosis increases with age, whereas lumbar lordosis decreases, leading to a net effect trending toward positive global sagittal balance with advancing age (385,386).

Coronal imbalance seems to be associated with deterioration in pain and function scores for nonoperated patients, but only when it is greater than 4 cm (384).

Other radiographic parameters correlated with pain in AS were lateral vertebralolisthry, L3 and L4 endplate obliquity angles, lumbar lordosis, and thoracolumbar kyphosis (387).

Beyond the standard clinical examination, patients with symptomatic AS may sometimes require interventional radiological procedures, such as sequential discograms, facet blocks, and epidural blocks to evaluate pain source. Finally, in elderly people with degenerative scoliosis and symptoms of claudication, leg pain and multilevel LSS, motor evoked potentials (MEP) may be helpful to identify the level responsible for the clinical presentation.
Regarding FP, kyphosis and vertebral deformities could be evaluated on lateral conventional x-rays. Kyphosis is measured by the traditional Cobb angle where the angle is derived from the slope of the superior vertebral endplate of T4 and the inferior vertebral endplate of T9.

Intervening vertebral deformity could be quantified by morphometry, which compares the height of the anterior, middle, and posterior parts of each individual vertebral body. When one of these heights is more than 20% lower than one of the other heights, the occurrence of a vertebral fracture is assumed (388). Because the number and severity of vertebral fractures are both of prognostic value regarding pain, disability, and risk of further fractures, a cumulative semi-quantitative index of the number and severity of vertebral deformities occurring between T4 and L5 has been developed (389,390).

**Surgical Treatment: Indication and Limits**

In AS and FP, surgery is an extreme option when the nonsurgical measures have no effect or do not promise any relevant long-term improvement.

Planning the surgical procedure requires a clear understanding of the prominent symptoms or clinical signs, and it is also influenced by the patient’s general health, age, conditions of bone quality, and the patient’s expectations.

The possible surgical technique can be divided into posterior, anterior, or combined procedures. In all these procedures, a simple decompression or stabilization can be done, or both can be combined. In some cases, additional correction may be performed by osteotomies or by sequential segmental correction through instrumentation (373).

Spinal surgery in these conditions is very demanding because it frequently deals with fragile elderly with poor bone quality. In these cases, reoperation rate and complications are significantly above the average for spinal surgery (373,391,392).

**Rehabilitative Approach**

Any treatment of LBP caused by vertebral deformity in adults must address two essential elements (see Table 33-8): degenerative instability and the reduction or loss of the physiological curvature of the lumbar segment in the sagittal plane. In fact, findings from various studies (387,393) have confirmed that there is no direct correlation between the degree of pain, the Cobb’s angle of curvature, and the cause of scoliosis (degenerative, de novo, etc.). Schwab et al. (394), however, noted a statistically significant correlation between the angle of L3 and L4 in the frontal plane, the extent of lumbar lordosis measured from L1 to S1, and pain and disability. The discomfort felt by the patient increased with the widening of the spinal angle and the decrease in lumbar lordosis.

Scoliotic curves over 30 degrees Cobb at the end of skeletal maturity often evolve slowly and insidiously, involving both the anatomical structure of the curve and the functional status of the patient (376). This worsening seems to be a postural collapse that, at least initially, is not a real deformity because it is not structured. However, over time the permanent asymmetric load tends to modify the vertebral structure and can no longer be recovered. The development of curvature is accompanied in a linear way by an increase in CLBP and psychological suffering, and in the most serious cases by a reduction of cardiopulmonary function (395).

**Scoliosis Progression**

Scoliosis is the most disabling deformity of the spine among adults (375). In addition to the concern for present disabilities, there is also the awareness that, as time passes, the condition will probably grow worse. Furthermore, when the major curve is at the lumbar and thoracolumbar levels, besides the worsening of rotation and lateral curvature there is the risk of a collapse into a disabling total round back and flexed anterior or lateral posture.

We find in the literature a growing amount of data confirming the possibility that exercise alone can in some cases slow down the development of scoliotic curvature, not only in the child but also in the adult (Fig. 33-14) (396). The reduction of scoliotic curvature certainly does not indicate a reduction of deformity but a recovery from the postural collapse, which is present in the upright posture. From a study by Torell and Nachemson, there is evidence that in adolescents, regardless of the magnitude of curvature, the mean difference between a standing radiography and a supine one is 9 degrees Cobb (397). There is no data in the literature to indicate precisely what this difference is (Duval-Beaupère called it “postural collapse”) (398) in adult scoliotic patients. Probably the recovery of this collapse is the key with which to avoid any worsening of adult curvature. On the other hand, the functional, cosmetic, and psychosocial damage caused by scoliosis is directly proportional to the magnitude of curvature (399) so an initial improvement, followed by stability over time, must be considered a remarkable success in the therapeutic treatment of AS.

The goals at the neuromotor and biomechanical levels are the recovery of postural collapse, postural control, and vertebral stability. The following are included as objectives of rehabilitative treatment:

- Becoming aware of pathology consequences and recovery possibilities for postural collapse
• Muscular strengthening and vertebral stabilization, always done in auto-correction, that is, in the position of maximum postural collapse recovery
• Global improvement of patient’s function, even with a partial recovery of possible deficits in the joint range of motion and of muscular retractions, if present
• Development of balance
• Postural integration, which includes the neuromotor integration of correct postures and an ergonomic education program
• Functional improvement, with aerobic and respiratory exercises in the case of reduced cardiopulmonary function
• Cognitive behavioral approach, even in the absence of pain

Back Pain and AS
The literature on spinal pain and scoliosis in adults is fairly uniform. In adult scoliotic subjects, researchers have found that the incidence of lumbar pain was similar to that of subjects without vertebral deviation (400), but the prevalence was higher (375,392,396,401). This pain seems to be more frequent in women after pregnancy or after a period of spinal mechanical overload (402), even if researchers have found no risk of debilitating LBP in adult patients whose lumbar scoliosis has not been treated. Moreover, there is a similar rate of surgery for lumbar pain in patients, either with or without scoliosis (403). Even if the symptom of pain is the main cause of surgical treatment requests for stabilization purposes, its extent cannot be correlated to the magnitude of curvature (404). Instead, there is a significant relationship between pain and the magnitude of lumbar lordosis. In fact, the increase of pain and reduction in quality of life are directly proportional to the flattening of the lumbar curve (394). For this reason, in the treatment of an adult patient with scoliosis and persistent lumbar pain, one of our goals is to recover/maintain sagittal curves with particular attention to the research of a good lumbar lordosis.

Strength-endurance training exercises that emphasize the extension of the spine can be particularly useful. In any case, the three-dimensional nature of scoliotic deviation requires that we pay attention to the starting position, which should be chosen after performing several tests to find the one most appropriate for the patient.

The scoliotic patient, like all subjects who report CLBP, also tends to develop a progressive fear-avoidance behavior, that is, a growing reduction of his/her activities for the sake of avoiding pain. In the acute phase this fear-avoidance behavior, as exemplified by rest, claudication or stick usage, has a protective effect against pain, given the reduced stress on the recovering structure.

Consequently, this behavior can persist in order to avoid pain, but it can also cause a progressive “disuse syndrome.” Therefore, the treatment schedule for a patient who experiences CLBP must be organized from the cognitive behavioral perspective.

Back Pain and Hyperkyphosis
Hyperkyphosis resulting from Scheuermann’s disease, and from idiopathic or postural hyperkyphosis among adults, involve a pronounced rigidity of the thoracic section that leads to dysfunctional strain on the spinal segments retaining mobility, meaning the cervical spine and lumbar spine. While the mediodorsal spine benefits from a type of “natural” arthrodesis, which is just as efficient as surgery, the increased pressure on the transition areas of the lordotic curvatures above and below the thoracic kyphosis can cause localized stress and eventually lead the patient to experience cervical and LBP. Additionally, the apex of the kyphosis itself can at times become a locus of sharp pain that is difficult to manage. Rehabilitative measures should include exercises to control posture and recover function similar for those for nonspecific LBP.

CONCLUSION
Today’s approach to LBP requires deep knowledge of the actual evidence and the clinical specifics of the individual circumstance. A superficial, nonexpert approach to these problems is no longer acceptable. That is particularly the case in our specialty, because the evolution of knowledge in the literature, if not the expert consensus, indicates a PRM leadership role in the LBP field. In fact, excluding ALBP, a rehabilitative team approach eventually combined with PO treatments (but clearly distinguishing the two) is mandatory in most of the cases (21,22).

REFERENCES


PART III

    de Graaf I, Prak A, Bierma-Zeinstra S, et al. Diagnosis of lumbar spinal
    Beattie PF, Meyers SP, Stratford P, et al. Associations between patient
    Stucki G, Liang MH, Fossel AH, et al. Relative responsiveness of condi-
    Katz JN, Dalgas M, Stucki G, et al. Degenerative lumbar spinal stenosis:
    Rauschning W. Pathoanatomy of lumbar disc degeneration and stenosis.
    Ray C. Extensive lumbar decompression: patient selection and results.
    Pheasant HC, Dyck P. Failed lumbar disc surgery: cause, assessment,
    Jonsson B, Stromqvist B. Symptoms and signs in degeneration of the
    Ullrich CG, Binet EF, Sanecki MG, et al. Quantitative assessment of the
    Haig AJ, Geisser ME, Tong HC, et al. Electromyographic and magnetic
    Johnsson KE, Rosen I, Uden A. The natural course of lumbar spinal
    Mazanec DJ, Podichetty VK, Hsia A. Lumbar canal stenosis: start with
    Johnsson KE, Uden A, Rosen I. The effect of decompression on the
    Sterner, M, Nilsson, S, Haggkvist, A, et al. Lumbar canal stenosis:
    Knutsson F. The instability associated with disk degeneration in the
    Dvorak J, Panjabi MM, Novotny JE, et al. Clinical validation of func-
    Dupuis PR, Yong-Hing K, Cassidy JD, et al. Radiologic diagnosis of
    Suda V, Saitou M, Shibaaki K, et al. Gait analysis of patients with neu-
    Simotas AC, Dorey FJ, Hansraj KK, et al. Nonoperative treatment for
    Fritz JM, Delitto A, Welch WC, et al. Lumbar spinal stenosis: a review of
current concepts in evaluation, management, and outcome measure-
    Leinonen V, Maatta S, Taimela S, et al. Paraspinal muscle denervation,
    paradoxically good lumbar endurance, and an abnormal flexion-exten-
    Bosio KD, Greuber RD. Lumbar spinal stenosis: anatomy and patho-
    Verbiest H. A radicular syndrome from developmental narrowing of the
    Ulrich CG, Binet EF, Saneeck MG, et al. Quantitative assessment of the
115–114.
    Geisser ME, Haig AJ, Tong HC, et al. Spinal canal size and clinical symp-
2007;23(9):780–785.
    Haig AJ, Geisser ME, Tong HC, et al. Electromyographic and magnetic
resonance imaging to predict lumbar spinal stenosis, low-back pain, and
    Benoist M. The natural history of lumbar degenerative spinal stenosis.
    Herno A, Airaksinen O, Saari T, et al. Lumbar spinal stenosis: a matched-
    Johnsson KE, Rosen I, Uden A. The natural course of lumbar spinal
    Porter RW, Hibbert C, Evans C. The natural history of root entrapment
    Johnsson KE, Uden A, Rosen I. The effect of decompression on the
natural course of spinal stenosis: a comparison of surgically treated and
    Mazanec DJ, Poddichery VK, Hsia A. Lumbar canal stenosis: start with
    Atlas SJ, Keller RB, Robson D, et al. Surgical and nonsurgical man-
agement of lumbar spinal stenosis: four-year outcomes from the main


Scoliosis and Other Spinal Deformities

DEFINITION

One of the unique features that distinguishes humans from the animal kingdom is our bipedal nature. Quadrupedal ambulation is more stable, with built-in redundancies. However, by standing on our hind legs, we are able to see above many other vertebrates. Our arms and hands are freed to manipulate tools. The unique balance between flexibility and stability in the human spine helped to realize this evolutionary innovation. It is a wonder that this balance is maintained for a lifetime of wear and tear for the vast majority of the world’s population. When that balance is altered, the result is deformity of the spine in one or more of its planes of movement. These conditions have likely been with man since bipedal mobility began. Hippocrates stated that “there are many variations of curvatures of the spine, even in persons who are in good health, for it takes place from natural conformations and from habit, and the spine is liable to be bent from old age and pains” (1).

We have advanced in the effectiveness of the treatment of spinal curvatures since the time of the ancient Greeks and Romans, when various types of traction devices and trun- cal wraps reportedly were advocated. However, Hippocrates’ description of the pathophysiology of spinal curvatures is fairly accurate. In this chapter, we will discuss the evaluation and treatment of people disabled by spinal deformity, including those whom Hippocrates might have described as in good health (idiopathic), those with natural conformations (con- genital), those with curves due to “habit” (neuromuscular), and those who are bent from old age and pains (degenerative and traumatic).

Colloquially, scoliosis is often used as a term to include all spinal deformities. The term was coined by Galen in the second century A.C.E (1). It is derived from the Greek word “skolios” (2), meaning bent, twisted, or curved. As such, it could include any spinal curvature. However, Galen defined it as the (abnormal) curve of the spine in the coronal (aka frontal) plane (Fig. 34-1). Sagittal plane pathology results from excess or reversal of the spine’s natural sagittal curves. Kyphosis, the word Galen applied to sagittal curves with a posterior vertex (concave anteriorly), derives from the Greek word “kuphos” (2), meaning bent forward or humped. Within a range of degrees, it is normal in the thoracic spine. Lordosis is the word used to describe a sagittal curve with an anterior vertex (concave posteriorly). It derives from “lordos” (2), meaning excessive posterior bending. Within a range of degrees, it is normal in the cervical and the lumbar spine (Fig. 34-2).

WHAT IS A PATHOLOGICAL CURVE?

Although the presence of any fixed frontal curve could be con- sidered outside of the design parameters of the spine and abnor- mal, it is generally accepted that some degree of frontal curve may be “within normal limits.” The Scoliosis Research Society (SRS) has defined a medically significant frontal plane curve (scoliosis) as any curve which is greater or equal to 10 degrees, with or without a rotatory component (3). The definition of a pathological sagittal plane curve is less settled. The spine is “balanced” in the sagittal plane along a plumb line dropped from the center of C7 to the sacral promontory (4). This bal- ance can be influenced by the position of the head and neck, as well as the pelvis and lower limbs. This dynamic balance may help explain some of the outstanding controversy about pathological sagittal curves. Numerous authors (1,5) give the normal range of thoracic kyphosis to be 20 to 40 degrees, oth- ers (1) give the range as 30 to 50 degrees. Fon et al. (6) have shown that kyphosis increases about 4 to 5 degrees over child- hood to the end of adolescence. Kyphosis is also known to increase with aging. The SRS has agreed that normal thoracic kyphosis ranges from 20 to 50 degrees. The range of normal for lumbar lordosis is even wider. Jackson and McManus (7) found the mean lordosis to be 59 degrees. Vaz et al. (8) found a mean lumbar lordosis of 46.5 degrees, with a range from 26 to 76 degrees. The SRS has agreed (4) that normal lumbar lordosis can range from 31 to 79 degrees, depending on measurement technique. However, in assessing any curve detected in a clinical situation, it is important to neither simply rely on curve measurements which are within the “normal” range nor to consider the curve amplitude as the only factor in assessing the patient and his/her complaints and impairments.

EPIDEMIOLOGY

As will be discussed later, spinal deformities can impair not only the musculoskeletal system but also many other organs and systems, often in people already coping with other impairments from chronic disabling conditions. Scoliosis
and the other spinal deformities are not uncommon. The prevalence of idiopathic scoliosis is reported to be between 0.3% and 2% of the population (9). The incidence of degenerative scoliosis is reported as 6% in people over 50 years of age and 36% in persons over 50 with osteoporosis (10). Scheuermann’s disease, a common cause of pathological thoracic kyphosis, has a prevalence of between 0.4% and 8.3% of the population. Up to a 100% incidence of spinal deformity can be found in people with many neuromuscular conditions (see section on “Neuromuscular Scoliosis”).

A physiatrist cannot avoid scoliosis by subspecializing. Spinal deformity is seen in all ages and almost every type of disabling condition (Table 34-1). Physiatrists who do not treat children cannot avoid childhood onset spinal deformities; they will encounter adults who had spinal deformities as children and need to help them address late sequelae of their deformities.

Hopefully, you have now gained an appreciation for the definition of scoliosis and its importance as an issue in the care of persons with disability and/or pain. We will now proceed to develop a construct to use to help better understand these conditions and effectively care for the persons with these disabilities.

---

**FIGURE 34-1.** The planes of the body. (Courtesy of National Cancer Institute’s SEER Program web-based training module.)

**FIGURE 34-2.** There are three basic types of pathological spinal curves. (Courtesy of Donald Bliss and Alan Hoofring of the Medical Arts and Publications Branch, NIH.)
and their families. By persevering through the remainder of this chapter, you will have the chance to

- Regain sufficient memory of the anatomy and physiology underlying the balance of the spine in order to appreciate the factors contributing to the various spinal deformity pathologies and to have a basis to develop a rational treatment plan
- Understand how to obtain the crucial historical, clinical, and radiographic data to form an educated diagnosis and prognosis
- Appreciate the impact of spinal deformity on a person's quality of life
- Understand the pathophysiology and etiologies of spinal deformity
- Understand the roles of physiatric modalities in the (re)habilitation of persons with spinal deformity
- Understand the role of surgical treatment
- Consider the issues involved in postsurgical rehabilitation

### SPINAL ANATOMY AND PHYSIOLOGY

#### Three Columns for the Price of One (Bony Functional Anatomy)

To achieve the elegant balance of flexibility and stability that the human spine offers, the design requires 7 cervical, 12 thoracic, 5 lumbar, and a fused set of sacral and vestigial coccygeal vertebrae. The stability comes from the fact that the spinal column is actually three columns in one. The anterior column is composed of the anterior longitudinal ligament and the anterior portion of the vertebral body. The middle column is made up of the posterior wall of the vertebral body and the posterior longitudinal ligament. The posterior column is formed by the posterior bony arch, which consists of the transverse processes, facets, laminae, and spinous processes. Each vertebra has the potential for six degrees of freedom (Fig. 34-3): translation in all three axes of movement and rotation around each axis. However, all vertebrae are not created equal. Movement at different levels of the spine is limited in some planes. In essence, in the function of the spine follows its form (with apologies to Louis Sullivan). The cervical vertebrae have the greatest freedom including free flexion, extension, lateral rotation, and lateral flexion (this is fortunate, since this freedom affords the eyes and ears a wider field of function). This mobility is due to larger disks, concave lower and convex upper vertebral body surfaces, and transversely aligned facet joints. Thoracic vertebrae have restricted flexion and extension and limited rotation but freer lateral flexion due to their attachments to the rib cage, smaller disks, flatter vertebral body surfaces, frontally aligned facet joints, and overlapped spinous processes. The lumbar spine has good flexion and extension and free lateral flexion due to large disks, posteriorly directed spinous processes, and sagittally directed facet joints. However, there is only limited lateral lumbar rotation due to facet alignment (Fig. 34-4).

#### Ties That Bind (Elastic Tissues)

The intervertebral disks form one fourth of the total length of the spinal column (11). Since they are a viscoelastic tissue, they serve as shock absorbers and take part of the load applied to the spinal column. They also redistribute the load applied on the vertebral bodies, moving more of the force to the stronger, cortical bone (arranged in parallel with the length of the spine) at the vertebral end plates than the thinner layer of cortex sandwiched between softer cancellous (and therefore weaker) bone in the center of the vertebral body. The disks also help connect the vertebra. And, as implied above, they participate in the articulation of the spine.

A number of ligaments also help to link the spine together. The anterior longitudinal ligament is a thick, strong ligament that stretches from the occiput to the sacrum along the anterior surface of the vertebral body. It attaches to the vertebral end plates and is confluent with the anterior portion of the annulus fibrosis of the disk. At the posterior aspect of

---

**TABLE 34.1** Types of Scoliosis and Typical Ages of Onset

<table>
<thead>
<tr>
<th>Pathophysiologic Type of Scoliosis</th>
<th>Typical Age of Onset</th>
</tr>
</thead>
<tbody>
<tr>
<td>Idiopathic</td>
<td>Infancy to adolescence</td>
</tr>
<tr>
<td>Neuromuscular</td>
<td>Childhood to adolescence</td>
</tr>
<tr>
<td>Bone or ligamentous dysfunction</td>
<td>Childhood</td>
</tr>
<tr>
<td>Traumatic/posttraumatic</td>
<td>All ages</td>
</tr>
<tr>
<td>Infectious or neoplastic</td>
<td>All ages</td>
</tr>
<tr>
<td>Degenerative</td>
<td>Middle age to seniors</td>
</tr>
</tbody>
</table>

**FIGURE 34-3.** Six degrees of freedom are possible at each vertebra. The vertebra can move in each of the planes and rotate around each of the axes. (Courtesy of Donald Bliss and Alan Hoofring of the Medical Arts and Publications Branch, NIH.)
the vertebral body, there is a similar ligament, the posterior longitudinal ligament. However, it has an hourglass shape and is thinner than the anterior ligament. There are also three ligaments which connect portions of the posterior column of the spine. The supraspinous ligament stretches across the tips of the spinous processes while the interspinous ligaments stretch from one spinous process to the next. The ligamenta flava (flava is Latin for yellow and the ligament is that color due to the high proportion of elastic tissue within it) connect the laminae of adjacent vertebrae from the anterior surface of the proximal lamina to the posterior surface of the distal lamina. They also blend in with the capsule of the zygapophyseal joints.

**Spinal Articulations**

The spine has two primary joints. The intervertebral joints (as noted above) are formed by the surfaces of the vertebral bodies with the disks. These joints (with the restrictions noted above in the discussion of vertebral anatomy)
allow limited movement in all six directions, although each individual joint only moves to a small degree. The zygapophyseal joints are the only true synovial joints in the spine and are found at the junction of the pedicle with the lamina. The form of their alignment contributes significantly to the functional movement at different levels. The transverse alignment in the neck permits good flexibility in all planes. In the thoracic spine, the facets are aligned frontally which, along with the linkage to the rib cage, limits flexion and extension. The lumbar facets are arranged sagitally, limiting lateral rotation.

Dynamic Balance
It would be possible to place a spine consisting of only its bony and elastic tissues on a base and have the spinal column statically balance in an upright position. However, if the point of the flexibility of the spine is to give humans the freedom to move on two legs, then the spine requires a complex system of dynamic control. The neuromuscular system provides this balance, actuated via muscles intrinsic and extrinsic to the spinal column. The paraspinal muscles are a series of layers of longitudinal muscles attached from the occiput to the sacrum. They are partially distinguishable by their fiber direction and length but “from a functional standpoint, there is little reason to subdivide the back musculature in any detail since the muscles work together in large groups” (11). They serve primarily as extensors to resist the pull of gravity in sitting and standing. In addition, the superficial layers will produce ipsilateral rotation when firing unilaterally, while the deeper groups will produce contralateral rotation with unilateral firing. The primary lateral flexors of the trunk are the quadrati lumborum and oblique abdominal muscles. The primary flexors are the psoas, assisted by the recti abdominus.

CLINICAL ASSESSMENT OF THE PATIENTS WITH SPINAL DEFORMITY

History
Like every other medical condition, most of the diagnosis can be made just by taking a good clinical, functional history and performing a physical examination. “As each symptom is given, it should be pursued with a dogged tenacity until all the details are known” (12). We have mentioned that the majority of patients with scoliosis have an idiopathic condition. It is critical to first rule out spinal deformity associated with another disabling condition. Much of this will be accomplished with good history taking which can then direct and focus the physical and radiological examinations.

The first clue to the etiology of the deformity will be the age of onset. If the scoliosis presents at or soon after birth, it is more likely to be associated with a congenital bony deformity or inherited disease involving the neuromuscular system. If the scoliosis begins in late childhood or adolescence, it is likely to be idiopathic. Spinal deformity that begins in young adulthood is most likely due to trauma or infection; in middle age or later, it is likely to be associated with degenerative disk or joint disease and/or osteoporosis.

It is also important to determine if the onset of the curve is associated with any acute problem such as trauma or infection. The degree of curvature and how fast the size of the curve increases also offer insights into etiology. An underlying neuromuscular or ligamentous condition may promote quicker progression. Children within their adolescent growth spurt may have greater progression due to the rate of overall growth. Progression in an older adult may be an indicator of poor bone health.

Details relating to the nature of the curve itself are insufficient to advance the development of a good diagnosis and rational treatment plan. It is vital to inquire about associated complaints, signs, and symptoms. A thorough review of systems should be performed. However, certain areas of inquiry are more likely to yield clues. Although pain is associated with spinal deformity in adults, it is uncommon in children and adolescents (13) and obligates the physiatrist to search for underlying pathologies including neoplasm or infection. Similar to other pain-evoking conditions, details about the quality, intensity, and temporal and spatial parameters of the pain will help distinguish among bony, soft-tissue–based, and neuromuscular-based pain. Cardiopulmonary dysfunction is the most serious morbidity of spinal deformity. Therefore, questions about fatigability, shortness of breath, palpitations, and decrease in endurance must be included in the history. Any suggestion of cardiopulmonary dysfunction should lead to a more rapid and aggressive workup.

It is important to inquire about the symptoms of central nervous system dysfunction. Problems with loss of balance or falls may indicate a coordination problem such as is seen in cerebellar syndromes. Complaints of muscle spasms may indicate spasticity. Conversely, complaints of weakness or sensory loss may indicate peripheral nervous and/or muscle pathology. Reports of frequent joint sprains may indicate the presence of a condition associated with ligamentous laxity. In children and adolescents, it is important to obtain a history of developmental milestones. Indication of delay in any of the developmental spheres—social/behavioral, communication and cognition, fine motor, gross motor—may provide clues to the presence of an underlying syndrome. Since spinal deformity may be initiated or exacerbated by conditions affecting other areas of the body, the physiatrist should ask about conditions that may affect the lower limbs, such as a history of hip dislocation or leg length discrepancy, and even conditions associated with asymmetry of the upper limbs and head.

Last, but certainly not least, a thorough social and functional history must be obtained. Further clues into etiology may be gleaned. In addition, the insights obtained will identify the barriers to successful functional outcome to the treatment plan and ensure that the goals of the rehabilitation team are in concert with the goals of the patient and his or her family. In some cases, it may be beneficial to use a standardized questionnaire to follow the impact of the person’s spinal deformity on the function and related aspects of quality of life. There
is a validated instrument focused specifically on persons with spinal deformity (14–16). Alternatively, one of the many generic “disability/burden of care” or “health-related quality of life” instruments may be used (see Chapters 11 and 18).

Physiatric Clinical Examination

Although school screening programs for idiopathic scoliosis may only look at a child’s spine, it is not sufficient for a physiatrist to limit the assessment of spinal deformity to the spine. A comprehensive physiatric examination must first be done. This should include assessment of pain, strength, endurance, sensation, balance (in sitting and standing), coordination, range of motion, reflexes, mobility, and dressing skills. Leg length discrepancy should be assessed, which is not always easy to determine, and corrected.

Once the general examination has been completed, a more focused assessment of the spine must be performed. You do not have Superman’s x-ray version; clothing must be removed so that the spine can be observed directly. The surface anatomy should be viewed in both prone and side lying. If the patient can sit, then the spine must be seen in sitting. If he or she can stand, even with support, then the spine must be viewed in standing. If there is concern about leg length discrepancy and/or pelvic asymmetry, the spine should also be viewed in standing with leveling blocks under the foot on the “short” side and the depth of the correction recorded. Any difference in the shape or degree of curve in different positions should be noted, it is a sign of the degree of flexibility of the curve. The spine should be viewed from both the frontal and sagittal planes. The spinous processes should be palpated to better appreciate the curves’ dimensions, including shape in both frontal and sagittal planes and presence of rotation. The curve should be described by its length (approximately which vertebrae start and end the curve), its shape (e.g., “C” or “S” shaped), and the direction of the vertex of the curve. Truncal range of motion in all planes—forward flexion, rearward extension, R & L lateral flexion, and R & L lateral rotation—should be viewed with special attention to asymmetry and spinal levels where movement is occurring. While the patient is flexing forward, the examiner should look at the shape of the rib cage for prominence or humping (Adam’s forward bending test (17), which is the test used in school screenings) (Fig. 34-5), which may be a sign of a rotatory component to the curve. The shape of the rib cage should be viewed from all angles to look for deformity. Rib cage deformity may not only be a sign of rotation of the vertebrae but also suggest underlying pathologies such as osteogenesis imperfecta. One must look not only at rib prominence but also for deformities such as pectus carinatum or excavatum. Symmetry at the top and bottom of the trunk should be assessed in sitting and standing. Shoulder

FIGURE 34-5. Adam’s forward bending test. (Courtesy of Donald Bliss and Alan Hoofring of the Medical Arts and Publications Branch, NIH.)
symmetry should be viewed at the level of the acromioclavicular joint (13). Pelvic symmetry can be viewed either at the brim of the pelvis at the iliac crests or at the posterior superior iliac spines, which is seen on the surface of the buttocks as the two sacral “dimples.”

Measurement of symmetry of the appendicular skeleton and head should be done to rule out asymmetries. Head size and alignment should be viewed since reports have associated these issues with scoliosis (18). There are reported associations between upper limb deficiency and scoliosis (19–21), and, anecdotally, it has been associated with neonatal brachioplexopathy (Erb’s palsy). Therefore, upper limb asymmetries should be assessed. Upper limb length should be measured in a consistent manner. One method is to measure from the acromioclavicular joint to the radial styloid. Lower limb discrepancy produces functional scoliosis in standing and it may contribute to the development of fixed scoliosis, as well (22). Lower limb length is commonly measured from the anterior superior iliac spine at the pelvis down to the medial malleolus. The physiatrist must also look for functional asymmetries that could result from joint contractures or deformities (especially at the hips and knees) or tendon contractures (especially tendons of two joint muscles such as the hamstrings and gastrocnemius). Functional asymmetry may also result from asymmetric neurological function including hemiparesis.

RADIOGRAPHIC ASSESSMENT OF THE SPINE

Measuring the Curve Angle

The standard technique for measuring the angle of a spinal curve is called the Cobb angle (23). Initially developed as a frontal plane measurement, the technique can also be used for assessing sagittal curves. We will review the method for assessing scoliosis, but the rules can be applied similarly in lordosis and kyphosis. The film from which the curve is viewed should, if possible, include the entire spine on one film and be taken posterior-anterior (to reduce the dose of radiation to the breast) (13). Even if the patient is only suspected of having a frontal curve, a lateral film must also be obtained to better assess the amount of rotation in the curve. If at all possible, the films should be done in standing. If the patient is unable to stand, then sitting is preferred to a supine film. The P-A film should be taken with the arms at the side; the lateral film should be done with the arms at 90 degrees of abduction to keep the upper limb from obscuring the spine. Although it is common to shield the perineum when films are done, the initial films should be done without gonadal shield in order to fully visualize the hips, pelvis, and proximal femurs (13). In children and adolescents, a sense of skeletal maturity can be gleaned by looking for Risser’s lines at the crest of the pelvis.

To perform the Cobb measurement, the examiner should first note the upper border, lower border, and vertex of the curve. The curve is classified by the position of the vertex into thoracic, thoracolumbar (vertex at the thoracolumbar junction), or lumbar. The upper and lower borders are identified by the “end vertebrae.” They are the first and last vertebrae that tilt into the concavity of the curve (13). To define the angle, a line is drawn on the radiograph parallel to the upper end plate of the upper end vertebra. A second angle is drawn parallel to the lower end plate of the lower end vertebra. In most cases, these lines will not intersect on the film and, therefore, the angle cannot be measured directly. Instead, using principles learned in high school geometry, a line perpendicular (90 degree angle) to each of the above lines is drawn. The angle formed by the intersection of the perpendicular lines is equal to the angle made by the upper and lower end vertebrae and is measured and reported as the Cobb angle (Fig. 34-6).

More and more facilities are converting to filmless, computer-based radiological information systems. Many of these systems offer tools to directly measure the Cobb angle on the screen. The semimanual methods still require the clinician to place the parallel lines at the top and bottom of the curves. Studies have shown these methods to have precision and
reliability similar to the film-based method (24–26). More recently, some systems have introduced fully automated Cobb angle calculations. However, it appears that this technology still requires further refinement to achieve adequate accuracy for curves of all magnitudes (27).

Assessing Rotation
As mentioned above, the spine has six degrees of freedom, including rotation along the axis of the spinal cord. Many spinal curves are associated with rotation of the vertebral. Curves with rotation may be more difficult to treat and may cause greater disturbance to the rib cage with resultant adverse effects on respiration. Therefore, it is important to assess the presence and degree of rotation associated with the curve, as well as its rate of progression. Radiographically, the position of the pedicles provides the best sense of degree of rotation on a plain x-ray. The examiner should note the symmetry and alignment of the pedicles in the P-A film. If asymmetry is present, the pedicle that is more prominent will indicate rotation away from that side of the vertebra. The degree of rotation can be graded using the system of Nash and Moe (13) (Table 34-2). Automated computerized systems to assess rotation are also under development (28).

New computer-assisted techniques are also under development to supply additional information about the three-dimensional characteristics of the curve using the data supplied by digitized plain films (29). Such information may be useful in assessing the results of nonoperative and for planning operative treatments. Despite the additional radiation dosage,

<table>
<thead>
<tr>
<th>Table 34.2 Moe-Nash Classification of Spinal Rotation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Nash and Moe Grade</strong></td>
</tr>
<tr>
<td>------------------------</td>
</tr>
<tr>
<td>Grade 0</td>
</tr>
<tr>
<td>Grade I</td>
</tr>
<tr>
<td>Grade II</td>
</tr>
<tr>
<td>Grade III</td>
</tr>
<tr>
<td>Grade IV</td>
</tr>
</tbody>
</table>
three-dimensional CT may be useful, especially in congenital disorders (30,31). MRI might offer similar information without radiation exposure (although still with the need for sedation in groups including young children and adults with claustrophobia); however, at the current state of technology, recommendations for its use are still limited to ruling out potential etiologies or neurological abnormalities (32). It is also important to remember that neither CT nor MRI is able to assess the impact of gravity on a curve, since both studies are done in a supine position.

Assessing Vertebral Integrity

It should be obvious by now that among the etiologies of spinal curves are conditions intrinsic to the bone. Most of these will become evident upon careful inspection of the radiographs obtained to determine the magnitude of curvature. The spinal elements should be inspected for signs of hemivertebra, unsegmented bars, and/or bifid spinous processes (see Fig. 34-10 later in this chapter). The ribs should be viewed to look for asymmetry or changes in the natural curve (Fig. 34-7). The pelvis should be examined for signs of asymmetry (Fig. 34-8). The hip joint should be examined for signs of subluxation. Femurs should be assessed for any asymmetry and separate femur films ordered if there is any suspicion of pathology.

THE IMPACT OF SCOLIOSIS ON LIFE FUNCTION (SECONDARY MORBIDITIES)

The reason that this chapter is found in Physical Medicine and Rehabilitation: Principles and Practice is because spinal deformity is not just a cosmetic problem; it impacts on many aspects of quality of life. The types of functional morbidities and their intensity may vary with curves of different etiologies; however, certain problems may be found in almost all types of spinal deformity. In this section, we will review the areas of most common functional concern in persons with spinal curves.

Pain

Pain may be the presenting symptom of spinal deformity; it is in 40% to 90% of adult scoliosis (33). In thoracic kyphosis caused by Scheuermann’s disease, it is an early symptom in 20% to 60% of patients. However, before attributing pain to idiopathic or degenerative etiologies, it is vital to remember that pain may be the sentinel of a critical, even fatal illness. Pain may be the presenting symptom in infectious processes such as epidural abscess or discitis (34). Rarely, one can have a spontaneous epidural hematoma of the spine. In these cases, due to nerve irritation, the pain may be radicular in character. Of greatest concern is to rule out pain due to tumor. In benign

FIGURE 34-7. Rib asymmetry associated with spinal deformity in an adolescent girl with osteogenesis imperfecta. (Courtesy of Department of Imaging Sciences, Warren G. Magnuson Clinical Center, NIH.)

FIGURE 34-8. Pelvic asymmetry (and rib deformity and osteopenia) associated with spinal deformity in a child with osteogenesis imperfecta. (Courtesy of Department of Imaging Sciences, Warren G. Magnuson Clinical Center, NIH.)
tumors, the delay between the onset of symptoms and the diagnosis can be up to 19 months. In malignant tumors, it can be up to 4 months between symptom onset and diagnosis. It is important to recall that tumor-related pain is often worse at night and is disassociated from any obvious initiating or exacerbating factors (35).

Chronic pain plays a role in many spinal curvature pathologies. As implied above, it is common in degenerative spine conditions with associated curvature. More than 50% of patients with Scheuermann’s disease have pain later in their course, especially if L₁ and L₂ are involved (36). There appears to be an increased risk of pain in adults who had scoliosis in childhood or adolescence (37). Pain is often a problem in persons with achondroplasia and is felt to be associated with the mechanical disadvantage of the lordosis on sagittal spinal alignment. Back pain is also found in osteogenesis imperfecta. The pain in osteogenesis imperfecta may also be related to the mechanical disadvantage of the lordosis.

**Mobility**
As noted above, normal spine function is a delicate dynamic balance between rigidity and flexibility. It is not surprising, then, that balance may be affected by spinal deformity. When the patient has an underlying neuromuscular condition, the spine asymmetry is likely to exacerbate the difficulties with balance control. What is more unanticipated is that balance is disturbed in idiopathic scoliosis as well. Reports have shown increased amounts of sagittal and lateral sway in standing, along with increased sway radius, leading to decreased stability in standing (38). This may result in increased losses of balance and falls as well as increased and asymmetric stress on lower limb joints, creating the potential for a vicious cycle. Recent reports suggest that there may be subtle underlying neurological deficits in persons with idiopathic scoliosis manifesting as balance difficulties (39,40).

**Cardiopulmonary Function**
The most critical of all the potential complications associated with spinal deformity is compromise to the cardiopulmonary system. There is a common pathway toward cardiopulmonary dysfunction. It begins with restrictive lung disease resulting from deformity of the bellows mechanism formed by the rib cage, thoracic spine, and diaphragm. On pulmonary function testing, this is demonstrated by greater decreases in vital capacity than residual lung volume (41). In untreated idiopathic scoliosis, it has been shown that once curves achieve an angle of 70 degrees there will be a mild decrease in vital capacity. By the time a curve reaches 90 to 100 degrees, the patient with have dyspnea on effort. With curves greater than 100 degrees, alveolar hyperventilation, CO₂ retention, pulmonary hypertension, and right-sided heart failure are seen (37,41). At this point, the patient may be irrevocably set on a downward slide of illness leading to death due to pneumonia and/or heart failure. It is likely that these problems will present with much smaller curves when there is also an underlying neuromuscular disease or syndrome affecting other organ systems. Prevention of serious cardiopulmonary decline is one of the major indications for surgical correction. Close monitoring of pulmonary function tests is an essential part of ongoing follow-up of any patient with spinal deformity, especially if progression is noted. It is quite important to remember that there is a “window of opportunity” in progressive curves and that surgery in such curves must be done before pulmonary function has deteriorated to such a degree that the risk of inability to wean from mechanical ventilation after surgery is greater than the potential benefits of the surgery.

**Emotional and Behavioral Health**
A physiatrist is obligated to approach his/her patient holistically. As with any other disability, the emotional impact of spinal deformity should be addressed. Spinal deformity has traditionally had tragic, negative connotations in Western society as evidenced by Shakespeare’s Richard III and Victor Hugo’s Quasimodo, The Hunchback of Notre Dame. Our modern society often appears obsessed with physical beauty and it seems as if these societal pressures are greater on women and girls. It is important, then, to be mindful of the potential impact that spinal deformity may play on the patient’s body image and emotional health. Studies of patients with idiopathic scoliosis bear out this concern, with findings showing poorer body image, greater unhappiness, lower self-esteem, increased participation in high-risk behaviors, and greater incidence of depression (42–44). Questions relating to the emotional and behavioral status of a patient with spinal deformity should be a regular part of the physiatrist’s ongoing follow-up of the patient. There should be no hesitation to enlist the help of psychological support professionals such as rehabilitation psychologists and social workers to address the emotional needs of the patients and their families.

**PATHOPHYSIOLOGY OF SCOLIOSIS**
Spinal stability is derived from the design of the bony, elastic, and muscular elements of the spine and how they are controlled by the nervous system. The bony anatomy has a fair degree of intrinsic stability engineered into it, but must allow movement. Stability is reinforced by the elastic structures but, by their very nature, they will allow movement. The motor control system is designed to adjust to the movements engineered into the bone and elastic tissue system. Understanding the anatomy enables one to determine factors contributing to spinal curvature and better formulate a plan of treatment.

One useful division is to separate factors intrinsic to the spinal column from extrinsic factors. In looking at the intrinsic factors, one should consider how aspects of each of the elements of the spine—bony, elastic, and neuromuscular—may be affected. Once one has identified which systems are involved, the type of involvement should be understood. Is there a failure in the system or is the system overactive? Then, consider the etiology of the impairment. Is it congenital or...
acquired? If it is acquired, is the etiology traumatic, infectious, neoplastic, autoimmune, or degenerative? Understanding the biology of the dysfunction is important, that is, is the problem due to a metabolic defect in the tissue or is the tissue reacting to an outside influence. The nature of how these systems interact should also be considered. For example, does the elastic pathology cause secondary bony deformity?

Factors extrinsic to the spine can influence the spine in one of two ways. They can directly impact on the mechanical function of the spine by applying asymmetric forces to the bony and elastic tissue complex of the spine. Disorders resulting in misalignment of the appendicular skeleton, rib cage, and/or head can also apply abnormal forces to the spine. The spine can also be indirectly influenced by abnormal, especially asymmetric, function of the neuromuscular system extrinsic to the spine itself, even if spinal neuromuscular system is intact. It should also be remembered that this extrinsic neuromuscular dysfunction can directly affect the spine or it can indirectly influence the spine by causing extrinsic mechanical failures such as joint contractures or dislocations.

Figure 34-9 is a graphical representation of how all these factors may be organized. However, the relationship between the all of the factors—intrinsic and extrinsic—is even more complex than it appears because of the potential for interactions between each and every factor. Table 34-3 is a list of pathological conditions associated with spinal deformity grouped by category, based on the taxonomy developed by the SRS (9).

**SCOLIOSIS DUE TO BONY DYSFUNCTION**

**Congenital**

A number of disorders of vertebral development can occur. These anomalies are present at birth, but they may not be detected immediately. The curvature may be subtle and not noticed until the child begins sitting or standing. Often, clinical evidence of curvature may not be apparent until the child grows and one part of the vertebral column grows faster than another, thus revealing the asymmetry associated with the anomaly. Winter (45) has developed a classification system for these anomalies based on the underlying embryological dysfunction. The dysfunction may be a failure of segmentation, a failure of formation, or a combination of both. The archetype of segmentation failure is a bloc vertebra, formed at two or more spinal levels. This is less likely to cause a curve than a unilateral failure to segment, known as a unilateral bar. The location of the bar will determine the type of pathological curve. A lateral bar will result in scoliosis, a posterior bar will produce kyphosis, and an anterior bar will result in lordosis.

The defect in formation that will result in curvature is usually called a hemivertebra. In reality, any amount of partial defect is possible. The partial vertebra may be “incarcerated,” meaning that the vertebrae above and below it have altered their shape to compensate for the abnormal shape of the partial vertebra, therefore limiting the degree of curvature. A nonincarcerated hemivertebra is more likely to result in curvature because its neighboring vertebrae have not adapted to the hemivertebra's abnormal shape. Figure 34-10 shows representations of the basic types of deformities according to the Winter classification.

Although congenital bony anomalies may be isolated, it is incumbent to first rule out associated genetic disorders. From 25% to 40% of children with congenital bony anomalies also have other problems including genitourinary anomalies, cardiac anomalies, Klippel-Feil syndrome, and spondylothetic dysplasia. However, the majority of curves due to congenital bony anomalies are not severe. Only 38% show severe progression, whereas 47% show only mild to moderate...
progression and 15% show no progression (46). The character of the curve will help determine likelihood of success of a given treatment approach. Long, flexible curves may respond to orthotic management. Short, stiff curves will require surgery if they progress.

### Acquired Bony Pathology

The bony spine can be directly injured by blunt force or projectile trauma (see Chapter 27). This trauma can result in fracture and/or dislocation of spinal elements. Each year, there are between 18,000 and 50,000 spine fractures in the United States (47, 48). The spine can also be injured as an unintended side effect of radiation. The instability resulting from spinal trauma can result in pathological spinal curvature, even without injury to the spinal cord. Development of pathological curvature is more common when the patient is managed without surgery, but curves can also occur after surgery, often as a result of non-union of fusion and/or implant failure. Kyphosis is the most commonly seen deformity after spinal trauma. When surgery is indicated, this may require both posterior and anterior fusion (48).

Historically, tuberculosis was a common cause of spinal deformity. TB osteomyelitis, known as Pott’s disease, has been found in Egyptian mummies (49). Most often, it produces a sharply localized kyphosis called a gibbus deformity. TB is still important to consider, especially due to the prevalence of atypical forms of mycobacteria in immunocompromised populations such as persons with HIV. Acute pyogenic osteomyelitis, most likely due to Staphylococcus aureus, is more common than TB. It will present with associated fever, malaise, and back pain. There will be elevations in white blood cell count and ESR. Unfortunately, plain x-rays may not reveal any pathology for 10 to 14 days after the onset of symptoms. Damage usually starts in the vertebral body and spreads via the disk to other levels. It is uncommon for the infection to start in the disk or posterior spinal elements. As in TB, kyphotic deformity is most common (49). Treatment for these deformities depends on the number of levels involved. A one-level infection may be managed with a spinal orthosis. If two or more levels are involved, surgical intervention will likely be needed; neurosurgical participation will often be required to prevent spinal cord compromise (49).
Although less than 10% of all neoplasms are primary bone tumors, 50% to 70% of persons who have died of neoplastic disease have shown some evidence of bony involvement (35). The spine is the most common site of bone metastases (35), making neoplasm an important potential etiology of spinal deformity. Pain will often be the first symptom. In adults, malignant tumors are more often found in the vertebral body or pedicle, whereas benign tumors are found more often in the posterior elements. Primary benign tumors include osteoid osteoma, osteoblastoma, giant cell tumor, and aneurysmal bone cyst (which is not a true tumor but rather a hyperemic osteolytic process) (35). Osteoid osteoma usually resolves after surgical removal, which should limit deformity. Osteoblastoma, although rare (<1% of all primary bone tumors), is present in the spine in 40% of cases. Half of those cases will have scoliosis (35). Aneurysmal bone cysts may involve the neural arch, body or posterior elements and can produce scoliosis or kyphosis due to osteolysis. While only 12% of giant cell tumors involve the spine, 50% of those that do will involve the vertebral body and can result in vertebral body collapse and subsequent kyphosis and/or scoliosis (35).

Primary malignant bone tumors are rare and do not commonly involve the spine. Less than 2% of osteosarcoma, only about 5% of Ewing’s sarcoma, and not many more of chondrosarcomas are found in the spine. If they are successfully removed before the spine becomes unstable, spinal deformity is unlikely (35). However, as noted above, metastases to the spine are common in solid tumors including those originating in the breast, prostate, thyroid, kidney, bronchus, esophagus, the gastrointestinal tract, and the rectum. Of white blood cell–derived tumors, multiple myeloma is most commonly found in the spine, although lymphoma is also found. As with other spinal neoplasms, metastases usually present with pain. If pathological spinal curvature is present in metastatic bone disease, it is quite possible that the tumor has caused spine instability and urgent evaluation and treatment should be performed to prevent neurological compromise.

The most common cause of acquired spinal deformity arising from the bony spine is degenerative. The overall incidence of degenerative spinal curvature is between 2% and 4% of the population (50). Since the presenting problem is usually pain, it is important to first rule out underlying pathologies. Temporary, functional curvature can result from neurogenic pain or disk or facet pathology (51) and should be distinguished from fixed curve. Spinal deformity is fairly common in women with osteoporosis, with scoliosis occurring in between 36% and 48% (52). The onset of curvature is thought to occur in the 5th decade of life, but it is possible that subtle changes occur earlier but are clinically undetectable (52). Primary adult spinal deformity is due to asymmetric vertebral collapse (51). There is a positive correlation between osteoporosis and the presence of thoracic kyphosis (5). However, the role of osteoporosis in frontal plane deformity is still controversial (51). Most biomechanists believe that degeneration of the intervertebral disks plays an important role. The loss of disk height and its ability to more evenly distribute force across the vertebral body is felt to increase the stress and load on the vertebral bodies, thus promoting collapse (53). Osteoarthritis and lax elastic tissues can result in abnormal spinal motion with resultant abnormal
stresses and this may also play a role (54). The most common primary curves are thoracic kyphosis and lumbar scoliosis. The scoliotic curves are usually short and focused at L2-3 or L3-4. They are detected by the 6th decade of life and are associated with a loss of lumbar lordosis. To better plan treatment, it is important to rule out curves which may have resulted from secondary degeneration caused by other bony pathology within and without the spine. This includes congenital anomalies such as facet asymmetry, idiopathic childhood or adolescent scoliosis, deformity resulting from trauma, infection, neoplasm, or autoimmune disease (e.g., kyphosis due to ankylosing spondylitis (50)), unilateral spondylolisthesis, and extraspinal asymmetries such as leg length discrepancy (51).

SCOLIOSIS DUE TO ELASTIC TISSUE DYSFUNCTION

Congenital
Many inherited disorders of connective tissues are associated with spinal deformity. They include diseases such as osteogenesis imperfecta, achondroplasia, Marfan syndrome, and Ehlers-Danlos syndrome. Of and achondroplasia are examples of complex conditions that impact on both bone and connective tissue. Along with the knowledge of unique aspects of each disease, the physiatrist should be able to apply his/her understanding of how “pure” bone and “pure” elastic tissue syndromes impact on spinal curvature. As such, we will focus this section on Marfan and Ehlers-Danlos syndromes as examples of conditions that more “purely” affect elastic tissues.

Marfan syndrome (listed in Online Mendelian Inheritance in Man as OMIM #154700) is an autosomal dominant syndrome that results from a defect in the fibrillin gene at chromosome 15q21.1 (55). In addition to diffuse joint laxity (including the spine) and long limbs, it produces ocular deficits including myopia and flat corneas, and valvular heart dysfunction including mitral valve prolapse, aortic valve regurgitation, and dilation of the aortic root (which can lead to potentially fatal aortic aneurysm). Scoliosis may be seen in 52% of Marfan patients (56). When present, the curves tend to be severe and double major curves (i.e., “S” curve with thoracic and lumbar components) are common (57). In addition to the prevalence of long limbs, one study has shown a 5% prevalence of leg length discrepancy greater than 2 cm in Marfan patients. In that group with leg length discrepancy, over 90% had scoliosis greater than 10 degrees (56). This suggests that the pathology of spinal curvature in Marfan syndrome is not only due to the diffuse joint laxity. Since these patients are at risk for severe curves, they should be followed closely. Bracing should be begun at the lower end of the range of curvature treatment. If the curve progresses despite bracing, decision to proceed to surgery should also be made at the lower end of the treatment range.

Ehlers-Danlos is a family of genetic disorders (types I to VIII are OMIM #13000, 130010, 130020, 130050, 225400, 130060, 130080; there are more associated conditions) phenotypically associated with diffuse joint laxity (58–65). Types VI and VII have most often been associated with kyphoscoliosis (58–65). Types VI and VII have most often been associated with kyphoscoliosis (63,64). The actual incidence of spinal curvature in these groups is not yet well-known (66). Type VI is an autosomal recessive disorder presenting with joint laxity and recurrent joint subluxations, Marfanoid habitus, risk of ocular rupture and a progressive, often neonatal onset, scoliosis. Decreased hydroxylysine is found on laboratory investigation and many of the cases have a deficiency in the enzyme lysyl hydroxylase (63). Double major curve appears to be the predominant pattern of scoliosis in this group (67). Type VII is an autosomal dominant disorder which results from defects in the COL1A1 and COL1A2 genes (which happen to also be involved in osteogenesis imperfecta). This type of EDS is associated with severely lax joints and frequently recurrent subluxations. They also have short stature and midface hypoplasia. Hypotonia and gross motor delay are also commonly seen. Kyphoscoliosis is the most common curve pattern in this group (64). Scoliosis is also reported in up to 50% of persons with one of the more benign types of Ehlers-Danlos, EDS III (67). Persons with EDS have vascular fragility often resulting in easy bruising. This aspect of EDS must be carefully considered in planning treatment of the curvature. Type of brace chosen and fit is thus more crucial than in other disorders. Surgery has the potential for increased blood loss, perioperative arterial rupture, wound hematoma, and wound dehiscence due to the fragile blood vessels (66,67).

Acquired
It is difficult to identify acquired conditions which purely affect the elastic tissues without affecting the bony spine. Trauma, infection, and neoplasm do not often limit themselves to only the elastic spinal column tissues. Rheumatologic “connective tissue diseases” also do not limit their damage to elastic tissues. However, as noted above in the discussion of degenerative spinal disorders, disk disease is considered to be a major contributor to bony degeneration. Additionally, there are some conditions where it is important to be mindful of the possibility of spinal deformity due to soft-tissue contractures. They include empyema with resultant lung and pleural scarring (68), contractures, due to burns or other injuries, and radiation fibrosis (9).

SCOLIOSIS DUE TO NEUROMUSCULAR DYSFUNCTION

Congenital
Scoliosis in inherited neuromuscular conditions can derive from three basic etiological categories: peripheral weakness, abnormal central control, and abnormal sensory afferent input. Weakness can cause failure of the dynamic balance of the spine and result in curvature. This occurs in lower motor neuron disorders such as spinal muscular atrophy and Charcot-Marie-Tooth disease (which also has a sensory component). It is also a factor in meningomyelocele, whose pathology is more complex, since it is also associated with bony deformities. Weakness is also the etiology in curves resulting from
myopathic conditions such as the muscular dystrophies, congenital hypotonia, and myotonic dystrophy. Abnormal central control of dynamic spinal balance occurs in conditions including cerebral palsy and the spinocerebellar disorders (although some of them, such as Friedreich’s ataxia, are more complex, with additional pathologies such as peripheral nervous system–derived loss of sensation). Riley-Day familial dysautonomia (an autosomal recessive disorder at chromosome 9q31-q33, which primarily affects Ashkenazi Jews) is a good example of a condition with scoliosis due to loss of sensation on the basis of dysfunction in peripheral afferent nerves (69).

All of the curves associated with congenital neuromuscular conditions have some common features. These include similar temporal characteristics: early onset, rapid progression during childhood and/or adolescent growth spurts, and curve progression, which continues into adulthood, since the underlying pathology is still present (70). They often also have an association with dependence in sitting with asymmetric trunk and/or pelvic obliquity (71). However, the pattern of the scoliosis may be similar to that seen in idiopathic scoliosis (72). The inevitability of progression in the more severe types of these disorders means that pulmonary compromise will eventually appear if the curve goes untreated. As mentioned in the section discussing the impact of curves on life function, this is the most serious of all curve-related complications. It is further complicated by the association with cardiac complications that many neuromuscular conditions have, including Duchenne muscular dystrophy and Friedreich’s ataxia (70). These factors preclude any hesitation to proceed to earlier surgical management in order to slow respiratory compromise. There appears to be only a limited role for bracing in neuromuscular patients. It appears to only be effective in preventing progression in patients who are still ambulatory, but it can also be used as a sitting support and/or comfort measure (73).

Spinal muscular atrophy is a large group of disorders that affect the motor neurons in the anterior horn of the spinal cord, resulting in progressive weakness. The prevalence of curvature is reported between 30% and 100% (70,74). They are more likely to present with long thoracolumbar frontal curves. Type I is usually fatal due to respiratory failure by age 2. This early mortality has historically limited the importance of spinal curvature in care of the child. However, given the prevalence of curvature and the underlying weakness of the respiratory muscles, treatment should begin early. Scoliosis in the X-linked Duchenne muscular dystrophy (a disorder in the production of the protein dystrophin due to a defect at chromosome Xp21.2, 12q21) has a similar prevalence, 33% to 100% (74) with a 95% likelihood of scoliosis by the time the patient becomes wheelchair dependent (70). The curves may progress as much as 16 to 24 degrees/year and are also often long thoracolumbar curves, although double major curves are also seen (70).

In addition to decreased pain perception due to afferent dysfunction, people with Riley-Day syndrome have decreased tear production, no flare reaction in response to histamine injection, absence of tongue papillae, decreased to absent deep tendon reflexes, hypertension, and postural hypotension. About 86% of patients will have some form of spinal deformity. Of that group, 53% will have scoliosis, 44% will have kyphoscoliosis, and 4% will only have kyphosis (75). The prevalence of deformity is the same in males and females. Curves present early and tend to be rigid. Left thoracic curves are more common than in patients with idiopathic scoliosis (70). In Friedreich’s ataxia, which has elements of afferent dysfunction and central dysfunction, there is also a very high prevalence of curvature, between 75% and 100% (70,76,77). In one study, the majority (57%) were double major curves. The remainder were thoracolumbar (14%), thoracic (7%), lumbar (4%), and small multiple curves (11%) (77). Of interest is the finding of one group of Friedreich’s ataxia patients followed up for 10 years; they appeared to divide into two subgroups. One (55% of patients) had classic neuromuscular curves with progression of at least 60 degrees. The other group appeared to have idiopathic scoliosis with nonprogressive curves of less than 40 degrees.

Cerebral palsy is the most common congenital/neonatal central nervous system disorder associated with spinal deformity. There is a 15% to 68% prevalence of spinal curves (70,78). Curves are more common in quadriparetic cases and when dysfunction is more severe. As the patient gets older, it is more likely that curve will be detected (78). There is a negative correlation between ambulation and spinal deformity. The role of pelvic obliquity in the development of spinal curves is often debated, especially given the prevalence of hip subluxation and dislocation in cerebral palsy. Given the association between limitations in ambulation and spinal curves, it is reasonable to consider the role of pelvic obliquity and to address conditions such as hip deformity which may produce the obliquity. The amount of progression in adulthood is almost doubled (from 0.8 to 1.4 degrees/year) in untreated curves that are more than 50 degrees by the time the child achieves skeletal maturity. Thoracolumbar curves are more likely to progress than other curves, and scoliosis is often associated with thoracic kyphosis and lumbar lordosis (78).

As mentioned above, meningomyeleocele can present with aspects of both peripheral and central nervous system dysfunction, depending on the level of the lesion, as well as bony deformities. Spinal deformity is present at birth in 20% of patients and 51% will develop some spinal deformity. If the meningomyeleocele is above L2, the likelihood of spinal curvature increases to 87%. Curve presentations include scoliosis in 77% of patients, lumbar lordosis in 57% of patients, and thoracic kyphosis in 23% of patients (79).

**Acquired**

The most common acquired cause of neuromuscular spinal deformity is spinal cord injury in children and adolescents. The likelihood of curvature increases as the age of onset decreases. Preadolescents have a 97% likelihood of acquired spinal deformity. The prevalence reduces to 78% in 14 year olds, 57% in 15 year olds, and 50% in 16 year olds (70). Curvature is not common in adults with spinal cord injury unless there is underlying bony instability.
Although syringomyelia may be congenital, it also may be due to trauma, with or without detectable spinal cord injury. It may also occur as a result of infectious processes such as meningitis or myelitis. Up to 85% of persons with syringomyelia will have scoliosis. Fortunately, most curves are less than 50 degrees in magnitude (70).

Neurofibromatosis is the most common benign tumor associated with spinal deformity. Neurological neoplasms causing spinal curvature are otherwise uncommon. Curvature is more common in NF-1 (von Recklenhausen), with scoliosis and/or kyphosis occurring in 5% of cases (80). There appears to be two types of curve associated with NF. One seems more like idiopathic scoliosis and may be associated with the presence of pseudomeningoceles (80). Curvature can also be due to paraspinal, plexiform neurofibromata impinging on neural structures and/or paraspinal soft tissues. Nonspine complications of NF such as spastic hemiparesis or asymmetric limb hypertrophy can result in spinal deformity. Curvature can also occur as a result of scarring due to radiation fibrosis after treatment of the neurofibromata.

**IDIOPATHIC SCOLIOSIS**

The etiology of the most prevalent cause of scoliosis is still unclear. Between 1.5% and 3% of the population has frontal spinal curvature greater than 10 degrees, meeting the SRS criteria for pathological scoliosis (81). This represents between 80% and 90% of all scoliosis (53,82). A familial association has been observed in a small minority of these cases, and, more recently, a number of genetic polymorphisms have been associated with some of these familial cases of idiopathic scoliosis (83,84). We previously mentioned many factors that can be involved in producing spinal curvature in the section on the general pathophysiology of scoliosis. Studies have suggested that a number of these factors may be involved in what is still classified as idiopathic scoliosis. Yarom et al. (85) have demonstrated asymmetric abnormalities in deltoid muscle including abnormal ratio of type I/type II fibers, with less type I fibers on the concave side of the curve and a myopathic recruitment pattern more prevalent on the concave side on needle EMG. Some association with rotational trunk strength has also been seen (86). Goldberg and Dowling (87) have shown differences in cerebral lateralization in girls with scoliosis. As opposed to controls, they were more likely to have uniform handedness and footedness. Other studies (38,40,82,88) have suggested that balance and vestibular function may be factors. Papaioannou et al. (89) have shown a positive correlation between leg length discrepancy between 1.2 and 5.2 cm and nonprogressive lumbar scoliosis. A recent study has suggested an association between sacropelvic bone morphology in the frontal plane and scoliosis (90).

Although the etiology of idiopathic scoliosis is not yet fully elucidated, a better understanding of the condition can be appreciated by categorizing the condition based on the age of onset of the curve (Table 34-4).

**Infantile**

By definition, the onset of this condition occurs before age 3. Most reported cases of infantile idiopathic scoliosis have been in Europe; the condition is quite rare in the United States (81). Especially given its rarity, it is important to first rule out one of the other causes for scoliosis discussed in the sections above, before making this diagnosis. The condition is usually benign, with 80% to 85% resolving spontaneously. Concern about progression should be raised if the curve is associated with a large difference in the angle formed at the connection of the rib to the vertebra (rib-vertebral angle) on the convex side versus the concave side of the curve (50). Smaller, more rigid curves are also more likely to progress (81). Since infants cannot stand and a supine film may result in false positives due to the child wiggling on the table, the definitive diagnosis of infantile idiopathic scoliosis requires a radiograph taken with the baby in a suspended position (81). Curves of less than 25 degrees with a rib-vertebral angle difference of less than 20 degrees can be monitored with radiographs every 4 to 6 months until the curve resolves and then every 1 to 2 years until the child reaches skeletal maturity. Larger curves are usually managed with a body cast and progressed to a molded orthotic once the child is older and the rate of growth slows. If the curve still progresses, surgery must be performed, but is delayed for as long as it is safe to wait, in order to allow for more spine growth. This is also one of the few instances where spinal surgery may be done.

**TABLE 34.4 Types of Idiopathic Scoliosis**

<table>
<thead>
<tr>
<th>Type</th>
<th>Age of Onset</th>
<th>Percentage of All Idiopathic Scoliosis (%)</th>
<th>Male:Female Ratio</th>
<th>Common Curve Pattern</th>
<th>Usual Side of Curve Vertex</th>
<th>Usual Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Infantile</td>
<td>&lt;3 y</td>
<td>1</td>
<td>Males &gt; females</td>
<td>T or T-L</td>
<td>L</td>
<td>May spontaneously resolve</td>
</tr>
<tr>
<td>Juvenile</td>
<td>3–10 y</td>
<td>12–21</td>
<td>Equal until 6 y, 1.8</td>
<td>T, double major, T-L</td>
<td>R</td>
<td>Usually will not spontaneously resolve</td>
</tr>
<tr>
<td>Adolescent</td>
<td>10 y-skeletal maturity</td>
<td>80–90</td>
<td>1.5</td>
<td>T, double major, T-L</td>
<td>R</td>
<td>Dependent on age of onset and size of curve</td>
</tr>
</tbody>
</table>

...
without fusion, with repeat surgeries to continue lengthening the spine (81).

Juvenile

The middle group of idiopathic scoliosis is more diverse. Although, by definition, this category includes children over 3 years of age and less than the age when signs of physical maturity start to become apparent, it may actually represent two populations. The first population can be thought of as “late infantile” and the other group as “early adolescent” (81). Additionally, it is important to note that between 10% and 15% of scoliosis in children between 3 and 10 years of age may be associated with spinal cord anomalies, such as syringomyelia (50). Thus, as in any presumed idiopathic scoliosis, underlying conditions must first be ruled out.

The rate of curve progression is variable. Some report that approximately one third of curves will not progress. Of the two thirds that do progress, half will only need orthotic intervention, while the other half will eventually need surgical fixation (81). In the juvenile group, clinical and radiographic monitoring every 4 to 6 months is considered acceptable for curves less than 25 degrees. Curves beyond 20 to 25 degrees will require active intervention, beginning with orthotic management.

Adolescent

Strictly defined adolescent idiopathic scoliosis consists of curves whose onset is after the initiation of puberty. As noted above, there are similarities between children with later onset juvenile scoliosis and adolescent scoliosis. This has led many to postulate that adolescent scoliosis likely begins in childhood but is not detected until the adolescent growth spurt begins (81). Given the prevalence of scoliosis in the population, the U.S. Preventative Services Task Force has recommended that adolescents be screened at school for scoliosis. The Adam’s bending test (see “Clinical Assessment” section above) is used for screening (50). Of those children screened positive, 3% to 9% have curves that progress to a need for active treatment (91).

To properly plan treatment for a patient with adolescent idiopathic scoliosis, it is important to be aware of the size of the curve at onset (which is positively correlated with progression) and the age and level of maturity of the patient (which is negatively correlated with progression) (81). Physical maturity can be assessed using the Tanner scales (which use external signs of sexual maturity), Risser’s sign (which is an indication of the degree of epiphysial closure at the iliac crest), or formal bone age radiographs of the hands. Curves can be observed until they get to 20 (50) or 25 (81) degrees. The patient should have clinical and radiographic follow-up every 4 to 6 months. Patients with curves that are over the limits mentioned above or progress more than 5 degrees during the interval between follow-ups should begin orthotic management. If the curve continues to progress to more than 40 degrees before the patient achieves skeletal maturity, planning for surgery should begin (50,81).

DEVELOPING A PHYSIATRIC TREATMENT PLAN

Is There a Role for Therapeutic Exercise?

Most research has shown that therapeutic exercise is unable to stem the progression of curves in idiopathic scoliosis (92), and the studies that show a lack of effect have shown that this is not simply attributable to problems with compliance with the exercise program (92). Exercise alone has also not been able to correct the posture in patients with Scheuermann’s kyphosis (93). Unfortunately, as is common in research into the effect of exercise modalities in the treatment of disabilities, the literature is neither extensive nor very definitive. A group of researchers, many affiliated with the Scroth Clinic in Germany, have recently formed an organization called the Scientific Society on Scoliosis Orthopaedic and Rehabilitation Treatment (SOSORT) (94). This group is supporting research into the use of neuromuscular exercise to prevent curve progression, much of it based on the method developed at the German clinic. It will be worthwhile to monitor the progress of this research and, depending on the results, consider integrating the techniques into clinical practice. Since the exercises appear to be benign, at least for patients with idiopathic scoliosis, it would be unlikely that they would cause any harm if administered early in curve progression, so long as other, more accepted treatments are not neglected. It would be less wise at this time to consider these exercise techniques to stem progression of scoliosis in more complex pathologies such as those with neuromuscular or inherited multiple system involvement.

Putting aside the controversy above, most experts would agree that there are roles for exercise in the treatment of spinal curvature. Although exercise cannot stop curve progression, no scoliosis is without the risk of secondary morbidities. For example, patients with idiopathic scoliosis (especially under orthotic management) run the risk of core muscle weakness. As such, abdominal and gluteal strengthening exercises should be done to prevent deconditioning and atrophy. In addition, range-of-motion exercise of the hip flexors should be done to prevent contracture (92). In adults with degenerative scoliosis, exercise to improve trunk posture and alignment may prevent the development of pathological curvature. Since many adult spinal conditions are associated with pain, exercise to restore range of motion and strength should begin as soon as pain has been controlled in order to return the person to his or her premorbid functional state (95). Therefore, it should be clear that exercise has a role in the treatment of spinal curvature. That role is primarily to prevent secondary morbidities and to reduce the adverse influence of extraspinal processes (such as asymmetric limb joint contractures) on a progressive spinal deformity.

Physical Modalities

Unfortunately, the literature on the use of physical modalities as a primary treatment to stem the progression of spinal curvature is at least as limited and pessimistic as that on the use
of therapeutic exercise. Traction was the traditional treatment for spinal curvatures before the development of spinal orthoses (96). The theory behind the use of traction was that it would “stretch out” the curves, and the hope in more modern studies was that it could make surgery more effective. However, research in traction in idiopathic curves has not shown any advantageous effect (53). In the 1970s and 1980s, interest in electrical stimulation as a treatment for scoliosis grew. The hope was that electrically induced contraction of the paraspinal muscles on the convex side of the curve would stabilize and straighten the spine (50). Studies have tried both surface and implanted electrodes. Although the early studies suggested a positive effect, more recent longer studies using natural history controls have shown no effect on curve progression (81). Instrumental EMG biofeedback has been even less convincing as an effective treatment option. This is not surprising, considering that direct stimulation has been ineffective. It is hard to develop a sound theory to postulate an effect from spinal manipulation and, in fact, there have also been no convincing studies demonstrating an effect on progression from treatment with spinal manipulation (81,97).

Therefore, similar to therapeutic exercise, the role of modalities is secondary in spinal deformity. Temperature modalities should be used symptomatically, keeping in mind the contraindications against the use of deep heating modalities if spinal hardware is present, the patient is a child under the age of skeletal maturity, or there is a neoplastic process. Massage, electrical stimulation, traction, and/or spinal manipulation should also be considered only for symptomatic treatment, applying the same principles that the physiatrist would use in the treatment of other conditions.

Orthotic Management

It should be apparent by now that the mainstay of nonoperative, physiatric treatment of spinal deformity is spinal orthotics. The goals of the use spinal orthoses are no different than orthoses used in the appendicular skeleton, which are to immobilize and/or correct the position of the joint(s) to be controlled. The simple laws of Newtonian physics, which requires three points of fixation to achieve joint stability, are no different in the spine. However, the very nature and complexity of the trunk make the biomechanics behind spinal orthotic design much more complex since (as the reader may recall from the discussions in previous sections) the spine is a series of semirigid linkages, which are connected by tissues that function like springs. Each link has six degrees of freedom, making control of movement from an external source quite difficult. In addition, these linkages have different stiffness properties at different spinal levels due to the effects of their attached structures, such as the ribs (which increase the stiffness of the thoracic spine to bending by 200%). To further complicate things, the spinal column is suspended within the body. Tissues including muscles, other bones, air, water, viscera, and fat separate the spine from the orthosis attempting to act upon it. These tissues are, of course, not of uniform stiffness. The softer tissues will not transmit as much of the force of the brace to the spine. In some diseases, such as osteogenesis imperfecta, even usually stiff structures such as the ribs cannot safely accept the forces required to be transmitted to the spine without risk of fracture and/or deformation. Lastly, the skin is the point of contact between the patient and the orthosis. Therefore, the forces must be applied through the skin, which has a limited tolerance to pain. Many braces also cover most of the trunk, limiting the ability of the skin to effectively perform its transpiration function, and the orthosis can produce contact dermatitis and/or fungal infections (53).

Given the challenges recounted above, it may be a wonder to the reader that orthoses work at all. Despite the challenges, spinal orthoses are able to apply balanced perpendicular forces to create a (un)bending moment in the curve, focused at the curve vertex. Force must be applied on the concave side above and below the level of the vertex of the curve and it is countered by a force applied on the convex side at the vertex of the curve (53) (Fig. 34-11). In the living human system, the forces not only provide an immediate effect on the spine but also may promote biological tissue “creep” causing adaptation of the spinal tissues to the forces applied, resulting in a long-term effect from the brace (53). As a result of

![Three points of fixation are required to exert a force to counter a curve. (Courtesy of Donald Bliss and Alan Hoofring of the Medical Arts and Publications Branch, NIH.)](image-url)
the geometry of the spinal column, as the orthosis corrects the spine, the correction itself will propagate improved spinal stability, which will also slow progression. It has been shown that decreasing a curve from 45 to 30 degrees will increase the stability of the spine from 20% to 50%. A decrease in curve from 30 to 20 degrees will further improve the stability to 80% (96). In addition to the primary effect of spinal orthoses, they also improve endpoint control, which reduces sway in the neck and/or pelvis and helps bear some of the spine’s transverse load. These effects may also be helpful in controlling spinal deformity (96).

Most of the research on brace effectiveness has been from the idiopathic scoliosis literature, which has shown that it is more effective in preventing curve progression than electrical stimulation or observation of progression (98). There is a dose-response relationship between wearing time and effectiveness (especially in curves >35 degrees), with studies showing that full-time wear time (23 hours/day) is more effective than 8- to 12-hour wear times (99).

In light of the biomechanical complexity, it is not surprising that there are many different bracing options for spinal curvature. Most of the literature on bracing options comes from studies of idiopathic curves, so the physiatrist needs to apply the knowledge and principles gleaned and apply them to other conditions, based on a good understanding of the biomechanics discussed above. However, to help with decision making, Table 34.5 reviews the role of orthotic management in scoliosis of differing etiologies and the type of orthotic which can be considered.

There are two basic types of rigid braces. The “gold standard” is the Milwaukee type cervical-thoracic-lumbar-sacral orthosis (CTLSO). In addition to applying a bending moment at the vertex of the curve, it produces a distraction force on the spine by virtue of being anchored at the chin and the pelvis (53). Meta-analytic studies have shown the Milwaukee to be the “most successful brace in affecting outcome and preventing progression” (99). The obvious question is, if the CTLSO is so good, why are there any other braces? The answer is that the features which give it the increased stability (the chin and pelvic anchoring) make it less comfortable to wear and less cosmetically acceptable.

As a response to the problems of cosmesis and comfort with the Milwaukee, lower profile thermoplastic molded thoracic-lumbar-sacral orthoses (TLSO) have been developed. The Boston brace is a prefabricated unit with added pad components to apply the forces (100). The Wilmington (101) and CAD/CAM braces are custom molded with the force “pads” built into the mold. These orthoses are not effective for curves whose vertex is higher than T8 because they cannot provide a long enough proximal lever arm to achieve effective proximal fixation (81). There have been no definitive studies comparing these orthoses to the Milwaukee, but they have all been shown to be more effective than electrical stimulation or observation. The Charleston Bending (102) (and similar Providence Night-Time (103)) brace is a specialized TLSO where the brace is shaped to overcorrect the curve. This brace is only worn while the patient is sleeping. The hope is that there will be improved compliance with decreased wearing time. It may be effective in long “C” curves but has been shown to be less effective than the Boston brace in such curves (96). However, it may be more effective in lower (thoracolumbar and lumbar) curves (104). Some groups (105,106) are experimenting with rigid orthoses with decreased quantities of materials and claim to get similar results as the Boston type braces. Whether these styles will be more effective and/or more comfortable is yet to be determined.

Recently, a new nonrigid TLSO has been developed. The SpineCor is considered by its inventors to be a “dynamic” orthosis (107). It is made up of a series of elastic bands and pelvic girdle and is to be worn 20 hours/day with two separate 2 hour rest periods out of brace. It is less restricting but more complicated than the rigid thermoplastic TLSOs. There is only limited research on this device to date. The available research is limited to the group which developed the brace.

### Table 34.5 Role of Orthotic Management in Spinal Deformity of Different Etiologies

<table>
<thead>
<tr>
<th>Curve Etiology</th>
<th>Role of Orthotic Management</th>
<th>Type of Orthosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Idiopathic</td>
<td>Correct and prevent progression</td>
<td>Rigid with three-point fixation</td>
</tr>
<tr>
<td>Myopathic/lower motor neuron</td>
<td>Limited, since it cannot stop progression</td>
<td>Rigid</td>
</tr>
<tr>
<td></td>
<td>Ambulatory patients—slow progression</td>
<td>Soft (and must take care that orthosis does not decrease pulmonary function)</td>
</tr>
<tr>
<td></td>
<td>Seated patients—comfort and support</td>
<td></td>
</tr>
<tr>
<td>Upper motor neuron</td>
<td>Slow progression in seated patients (likely to restrict function in standing patients too much)</td>
<td>Rigid</td>
</tr>
<tr>
<td>Congenital bony dysfunction</td>
<td>Slow or stop progression in long flexible curves</td>
<td>Soft</td>
</tr>
<tr>
<td>Ligamentous dysfunction</td>
<td>Insufficient literature; might be able to control curve but effect would be limited by deformation of intervening structures. Consider limiting use for support and comfort</td>
<td></td>
</tr>
<tr>
<td>Degenerative</td>
<td>For comfort and support, will not likely affect curve</td>
<td>Rigid or semirigid</td>
</tr>
</tbody>
</table>
Their initial study was limited to children with idiopathic curves between the ages of 6 and 14, with initial curves of 15 to 50 degrees. The majority of the curves were less than 30 degrees at diagnosis. Only 29 patients completed 2 years of follow-up after treatment. The results were not compared to natural history controls. The results, however, are promising, with decreased curves in 55% of the patients, stability in 38%, and progression in only 7%. This same group has continued to follow an expanded cohort, again with promising results but not in comparison to other treatment or bracing options (108). (Fig. 34-12 shows examples of each of the brace archetypes.)

**SURGICAL TREATMENT OPTIONS**

Extensive discussion of surgical techniques, approaches, and instrumentation is beyond the scope of this chapter. The reader who is interested in more details of surgical treatment of spinal curvature is urged to reference an orthopedic surgical textbook focused on spine surgery, many of which can be found in the reference list for this chapter. Physiatrists need to have a sufficient understanding of spinal surgery to advise and advocate for their patients both preoperatively and postoperatively. This section focuses on the general principles behind surgery for spinal deformity and archetypical examples of spinal instrumentation.

The goal of most spinal surgeries is to realign the spine as much as possible with rigid surgical instrumentation and then allow the body to heal, fixed in a position of better alignment. This occurs by the postoperative fusing of the vertebrae. The surgeon must balance the need to fuse the spine for stability with the need to preserve as much motion as possible (109). For stability, fusions are preferably done beyond the ends of the curve at points within the stable zone (defined by Harrington as the area between two lines drawn upward and perpendicular to the pelvis beginning at the sacral pedicles) (9). In some conditions, there are additional considerations. For example, due to the loss of motor control in neuromuscular pathologies, the surgeon may choose to start the instrumentation and fusion at the pelvis. In congenital bony scoliosis, the surgeon must not only stabilize the spine but also correct the imbalance in bony growth. In conditions where the bone stock is abnormal, such as in osteogenesis imperfecta, the surgeon needs to consider how to best stabilize his/her instrumentation until good bony fusion is achieved.

Spinal instrumentation can apply only a limited range of forces. In the vertical plane, hardware can distract the concave side of a curve and/or compress the convex side of the curve. Distraction becomes more effective as the curve gets larger, whereas compression becomes more effective as the curve gets smaller. Translational forces can also be applied at the vertex of the curve in a manner similar to the way a spinal orthosis works, but without the complications of the intervening tissues. In kyphosis, cantilevered (like a springing diving board) forces can be applied to promote extension at the proximal portion of the curve. Torsional forces can also be applied to attempt to derotate the vertebrae by applying asymmetric force to the hardware anchors at a given spinal segment (109).

The Harrington system of rods, anchored by hooks on each end, was the first successful spinal instrumentation system. At first, one rod was placed on the concave side to distract the curve. As the procedure matured, a second rod was added on the convex side to compress the curve. The Harrington system only produces the vertical forces of compression and distraction. Since it is anchored only at the top and bottom of the curve, it is not as stiff as more modern systems and, thus, more prone to hardware failure and requires longer postoperative stabilization. Cotrel DuBosset instrumentation is the archetype of the category of hardware developed to address the stiffness issue of the classic Harrington rods. It has interlinked rods with hooks at multiple levels and is quite stiff. Luque rectangular instrumentation was one of the first systems with an option to apply a derotational torsion force. It is composed of two L-shaped rods, which form a rectangle encompassing the curve. It is attached by wires wrapped around and anchored to the left and right lamina at each segmental level. It is also stiff but can fail due to

![Milwaukee](Image1.png) ![Boston](Image2.png) ![Charleston](Image3.png) ![SpineCor](Image4.png)

**FIGURE 34-12.** The archetypes of spinal orthotic options. (Courtesy of Milwaukee-Bertil Holmberg, Boston-Boston Brace International, Inc., Charleston-Charleston Bending Brace Foundation, SpineCor-Spine Corporation.)
wire breakage. Halm-Zielke instrumentation is one archetype of the newer systems which are fixed at each segmental level by left and right pedicle screws with short linked rods connecting each spinal segment. It not only is quite stiff but also can more easily control curves with complex segments owing to the flexibility in how each rod is connected to the next. However, it is a more complex system with the increased surgical challenge of safely installing the pedicle screws (Fig. 34-13).

The Halm-Zielke system is also representative of a surgical trend toward an anterior approach to stabilization. Although recent studies have suggested greater curve correction (110,111), it is unclear whether this difference is functionally significant and worth the potentially increased morbidity associated with opening the chest cavity and possibly damaging the sympathetic nerve chain (112). However, surgical techniques are constantly advancing: morbidity from anterior approaches may decrease as a result of thoroscopic techniques (113), and morbidity from posterior approaches may decrease as a result of percutaneous techniques (114). Surgeons have also had a renewed interest in fusionless techniques, primarily in growing children who would lose height and chest cavity size if fused at a young age (115–117).

**CONCLUSION**

Physiatric management of spinal curvature requires a solid understanding of spinal anatomy, physiology, and biomechanics. This knowledge should be applied to understanding the way various disease mechanisms may affect the spine. Applying the principles of comprehensive history taking and physiatric examination with objective measurements, including appropriate radiographs, should allow physiatrists to sufficiently understand the diagnosis causing their patients’ spinal curvature and the pathophysiology of that condition and how it is affecting the spine. With this knowledge and an understanding of the role of physiatric modalities and surgery in the treatment of the patient, the physiatrist should be able to develop an appropriate rehabilitation program to help control and reduce the patient’s spinal curvature and optimize his or her quality of life.

**REFERENCES**


Upper Extremity Soft-Tissue Injuries

Upper extremity pain is a leading cause of physician visits in the United States. This is particularly prevalent in sports and the workplace, contributing significantly to disability and lost productivity. Along the spectrum of potential conditions, musculoskeletal soft-tissue injuries are the most common. Despite a wealth of investigations on soft-tissue injury, there remains an incomplete understanding of what these injuries are and the sequence of events leading to their development. A detailed understanding of the relevant functional anatomy and potential mechanisms of injury allows for the most accurate diagnosis and is directly related to the likelihood of successful treatment. This chapter describes a classification of soft-tissue injury, identifies potential mechanisms of injury, and presents a model for successful rehabilitation. Injuries specific to the shoulder and elbow are considered.

The soft tissues constitute a diverse group of connective tissues, including ligament, tendon, muscle, cartilage, fascia, synovium, articular cartilage, adipose tissue, and intervertebral disc. Collectively, connective tissues function as support for the soft tissues, transmit and distribute mechanical forces, and act as a conduit for neurovascular structures. They are composed of collagen fibers; ground substance, primarily in the form of proteoglycans and water; and specialized cellular components. Although the connective tissues share common structural features, each has unique cellular and biomechanical properties and, therefore, demonstrates individual responses to injury and healing (1). Loose connective tissue, for example, is poorly organized and typically seen in subcutaneous tissue, small joint capsules, and fascial borders. Tendons and ligaments are composed of dense, organized connective tissue, with a larger proportion of collagen fibers arranged according to their biomechanical requirements (2) (Fig. 35-1). Tendons, in particular, are subjected to great tensile stress because they function to transmit force from muscle to bone. The tensile strength of tendon is generally double that of its associated muscle (3,4). This likely accounts for the greater incidence of injury at musculotendinous junctions. There are numerous factors influencing the integrity of soft tissues, including age, sex, temperature, body weight, exercise, nutrition, drugs, immobilization, injury, and systemic illnesses such as inflammatory arthritis and the collagen vascular disorders.

Soft tissues are susceptible to failure under conditions of stress and strain. The most common mechanisms of injury include acute trauma and repetitive overuse or overload. Musculotendinous structures are especially vulnerable to failure from sudden overloading, as with forceful muscular contractions, particularly when weakened as a result of concurrent illness (connective tissue disorders) or medications (steroids). More commonly, repetitive overuse leads to an insidious onset of pain, inflammation, and ultimately, structural failure. In the context of persistent or uncontrolled stress, a cycle may occur in which structural maladaptations develop in the damaged tissue, setting the stage for further injury and chronic inflammation (Fig. 35-2).

The three phases of soft-tissue healing include the cellular response to injury; repair and regeneration, as immature collagen is laid down; and scar remodeling, a process that may continue for years (5). Understanding this sequence of events has direct implications on the success of treatment and rehabilitation. The type of injury, age, vascularity, nutrition, genetic and hormonal factors, and activity level influence successful healing of soft tissues. Acute injuries typically have a sudden onset, are associated with a classic inflammatory reaction, and tend to follow a predictable course toward resolution. Chronic injuries (duration longer than 3 months) are often marked by an insidious onset of pain, less intense inflammation, progressive functional impairments, and a tendency toward reinjury. Advanced age is associated with decreased collagen synthesis and impaired tendon healing. Exercise improves both the mechanical and structural properties of tendon as opposed to inactivity, which favors collagen degradation and decreased tendon strength.

The nature and classification of soft-tissue injury, as well as the injury response, are specific to the involved structure (1). Ligament sprains have been classically defined as grade I, II, and III based on the degree of tissue damage and separation (Table 35-1). Grade I sprains demonstrate negligible loss of structural integrity, display minimal signs of inflammation, and generally recover quickly and completely. In grade II sprains, partial rupture of the ligament is associated with significant pain and inflammation. Functional recovery generally occurs within 4 to 6 weeks; however, pain is often experienced for months after the injury, and the tensile strength of the involved tissue is reduced up to 50% (6). Grade III injuries are often associated with prolonged healing time, chronic instability, susceptibility to recurrent sprains, and impaired proprioception of the involved joint.

Tendons are susceptible to injury through the same mechanisms of tensile overload and repetitive overuse (7,8). Tendons are particularly vulnerable to failure when tension is applied quickly or obliquely, the tendon is tense before trauma, the
attached muscle is maximally innervated, the muscle group is stretched by external stimuli, and the tendon is weak in comparison to its muscle (9). Tendon injuries, or tendinopathies, can be classified according to a sequence of overlapping pathologic conditions: inflammation, degeneration, and rupture (Table 35-2). Paratenonitis refers to inflammation of the paratenon. When the structure is associated with a synovial lining, the condition is described as tenosynovitis. The term tendinitis refers to injuries and inflammation specifically involving tendon. Tendinosis describes a chronic process of intratendinous degeneration and atrophy, minimal or no inflammation, and loss of structural integrity, potentially leading to tendon failure (10–12). Inflammation of the paratenon may occur concomitantly with tendon atrophy, referred to as paratenonitis with tendinosis. A functional classification of traumatic tendinitis is particularly useful because the degree of disability correlates well with the extent of injury (Table 35-3). This grading system also provides objective parameters for following treatment and rehabilitation.

Muscular injuries are particularly common in sports. These are typically classified as contusions, strains, avulsions, and delayed-onset muscle soreness. Contusions result as a direct blow and are graded as mild, moderate, or severe based on the degree of soft-tissue swelling, motion restriction, and functional impairment. Intramuscular contusions tend to be more severe, resolve slowly, have more extensive scarring, and are susceptible to developing myositis ossificans (13–16). Muscular strains result from overstretching or peak contraction of the musculotendinous unit, particularly during eccentric muscular contractions. These injuries tend to occur more commonly at the musculotendinous junction. The classification is similar to that of muscular contusions.

<table>
<thead>
<tr>
<th>Grade</th>
<th>Signs and Symptoms</th>
<th>Functional Implications</th>
</tr>
</thead>
</table>
| First degree | Minimal pain and swelling  
No ligamentous instability  
Minimal localized tenderness | No significant loss of function  
Bracing often unnecessary  
Rapid return to activity |
| Second degree | Significant pain and swelling  
Subtle structural instability  
Occasional joint effusion  
Marked swelling and hemarthrosis  
Structural instability | Protective bracing indicated  
Significant activity limitations  
Susceptible to recurrence  
Immobilization indicated  
Surgery may be required  
Prolonged functional limitation |
| Third degree | | |

FIGURE 35.1. Tendon: structural characteristics.

FIGURE 35.2. Repetitive soft-tissue injury cycle.
The injured muscle generates reduced peak tensile loads and contractile forces, up to 50% of its strength. Contractile forces begin to recover as soon as 1 week after an acute strain; however, there is often prolonged and sometimes permanent loss of complete contractile ability (17–19). Therefore, the functional ability of a patient recovering from a muscular strain may be significantly limited, and premature return to activity may increase the risk for further injury. Delayed-onset muscle soreness typically occurs within the first 24 to 48 hours after an intense bout of exercise that often involves repeated eccentric muscular contractions. Both inflammatory and metabolic mechanisms have been proposed for muscle damage in this condition (20,21). In most cases, the condition is self-limiting but occasionally requires activity modification and anti-inflammatory medication.

It is useful to approach the treatment and rehabilitation of soft-tissue injuries through a conceptual framework that begins with establishing the most anatomically correct diagnosis and ultimately returns the patient to normal athletic or occupational performance (22–26) (Table 35-4). Protecting the injured site and controlling pain and inflammation set the stage for an active therapy program geared toward improving the flexibility, strength, and endurance of the damaged tissue and entire kinetic chain. This also takes into account other factors that may be associated with the development and persistence of symptoms (Table 35-5).

### SOFT-TISSUE INJURIES OF THE SHOULDER

Two primary factors make the shoulder joint particularly susceptible to soft-tissue injuries. First, the surrounding soft tissues constitute the main support system for the upper extremity. Second, the small glenoid fossa allows for a large range of motion that often permits excessive mobility, thereby straining the soft tissues and stressing the joint. Injuries of the shoulder can be broadly classified as those that result from more acute processes, such as direct trauma, and those that occur from more repetitive tasks, such as bursitis and tendonitis. Overall, shoulder injuries are so common that they are second only to back injuries in disability costs associated with lost time (27). Understanding the complex anatomy and kinesiology of the shoulder is essential for accurate diagnosis and appropriate treatment.

### Anatomy of the Shoulder

The shoulder is composed of many joints, including the scapulothoracic joint, which is considered a functional joint. The glenohumeral joint is a synovial joint lined by the glenoid labrum, which provides a large contact surface to the glenoid fossa. Despite the large labrum, the humeral head comes into contact with only about one third of the glenoid fossa at any one time (28). The capsule of the glenohumeral joint is divided into three functional bands that are considered ligaments, aptly named the superior (SGHL), middle (MGHL), and inferior glenohumeral (IGHL) ligaments. Additional support is provided by the coracohumeral ligament (CHL) originating on the coracoid and inserting into the greater and lesser tuberosities. The acromioclavicular (AC) joint is another synovial joint made up of the distal aspect of the clavicle and the

### TABLE 35.2 Classification of Tendon Injuries

<table>
<thead>
<tr>
<th>Injury</th>
<th>Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paratenonitis or tenosynovitis</td>
<td>Inflammation of the paratenon with associated pain, swelling, and tenderness</td>
</tr>
<tr>
<td>Tendinitis</td>
<td>Inflammation of the tendon with associated vascular disruption and inflammation</td>
</tr>
<tr>
<td>Tendinosis</td>
<td>Intratendinous atrophy and degeneration with a relative absence of inflammation; a palpable nodule may be present over tendon</td>
</tr>
<tr>
<td>Paratenonitis with tendinosis</td>
<td>Acute inflammation superimposed on chronic tendinitis</td>
</tr>
<tr>
<td>Partial or complete rupture</td>
<td>Acute inflammation is often superimposed on chronic inflammation with tendinosis</td>
</tr>
</tbody>
</table>

### TABLE 35.3 Tendon Injuries: Functional Scale

<table>
<thead>
<tr>
<th>Grade</th>
<th>Symptoms</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Pain after exercise, subsiding within 24 h</td>
</tr>
<tr>
<td>2</td>
<td>Minimal discomfort with exercise, without activity limitation</td>
</tr>
<tr>
<td>3</td>
<td>Pain that interferes with exercise</td>
</tr>
<tr>
<td>4</td>
<td>Pain interfering with activities of daily living</td>
</tr>
<tr>
<td>5</td>
<td>Rest pain that interfering with sleep</td>
</tr>
</tbody>
</table>

### TABLE 35.4 Rehabilitation of Soft-Tissue Injuries

- Prescribe activity-specific training
- Correct functional deficiencies
- Improve strength, flexibility, general conditioning
- Address structural deficiencies
- Promote healing
- Control pain and inflammation
- Establish correct diagnosis

### TABLE 35.5 Overuse Injuries

#### Intrinsic variables
- Age
- Flexibility imbalance
- Muscle imbalance/weakness
- Anatomic malalignment
- Genetic predisposition

#### Extrinsic variables
- Training errors/poor technique
- Environmental factors
- Equipment factors
acromion and is supported by the coracoacromial ligament, the AC ligament, and the coracoclavicular (CC) ligament (composed of two smaller ligaments—the conoid and the trapezoid). Motion at the AC joint requires not only translation but also rotation for smooth movement of the shoulder. The last synovial joint involved in the shoulder is the sternoclavicular joint. The joint is bordered by the medial aspect of the clavicle and the manubrium of the sternum. There are four ligaments surrounding the joint: the anterior and posterior sternoclavicular ligaments, the costoclavicular ligament, and the interclavicular ligament.

The muscles of the shoulder and shoulder girdle can be divided into two major groups: those that stabilize the scapula and those that attach to the humerus. The stabilizers include the trapezius, levator scapula, rhomboids, serratus anterior, and pectoralis minor. These muscles allow for the stability of the shoulder girdle and provide a foundation for movement and force generation that is passed along the trunk into the arm for functional use. The muscles that attach to the humerus include the rotator cuff muscles (supraspinatus, infraspinatus, teres minor, and subscapularis), deltoid, teres major, pectoralis major, coracobrachialis, biceps brachii, and latissimus dorsi. These muscles provide the arm with most of its motion.

**Kinesiology of the Shoulder**

Range of motion of the shoulder is accomplished by glenohumeral and scapulothoracic motions. The first 30 degrees of abduction is initiated by the deltoid muscle followed by a 2:1 ratio of movement, with the glenohumeral joint responsible for 120 degrees and the scapulothoracic motion supplying the additional 60 degrees. The humerus, however, needs to be in an externally rotated position to be able to obtain full abduction; otherwise, the tuberosity on the humerus impinges on the undersurface of the acromion.

Several authors have evaluated the kinematics of the scapula and its interaction with the humerus. Borstad and Ludewig (29) looked at scapular motion in both symptomatic and asymptomatic individuals. Through electromagnetic tracking, they were able to evaluate scapular tipping and internal rotation. They noted that in symptomatic individuals there was less upward rotation at lower humeral elevations and increased tipping at higher elevations when compared to controls. These same scapular changes were also noted in patients with multidirectional instability (30). There does not appear to be consistent findings with glenohumeral motion (31).

The muscles noted previously can be divided into functional groups. For example, internal rotation is accomplished by the subscapularis, latissimus dorsi, anterior fiber of the deltoid, pectoralis major, and teres major. External rotators include the infraspinatus, teres minor, and posterior fibers of the deltoid. Adductors include the deltoid, supraspinatus, trapezius, and serratus anterior. Adduction is accomplished by the subscapularis, infraspinatus, teres minor, pectoralis, latissimus dorsi, and teres major. Flexion of the arm involves the pectoralis major, biceps brachii, and anterior deltoid. Extension is accomplished by posterior deltoid, teres major, and latissimus dorsi. Some muscles may contribute to motion based on the initial position of the humerus. For example, if the humerus is in a flexed position, the pectoralis may assist in early extension to the neutral plane.

**Disorders of the Shoulder**

**Contusion and Myositis**

Contusions are a result of blunt trauma to the soft tissue and may be difficult to differentiate from tears because both involve a significant amount of injury to the muscle fiber. Contusions can be divided into intermuscular and intramuscular types. Diagnostically, ultrasound or magnetic resonance imaging (MRI) may be helpful in determining the location of the injury. This may have clinical value because contusions and hematomas occurring between the muscles tend to disperse more easily along the fascial plains and result in an earlier return to function. Intramuscular lesions take longer to resolve because they are confined, and the inflammatory response is greater. These injuries also have a higher incidence of compartment syndrome (rare) and myositis ossificans associated with them. Myositis ossificans is the invasion of calcium and bony islands within the muscle. This process is usually accompanied by erythema, swelling, and a significant amount of pain. Radiographs early in the process may not reveal the ossification. If a compartment syndrome is suspected, release and drainage are critical to prevent long-term injury to the neurovascular structures.

**Instability**

Laxity of the shoulder joint varies tremendously between individuals. A certain amount of laxity or movement of the joint is normal; this translation is needed for the joint to have the mobility and function that it needs. Laxity is also noted to be present in most asymptomatic adolescents and decreases with age (32). Instability as a pathologic process should only be explored if there are symptoms that correlate to the degree of motion noted at the joint. The labrum contributes to about 20% of the glenohumeral stability with the joint loaded (33). Anterior stability is maintained not only through the forces at the labrum and rotator cuff muscles but also along the capsule particularly at the MGHL and IGHL. The anterior forces when the arm is abducted 90 degrees are mainly counteracted by the IGHL. At angles closer to the body, the MGHL plays a larger role. Most dislocations occur in an anterior plane when the arm is abducted, thereby indicating that the area most likely to be involved in a traumatic dislocation is the IGHL (34). Posterior translation of the joint is significantly less common. The positions of the humerus that place the joint at greatest risk are flexion, adduction, and internal rotation. In this position, failure of several soft tissues is still required for dislocation to occur. The ligamentous capsule on the posterior aspect of the glenohumeral joint is underdeveloped; therefore, the support appears to be a combination of the anatomic position of the glenoid and the nearby SGHL, IGHL, and CHL along with the rotator cuff muscles themselves. Inferior translation is noted in the neutral
position by the sulcus sign. As stated earlier, a certain degree of laxity is normal. The counteractive force of inferior translation in the neutral position is mainly provided by the SGHL and the CHL. As the arm is abducted, the resistive force at the IGHL is increased and is related to the amount of abduction. If the arm is also internally rotated, additional force is placed on this ligament. Superior translation is clinically insignificant because of the bony structure of the acromion.

Diagnosis of instability relies on the history and reproduction of symptoms with specific physical maneuvers. An apprehension maneuver is typically completed in the supine position, and anterior translation-type forces are applied to the humerus with the arm abducted to 90 degrees and externally rotated. A positive test is noted with pain and fear of subluxation expressed by the patient (Fig. 35-3). A relocation maneuver is also helpful to confirm the diagnosis. This is accomplished by a posterior force applied on the humerus with a reduction in symptoms or additional motion allowed in external rotation with a reduction in pain. Diagnostic tests that may be helpful include a scapular Y film and an axillary view for anteroposterior location of the humerus in relation to the glenoid.

**Acromioclavicular Joint Sprain**

AC joint sprain usually occurs from a fall on an outstretched arm or a direct blow to the shoulder. Pain over the joint and increased joint mobility is frequently noted and may help to guide the diagnosis. In addition, there is pain noted on crossed adduction from compression of the joint and in the end range of abduction from the pull of the deltoid and trapezius. Typically, the diagnosis can be made clinically, but in some cases, an x-ray with a weighted view may be helpful. If gapping of the joint is seen, it can support the diagnosis. However, it should be noted that a negative x-ray does not preclude the diagnosis. AC joint strain and separations are divided into grades. Grade 1 is defined as pain at the joint; however, the ligaments are intact, and there is no subluxation of the joint. Grade 2 is movement of the joint related to a tear in the AC ligament but not in the CC ligaments. Grade 3 indicates a tear through both AC and CC ligaments. Grades 4 to 6 are defined by displacement of the clavicle posteriorly, superiorly, and anteriorly, respectively. In addition, grade 6 displacement involves entrapment of the distal clavicle in the surrounding muscles. Typically, grades 1 to 3 are treated conservatively, whereas grades 4 to 6 require reconstruction. Reconstruction can also be considered in individuals with persistent pain from a grade 3 injury. Treatment of a grade 1 separation would include ice, sling and swath, or resting of the arm for a short period of time. Range of motion is progressed as tolerated, the supporting musculature is strengthened, and weight loading of the joint is avoided for about 6 weeks. A grade 2 injury requires a slightly longer rest period (up to 2 weeks) before returning to range and strengthening activities. Grade 3 injuries do not require surgical reconstruction. Although there may be some controversy regarding nonsurgical management of these injuries, there are several studies that have demonstrated good outcomes with conservative treatment, and one study showed an earlier return to work and play when compared with the surgically treated group (35–38). Grades 4 to 6 are still thought to do better with surgical intervention.

**Bicipital Tendonitis**

Activation of the biceps muscle is noted with elbow flexion and supination of the forearm. There is also some activity with abduction of the arm in an externally rotated position. The long head of the biceps is integrally linked to shoulder function and may assist the rotator cuff muscles with counteracting anterior and superior forces at the humerus. Its origin is at the supraglenoid tubercle and the adjacent labrum. It courses within the capsule and along the bicipital groove between the tubercles and is held in place by the transverse ligament. The biceps tendon is frequently involved with rotator cuff pathology and anterior laxity of the humerus. Tendinitis is common and usually associated with overhead activities. Clinically, palpation of the tendon and pain produced with resisted supination while the elbow is flexed and held against the trunk (Yergason’s test) or with resistance of forward flexion with the elbow extended and supinated (Speed’s test) are indicative of tendinitis (Table 35-6). The tendon has been noted to be swollen, stenotic at the transverse ligament, and frequently hemorrhagic. In addition, adhesions may develop in the area with ongoing inflammation. If the transverse ligament is lax or ruptured, the biceps tendon can sublux. This can be palpated, and the patient frequently reports a snapping sensation. Rupture of the tendon most commonly affects the proximal portion of the long head. Symptoms range from discomfort and weakness to a painful snap. Surgical intervention is considered in young, active individuals. Mariani and Cofield (39) evaluated conservative versus surgical intervention and noted the following: there was no difference in residual pain, elbow motion, strength of elbow extension, forearm pronation, or grip strength; the nonsurgical group usually returned.
to activities, including work, more rapidly, albeit initially at a reduced capacity; and there was usually an associated loss of strength with elbow flexion (10%) and supination (20%), and many individuals were aware of this weakness. Therefore, if the patient is a young and active individual, surgical consultation is generally recommended.

Bursitis
There are multiple bursae in the shoulder. The subdeltoid, subacromial, and subcoracoid bursae separate the rotator cuff and the acromial arch. They often communicate with each other, and the subacromial bursa can become inflamed in the younger athlete by direct trauma or in the older individual with overuse and cuff problems. This syndrome is typically painful and significantly limits abduction of the arm. Movements in other planes are typically unaffected. An injection of lidocaine and crystalline steroid into the subacromial bursa can be helpful as a diagnostic tool and therapeutic treatment for the bursitis.

Impingement
The mechanism of injury involves repetitive motion from an abducted and externally rotated position to an internally rotated flexed position. This causes the tuberosity of the humerus to come under the arch of the acromion or along the coracoacromial ligament. These injuries are common in throwing sports and swimmers. It is usually associated with a painful arc (pain noted with abduction of the arm in the range of 60 to 120 degrees). Impingement signs include Hawkins’ maneuver, with the arm flexed, slightly abducted, and internally rotated, or Neer’s sign, with the arm in full flexion at 90 degrees, internally rotated while resisting flexion (Fig. 35-4; see Table 35-6). These tests may help with the clinical diagnosis but are not specific to impingement. In addition, impingement is frequently associated with tendonitis of the supraspinatus and biceps tendons as well as degenerative changes in the rotator cuff. Neer (40)

---

**TABLE 35.6** Tests Used in the Diagnosis of Shoulder Pain

<table>
<thead>
<tr>
<th>Maneuver</th>
<th>Description</th>
<th>Implication</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sulcus sign</td>
<td>Inferior translation force at the humerus with retraction of the soft tissue above the humeral head</td>
<td>Laxity in the inferior capsule⁴</td>
</tr>
<tr>
<td>Apprehension</td>
<td>Patient supine with arm abducted to 90 degrees, externally rotated, and elbow flexed; anterior translation placed at proximal humerus</td>
<td>Anterior instability</td>
</tr>
<tr>
<td>and relocation</td>
<td>Relocation with posterior force at proximal humerus with reduction in pain and greater external rotation</td>
<td></td>
</tr>
<tr>
<td>Yergason’s</td>
<td>Elbow is flexed and forearm supinated</td>
<td>Bicipital tendonitis</td>
</tr>
<tr>
<td>Speeds’</td>
<td>Arm is flexed with elbow straight and supinated; resistance is applied to arm flexors</td>
<td>Bicipital tendonitis</td>
</tr>
<tr>
<td>Hawkins’</td>
<td>Arm is abducted to 90 degrees and internally rotated, causing pain</td>
<td>Impingement of the supraspinatus tendon</td>
</tr>
<tr>
<td>Neer’s</td>
<td>Arm is flexed to 90 degrees and internally rotated; patient resists downward force</td>
<td>Humerus slides superiorly and causes impingement</td>
</tr>
<tr>
<td>Painful arc</td>
<td>Pain noted on abduction from 60 to 120 degrees</td>
<td>Supraspinatus tendonitis</td>
</tr>
<tr>
<td>Drop arm</td>
<td>Arm placed in full abduction and patient asked to lower arm slowly; arm drops from midabduction</td>
<td>Supraspinatus full-thickness tear</td>
</tr>
<tr>
<td>Clunk</td>
<td>Translation of the humerus over the edge of the labrum</td>
<td>Glenohumeral instability or possible labral tear</td>
</tr>
<tr>
<td>Adson’s</td>
<td>Pulse is palpated, and a drop in the pulse may be noted during inhalation, coupled with rotation of the head activating the scalene muscles</td>
<td>Possible thoracic outlet</td>
</tr>
<tr>
<td>Roos’s</td>
<td>Hands repetitively opened and closed with arms abducted and elbows bent increasing vascular demand</td>
<td>Possible thoracic outlet</td>
</tr>
</tbody>
</table>

⁴A certain degree of laxity is considered normal in individuals. It is best to compare side to side for significant differences.
described three stages of impingement. The first is noted to have edema and hemorrhage along the supraspinatus insertion. The changes usually occur in younger individuals (12 to 25 years of age) and are typically reversible. Stage II is associated with fibrosis, thickening of the coracoclavicular and bony changes of the acromion. The last stage is most commonly seen in people older than 40 years of age and is associated with partial or complete cuff tears. Biomechanically studies have been done looking at scapular motion, specifically dyskinesia in the setting of impingement. Scapular dyskinesia is defined by Kibler (41) as the loss of control in the motion of scapular with respect to external rotation and retraction. This results in an anterior tilt of the glenoid, loss of maximal rotator cuff muscle activation, and reduced rotation of the acromion. All of these factors contribute to clinical impingement. His work on impingement and muscle coupling, which will be discussed later in the chapter, has contributed to effective rehabilitation of the shoulder.

**Rotator Cuff Tendonitis and Tears**

The two main categories of rotator cuff tears are related to their mechanism of injury: traumatic and degenerative tears. Traumatic tears can happen at any age, but older people are more susceptible because of underlying degeneration. McNab (42) has defined an area in the cuff that is noted to have less vascularization, which may contribute to some of the changes noted. As we age, the cuff thins and becomes frayed. There is a 54% incidence of tears within the cuff in asymptomatic individuals older than 60 years (43). Calcific tendonitis is also a risk factor for cuff tears and frozen shoulder. This occurs more commonly in diabetic patients. The cuff can be evaluated with diagnostic ultrasound, arthrography, and MRI. Ultrasound is the least sensitive and can miss incomplete tears. MRI with T2-weighted images is highly specific for a full-thickness tear, whereas its sensitivity is less for partial-thickness tears. The use of arthrography can sometimes demonstrate extravasations of contrast in a partial-thickness tear, often missed with other imaging techniques. Symptoms frequently associated with rotator cuff tendinitis and tears are pain at the site along the tuberosity, night pain, exacerbation of the symptoms with lying on that shoulder at night, pain along the lateral aspect of the arm toward the insertion of the deltoid, and pain with overhead activities. A painful arc is present, and sometimes a drop arm test (pain on lowering the arm causing the individual to drop the arm rapidly) is noted. Manual strength testing can be done to isolate the supraspinatus, performed along the plane of the scapula (about 30% anterior to the frontal plane) with the arm internally rotated. Significant weakness will be noted if there is a full-thickness tear, and mild or no weakness is noted with partial tears, secondary to muscular compensation.

**Labral Tears**

People with labral tears often report a history of trauma or falling on an outstretched arm. The pain is typically over the anterior shoulder, and clicking is sometimes noted with abduction and external rotation. A clunk test may be positive with translation of the humerus over the edge of the labrum and may indicate instability of the shoulder or a labral tear. O’Brien’s test loads the anterior labrum with the arm extended at the elbow, flexed and internally rotated at the shoulder in the Neer’s position. It is then brought into adduction and a positive finding is noted with pain as the arm crosses the midline. Sensitivity of the O’Brien’s test appears to be limited in isolation (approximately 63%); however, in combination with other tests, the sensitivity is better (45). One study revealed a 90% sensitivity of the O’Brien’s test, and compression test and apprehension tests were all positive. In isolation, the most sensitive and specific test for labral tears appears to be the biceps load test II described by Kim et al. (44). The test is performed in the supine position, and the arm is elevated to 120 degrees and externally rotated to its maximal point, with the elbow in the 90 degrees flexion and the forearm in the supinated position. The patient is asked to flex the elbow while resisting the elbow flexion by the examiner. The test is considered positive if the patient complains of pain during the resisted elbow flexion. The test is negative if pain is not elicited or if the pre-existing pain during the elevation and external rotation of the arm is unchanged or diminished by the resisted elbow flexion. Diagnostic testing with an MRI may reveal a superior labrum anterior posterior (SLAP) lesion and signifies the direction of the tear. MRI arthrogram increases the sensitivity for finding a labral tear and is the test of choice if a labral tear is suspected.

**Adhesive Capsulitis or Frozen Shoulder**

The pathophysiology of the frozen shoulder can be either idiopathic or associated with internal derangement such as trauma, tendinitis, and tears. Prolonged immobilization following one of these injuries is a significant risk factor. However, a frozen shoulder can occur in the absence of these diagnoses. Middle-age women and diabetic patients appear to be at a higher risk for spontaneous idiopathic adhesive capsulitis. The loss of range is multiplanar, with external rotation and abduction being the most affected. Clinically, the diagnosis should be suspected with progressive loss of range and diffuse pain despite conservative treatment measures. The syndrome is typically painful to treat and has a natural recovery that can be prolonged lasting up to 2 years. Early intervention may help with range in the pain-free zone (Codman’s or pendulum exercises), joint mobilization with or without ultrasound, and a non-steroidal anti-inflammatory drug (NSAID) trial (Fig. 35-5). An intra-articular or subacromial injection of anesthetic and steroid may help (46). In addition, oral steroids may play a role in the acute management of adhesive capsulitis. According to Buchbinder and the Cochrane reviews, oral steroids provide significant short-term benefits in pain, range of movement of the shoulder, and function in adhesive capsulitis but the effect may not be maintained beyond 6 weeks (47). Distention of the capsule can be attempted but requires larger volumes to be injected (25 to 35 mL is the typical joint capacity). Initial infusion of lidocaine (3 to 5 mL) may make distention easier.
and less painful. Manipulation under anesthesia can be done but is generally a treatment of last resort.

**Neurovascular Compromise**

**Plexus Injury**

Plexus injuries typically occur as a result of traction-type injuries. Sports-related injuries to the plexus are typically neuropraxic and have a good prognosis. There is some controversy regarding whether some of these neuropraxic injuries occur at the nerve root or plexus level. They are most commonly seen in football, wrestling, and hockey. These injuries are commonly referred to as "stingers" or "burners" and usually affect the upper part of the plexus. Weakness may last minutes. More direct trauma may result in a more severe injury to part or all of the plexus, and prolonged weakness may occur. Avulsion injuries, which are usually traumatic and require significant force, are associated with a poor prognosis.

**Axillary Nerve**

The axillary nerve is derived from the fifth and sixth cervical roots. It arises from the posterior cord of the brachial plexus. It courses in front of the subscapularis muscle and passes along the lower part of the capsule along with the posterior circumflex vessels. The nerve then travels through the quadrangular space. This space is bordered by the teres major inferiorly, the long head of the triceps medially, the neck of the humerus laterally, and the subscapularis and teres minor superiorly. This nerve is at greatest risk with dislocations of the shoulder and fracture of the surgical neck of the humerus. Testing for the sensory branch can aid in the evaluation for dislocation to detect nerve involvement. The sensation along the lateral aspect of the shoulder is innervated by the axillary nerve. This area of sensation can also be assessed after the dislocation is reduced to evaluate possible nerve involvement. Prognosis is good unless there has been extensive damage to the nerve, and recovery can take 6 to 12 months.

**Suprascapular Nerve**

The suprascapular nerve typically arises from the upper trunk of the brachial plexus. It runs along the middle scalene and under the trapezius. It passes over the superior aspect of the scapula along with the artery and vein and rests in the suprascapular notch. These structures are frequently held in place by a transverse ligament. This ligament may be ossified, therefore creating a foramen. The nerve then sends a branch to the supraspinatus muscle and wraps around the spinous process of the scapula at the supraglenoid notch before innervating the infraspinatus. This nerve is sensitive to entrapment at the notch and along the spine of the scapula. These syndromes are more commonly seen in throwing athletes, volleyball players, weight lifters, swimmers, and martial arts participants. Symptoms may include diffuse shoulder pain, impingement, weakness, atrophy, or scapular dysfunction. Electrodiagnostic studies may be helpful in the diagnosis.

**Long Thoracic Nerve**

The long thoracic nerve arises from the roots at C5, C6, C7, and C8. It passes along the base of the neck and along the inside of the serratus anterior. The serratus is essential for stabilizing the scapula and maintaining contact of the scapula with the thorax. It assists in protraction and upward rotation of the scapula. The nerve can be injured by either a direct blow to the neck or thorax or by repetitive and prolonged pressure at the base of the neck. This can be seen with backpacks or briefcases being carried on one shoulder. Patients typically complain of pain in the neck and shoulder region and weakness or heaviness with overhead activities. Physical findings reveal scapular winging of the inferior medial border with resisted protraction. Recovery is generally good because of the multiple nerve roots contributing to the nerve.

**Thoracic Outlet and Vascular Impingement**

Thoracic outlet is a controversial diagnosis. Its presence is thought to result from neurologic or vascular compromise of the structures at the base of the neck as they pass into the upper extremity. Entrapment may be noted at the scalene muscles where the trunk of the plexus passes through the muscle, along the first rib in the subclavicular space, or in the infraclavicular space along the coracoid and the pectoralis minor. The symptoms may be vague arm pain, numbness and tingling usually along the medial aspect of the arm, pain along the neck and shoulder region, weakness that is diffuse and does not follow a nerve root distribution, and muscle spasms. Provocative testing includes Adson’s maneuver, which is positive when there is a drop in pulse palpated distally while the head is rotated to the involved side, resulting in contraction of the scalenes. Roos’ maneuver, for more functional induction of symptoms, demonstrates increased vascular demand of the arms in an overhead position (see Table 35-6). A positive test is noted with reproduction of symptoms. Although this diagnosis is controversial, the treatment of the symptoms is rehab based. The exercise prescription for treatment of the neurogenic
symptoms is usually focused on myofascial release, stretching of the scalenes and pectoralis muscles, mobilization of the first rib, and postural exercises in addition to the general guidelines for shoulder rehabilitation. Cervical x-rays can be obtained to assess the presence of a cervical rib, which can contribute to the development of this syndrome. Electrodiagnostic studies may help to rule out other entrapment neuropathies and to support this clinical diagnosis.

**General Principles of Shoulder Rehabilitation**

There are two general types of muscle coupling that occurs in the shoulder. One is the co-activation of muscles that allows for control of the extremity through a range. Examples of this type of coupling would be the upper trapezius, lower trapezius, and serratus for the rotation of the scapula. The second type is agonist/antagonist activity that allows for the transfer of forces. There is inhibition of the antagonist group of muscles with contraction or activation of the agonist group. Examples of this type include the pectoralis and latissimus muscles versus the infraspinatus and deltoid for control of internal rotation.

Nonoperative rehabilitation of the shoulder uses these principles for progression of an exercise program. Muscular substitution and inhibition particularly of the serratus anterior and upper trapezius muscles may lead to scapulothoracic dysfunction. In addition, position and movement of the scapula (with measurements of the “scapular slide”) will illustrate shoulder dysfunction, often seen early in shoulder injuries (48–50). Scapular motion may be restored with “scapular clock” exercises and include elevation, depression, protraction, and retraction. Scapular elevation/depression coupling utilizes trapezius, levator scapula, and rhomboids. Scapular protraction/retraction coupling is accomplished with rhomboids, mid trapezius, serratus, and pectoralis. Establishing good scapular position is essential for proceeding with a progressive exercise program.

Once scapular kinetics has been restored, glenohumeral range can be advanced and a progressive strengthening program should be prescribed. Strengthening of the scapular muscles, including the middle and lower trapezius, rhomboids, and serratus anterior, can be accomplished with either open- or closed-chain exercises. Open-chain exercises are more common but produce more shear forces along the joint. Frequently, these exercises are done with resistance. They help to facilitate scapulothoracic and glenohumeral control. Closed-chain exercises have less tensile stress along the capsular ligaments. They can be initiated with axial loading on a stable surface such as a table or an unstable surface such as a ball. They should be started with the arm in a pain-free position, typically slightly flexed with the trunk leaning on the arm (Fig. 35-6). This can then be advanced to loading the arm in an upright position on a table with the arm either flexed or abducted. Strengthening of the rotator cuff may follow scapular stabilization and may focus on particular deficits noted within the cuff (Figs. 35-7 and 35-8).

Eventually, pyelometrics should be added to the exercise regimen. Most sporting activities require plyometric use of a muscle with eccentric loading just before concentric forceful contractions. Because this activity is seen in many shoulder functions, such as throwing, golf, racquetball, and baseball bat swings, it is essential to advance the athlete’s program to include this in order to prevent reinjury. Medicine ball work with a rebounder will reproduce this type of exercise. Other examples are dynamic challenges that can be applied by a therapist in a stable position such as a tripod.

Rehabilitation of the shoulder should encompass a functional assessment of how the shoulder interacts and reacts with the rest of the body. Treatment may need to extend beyond the shoulder area. The legs are the foundation and are required for stability as well as to initiate the translation of energy from the larger muscles of the trunk to the shoulder girdle and eventually to the arm. There are several key points in the rehabilitation of the shoulder of a throwing athlete.
It should start with pelvic control over the planted leg, effective hip and trunk extension, scapular control (especially retraction), and normal glenohumeral rotation. Definition and use of the kinetic chain in the principles of rehabilitation are essential for an optimal outcome. Examination of the patient’s posture, spinal alignment, and proximal hip strength should be included in the initial assessment for a patient with shoulder dysfunction.

Postoperative rehabilitation of the shoulder follows the same general guidelines. Immediate postoperative care includes aggressive treatment of pain, swelling, and range if and when the patient has been released to move the joint. Most surgeries for labral or capsular repair can begin with the range of motion early and progress to isometric exercises within the first few weeks. Isometrics are not allowed in rotator cuff repairs in the early phases of rehabilitation. These can be introduced usually at the 6-week postoperative period. Progression of the exercise occurs when the range of motion is 65% to 70% of the opposite side. The exercise protocols described by Kibler (51) include diagonal plane proprioceptive neuromuscular facilitation (PNF) exercises for scapular and glenohumeral motion in this phase of rehabilitation. In addition, closed-kinetic chain exercises are used to progress scapular retraction and depression. Once normal patterns are noted, strength can progress to open-kinetic chain exercises, pyelometrics, and trunk or core stabilization.

**SOFT-TISSUE INJURIES OF THE ELBOW**

Soft-tissue injuries adjacent to the elbow are relatively common. Proper biomechanical functioning of the elbow is critical to ensure coordinated movement patterns and force transmission throughout the upper extremity. Assessment of the elbow likewise requires a functional understanding of the distal and proximal components in the upper extremity kinetic chain. In the context of soft-tissue elbow injuries, it is not unusual to observe motion restriction due to joint contracture and musculotendinous shortening or weakness involving the shoulder, cervical spine, wrist, and digits. The success of rehabilitation is often based on normalizing function throughout the involved extremity.

The elbow joint is structurally complex. Although commonly referred to as a hinged joint, various articulations contribute to its dynamic movement. The ulnar-humeral joint is the primary articulation and contributes greatly to joint stability. It is a typical, single-axis, hinged joint allowing for 150 degrees of flexion. A few degrees of hyperextension is often seen, primarily in women, related to relative ligamentous laxity. The radiohumeral and proximal radioulnar articulations allow for axial rotation and can be considered pivot-type joints (52). Up to 75 degrees of forearm pronation and 85 degrees of supination are conveyed through these articulations. The radial and ulnar collateral ligaments, olecranon fossa, and anterior joint capsule act as the main static stabilizers of the elbow. Dynamic stability is conveyed through neuromuscular control about the elbow, with the biceps, triceps, pronator teres, and supinator providing the main stabilizing function. The radial, median, and ulnar nerves all become relatively accessible around the elbow and are susceptible to injury through direct trauma, repetitive overuse, compressive forces, or entrapment. These syndromes are often overlooked in the context of chronic elbow pain.

### Lateral Epicondylitis

Lateral epicondylitis is the most common presentation of lateral elbow pain. The annual incidence of the disorder is between 1% and 3% in the general population (53–55). Although often referred to as tennis elbow, 95% of reported cases occur in individuals other than tennis players. It is more frequently seen in occupations requiring repetitive upper extremity activities and particularly those involving computer use, heavy lifting, forceful forearm pronation and supination, and repetitive vibration. In athletics, racquet sports are most commonly associated with lateral epicondylitis; however, the injury is also seen in golf, baseball, and swimming (52,56). In tennis, the incidence appears to be between 30% and 40%. In general, tennis players with symptoms are older, play more frequently, and tend to be more skilled (57,58). This lends support to the mechanism of injury being one of repetitive overuse.

The clinical presentation of lateral epicondylitis is often clear (59,60). Most patients present with localized tenderness over the common forearm extensor tendon insertion at the lateral epicondyle, often extending into the extensor mass. Less commonly, there is discomfort over the radiohumeral joint and annular ligament. Pain is generally reproduced with resisted wrist and middle finger extension and with gripping activities. Passive wrist flexion with elbow extension often generates symptoms. Flexibility and strength deficits are often seen in the wrist extensor and posterior shoulder muscles.

The differential diagnosis of lateral epicondylitis is most relevant in the context of chronic, nonresponsive symptoms.
CHAPTER 35  |  UPPER EXTREMITY SOFT-TISSUE INJURIES  |  918

Other soft-tissue considerations include radiohumeral bursitis, annular or lateral ligament strains, impingement of the synovial fringe, and myofascial strain involving the extensor forearm muscles. Neurologic considerations include cervical radiculopathy, radial tunnel syndrome, entrapment of the posterior interosseous nerve, and carpal tunnel syndrome. Radiocapitellar and proximal radioulnar arthritis are not uncommon in older adults. In the skeletally immature, Panner’s disease, involving fragmentation of the capitellum, can be observed.

The site of pathology in most cases involves the insertion of the extensor carpi radialis brevis and less commonly the extensor carpi radialis longus and anterior extensor communis tendons. Due to repetitive overuse of the extensor forearm muscles, microscopic failure at the enthesis has been described. Pathologic findings have included mucinoid degeneration and granulation tissue in the subtendinous space. Nirschl (61) has reported extensively on the topic and proposes a theory that repetitive overuse is likely to cause vascular compromise resulting in anoxia and angiofibroblastic hyperplasia. These pathologic changes in the tendon have been defined as tendinosis (8).

The treatment of lateral epicondyliitis follows the same principles for soft-tissue rehabilitation: control of inflammation, promotion of healing, reduction of abusive forces, and improvement of soft-tissue flexibility, strength, and endurance. In the context of acute injuries, resting the involved extremity and avoiding activities that reproduce symptoms are critical. Occasionally, a wrist cock-up splint may be necessary. NSAIDs and cryotherapy are useful in the acute phase. Therapeutic modalities are often administered to relieve pain and inflammation and possibly promote soft-tissue healing. The most common applications include cryotherapy, high-voltage galvanic stimulation, transcutaneous electrical nerve stimulation (TENS), ultrasound, phonophoresis, and iontophoresis (64–66). Acupuncture has been used successfully to treat soft-tissue injuries (67–69). Although the mechanisms of action remain unsubstantiated, acupuncture likely has both analgesic and anti-inflammatory effects. The literature also remains inconclusive with respect to the clinical efficacy of many of the therapeutic modalities. Klaiman et al. (64) compared phonophoresis, using a corticosteroid gel, to ultrasound in a double-blinded design in a variety of soft-tissue conditions, including lateral epicondyliitis. Although both modalities had a positive impact in terms of pain reduction and pressure tolerance at the site of maximal tenderness, there was no significant difference between the two therapies. Used conscientiously, however, in the initial stages of therapy or in the context of an acute exacerbation of symptoms, modalities may facilitate rehabilitation and allow the patient to advance to a more active program.

Corticosteroid injections are most appropriately used after failure of more conservative therapies or if acute symptoms are impeding other rehabilitation strategies (70). Intratendinous injections should be avoided, activity should be restricted for at least 7 to 10 days after the procedure, and patients should be well informed of the potential side effects, including skin depigmentation and atrophy, structural weakness, and in rare circumstances, tendon rupture. Injections can be repeated at monthly intervals but not more than three times in a year.

Therapeutic exercise is a necessary and critical component of soft-tissue rehabilitation (71–73). The injured site must be prepared to accept the demands of daily functional activities; otherwise, reinjury is likely to occur. Soft-tissue manipulation is often a valuable technique to improve tissue extensibility and promote tendon healing. Progressive loading of the wrist extensors is initially performed with passive stretching, advancing to progressive resistance training as the musculotendinous unit is able to accept greater demands without an increase in pain (Figs. 35-9 and 35-10). Strengthening is generally initiated with isometrics, advancing to concentric and ultimately

---

**TABLE 35.7**

<table>
<thead>
<tr>
<th>Differential Diagnosis of Lateral Elbow Pain</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Articular</strong></td>
</tr>
<tr>
<td>Radiohumeral joint disease</td>
</tr>
<tr>
<td>Radioulnar joint disease</td>
</tr>
<tr>
<td>Periostitis</td>
</tr>
<tr>
<td>Fibrillation of the radial head</td>
</tr>
<tr>
<td>Ectopic calcification</td>
</tr>
<tr>
<td>Panner’s osteochondrosis</td>
</tr>
<tr>
<td><strong>Soft tissue</strong></td>
</tr>
<tr>
<td>Radiohumeral bursitis</td>
</tr>
<tr>
<td>Annular and lateral ligament sprains</td>
</tr>
<tr>
<td>Rupture of the common extensor origin</td>
</tr>
<tr>
<td>Myofascial pain syndrome</td>
</tr>
<tr>
<td><strong>Neurologic</strong></td>
</tr>
<tr>
<td>Radial tunnel syndrome</td>
</tr>
<tr>
<td>Posterior interosseous nerve compression</td>
</tr>
<tr>
<td>Cervical radiculopathy</td>
</tr>
<tr>
<td>Carpal tunnel syndrome</td>
</tr>
</tbody>
</table>

---

**FIGURE 35-9.** Passive wrist extensor stretching.
eccentric exercises (74). The emphasis is on the wrist extensors, flexors, and forearm supination and pronation (Figs. 35-11 to 35-14). It is also important to identify structural deficits in the kinetic chain extending from the cervical spine to the distal extremity. Weakness in the shoulder external rotators and inflexibility in the internal rotators should be identified and corrected. Aerobic conditioning should be maintained during the course of rehabilitation. This likely has a positive impact on soft-tissue healing.

Returning to sport or occupation requires that abusive forces be modified and activity specific exercises initiated. Factors pertaining to equipment and training are well recognized as contributors to overuse injuries in sports. Poor backhand stroke mechanics in tennis, for example, have been associated with lateral epicondylitis. Corrective strategies in tennis have included switching to a two-handed backhand, using the largest comfortable grip, lowering string tension below 55 lb, switching to a more flexible and lightweight frame, playing with a softer ball, and wearing a counterforce forearm brace (60). The counterforce brace is believed to decrease the force of contraction from the forearm extensors and redistribute the force distal to the insertion at the lateral epicondyle (75,76). Injuries in the workplace necessitate a detailed ergonomic evaluation. Workstation modifications, education, and postural retraining
are important interventions. A graded return to sport reduces the likelihood of reinjury. Tennis players are advised to begin hitting forehand strokes, advancing to the backhand and volley provided symptoms remain under control.

Surgery is generally reserved for chronic conditions that are unresponsive to conservative management and are associated with significant limitations in functional performance. The approach most commonly involves debridement of degenerative granulation tissue adjacent to the extensor carpi radialis brevis insertion and partial release of the common extensor aponeurosis (77,78). Postoperative rehabilitation employs the principles previously discussed and often requires 4 to 6 months before resumption of full activity.

Medial Epicondylitis

Overuse injuries of the medial elbow typically involve the musculotendinous structures of the forearm flexor and pronator group, particularly the flexor communis and pronator teres (52,56,79). It is commonly seen in the throwing athlete as well as in racquet sports and particularly among more competitive players. Tennis players who use a topspin forehand and serve seem to be the most susceptible to medial elbow pain. Golfer’s elbow refers to the same clinical syndrome. Patients complain of localized pain at the medial elbow with occasional radiation into the forearm. The symptoms are typically reproduced with resisted wrist flexion and forearm pronation as well as passive wrist extension and supination. In chronic situations, soft-tissue contractures may develop with loss of full elbow extension and supination.

The differential diagnosis of medial elbow pain includes medial collateral ligament sprain, muscular strain of the forearm flexors or pronator teres, anterior interosseous nerve entrapment (pronator teres syndrome), cervical radiculopathy, and thoracic outlet syndrome. Pain reproduction with a valgus stress while the elbow is extended 20 degrees is suggestive of ulnar collateral ligament sprain. This is more commonly seen in the throwing athlete and in severe cases may be associated with a triad including ulnar collateral ligament strain, posteromedial capsuloligamentous strain, and ulnar neuropathy. Localized weakness in the flexor pollicis longus, the flexor digitorum profundus to the index and middle fingers, and the quadratus lumborum, in the absence of sensory deficits, suggests entrapment of the anterior interosseous nerve.

The treatment protocol for medial epicondylitis is similar to that already described for lateral epicondylitis. Managing pain and restoring full, pain-free range of motion at the elbow is necessary before initiating strengthening. Restriction in shoulder internal rotation should also be addressed. Resistance training focuses on the wrist flexors and pronator teres, advancing to eccentric training as symptoms permit. Counterforce bracing may assist in redistributing forces distal to the medial elbow and help to facilitate rehabilitation. Corticosteroid injections should be used with caution around the medial elbow because of the proximity of the ulnar nerve. Activity and sport-specific training focuses on biomechanical adaptations that may predispose to injury. In tennis, the athlete may be encouraged to hit more of a flat serve and use less wrist snap in the forehand stroke. Similarly, the golfer may be encouraged to use a larger and looser grip, to strike the ground less forcefully, and improve trunk rotation. Surgery is pursued only for intractable cases.

Olecranon Bursitis

The location of the olecranon bursa, being superficial to the triceps insertion and olecranon, makes it susceptible to injury from acute trauma and repetitive friction. In sports, it is more commonly seen in wrestling, weight lifting, and gymnastics. It is also observed in skateboarders and rollerbladers who do not wear elbow protection. The differential diagnosis includes posterior elbow impingement syndrome, triceps tendinitis, and fracture of the olecranon. Patients with systemic arthropathies often develop bursal reactions but rarely is the elbow involved alone. Patients with acute olecranon bursitis present with pain and swelling, normal elbow range of motion, and pain reproduction with passive elbow flexion beyond 90 degrees. It is important to rule out an infected bursa and fracture of the olecranon or trochlea. Chronic conditions may not be painful, and the bursa often feels thickened and boggy. Treatment includes pressure relief over the elbow, cryotherapy, anti-inflammatory drugs, and compression, particularly in the acute phase. A suspected infected bursa should be aspirated and analyzed before instituting appropriate antibiotic therapy. Aspiration may also be indicated in a severely swollen sterile bursa, and in chronic conditions, surgical excision may be necessary.

Triceps Tendinitis

Triceps tendinitis generally occurs as a result of repetitive overuse or extension overload at the elbow. The injury is seen in boxers, pitchers, weight lifters, bowlers, and gymnasts. The injury, associated with inflammation and microtrauma, typically occurs at the insertion of the triceps to the olecranon process. Pain is reproduced with resisted elbow extension and passive elbow
flexion with shoulder forward flexion. In the throwing athlete, loss of flexibility may be seen in elbow flexion and shoulder internal rotation. In atypical cases, patients will report a snapping sensation at the elbow caused by subluxation of the medial head of the triceps over the medial epicondyle. This is distinguished from subluxation of the ulnar nerve at the medial elbow owing to the absence of sensory symptoms in the hand.

The treatment of triceps tendinitis requires activity modification; therapeutic modalities including cryotherapy, ultrasound, or phonophoresis; electrical stimulation, and soft-tissue manipulation. Therapeutic exercise emphasizes flexibility training through the kinetic chain with a focus on elbow flexion and shoulder flexion and internal rotation. Progressive resistance training and a graded return to activity are addressed, monitoring for any increase in symptoms. An orthosis, limiting full extension of the elbow, may be useful during this period of reconditioning.

**PERIPHERAL NERVE SYNDROMES**

**Ulnar Neuropathy**

The ulnar nerve, originating from the medial cord of the brachial plexus, is well protected in the upper arm behind the medial head of the triceps and, except for associated trauma in the presence of humeral shaft fractures, is relatively spared of injury in the upper arm. As the nerve emerges into the cubital tunnel, bordered by the medial epicondyle, olecranon, medial collateral ligament, and aponeurosis of the flexor carpi ulnaris, it becomes more superficial and susceptible to trauma and repetitive overuse. In cubital tunnel syndrome, the nerve is particularly vulnerable as it passes through the fibrous arch connecting the two heads of the flexor carpi ulnaris (80–82). The arch is taut with elbow flexion and situations of repetitive or sustained elbow flexion increase the risk of nerve irritation and injury.

The patient with cubital tunnel syndrome may present with vague elbow pain but typically reports pain, paresthesias, and numbness in the fifth and ulnar aspect of the fourth fingers. Even though the branches of the ulnar nerve to the flexor carpi ulnaris and flexor digitorum profundus originate below the cubital tunnel, these muscles are often spared in cubital tunnel syndrome, as the nerve fascicles to the intrinsic hand muscles tend to be most affected. Percussion at the medial elbow (Tinel’s sign) will often reproduce symptoms in the hand. Maintaining the elbow in a fully flexed position for up to 5 minutes (elbow flexion test) may also induce symptoms. Patients can demonstrate impaired light touch sensation and two-point discrimination along the fifth and ulnar aspect of the fourth fingers. In more advanced cases, muscle atrophy can be seen in the intrinsic hand muscles and particularly the first dorsal interosseous. Weakness in the adductor pollicis can be demonstrated by a positive Froment’s sign, in which patients will compensate with the flexor pollicis longus when attempting to squeeze a piece of paper between the thumb and the index finger.

Electrodiagnostic studies are useful in confirming ulnar neuropathy at the elbow in addition to ruling out other diagnostic considerations including cervical radiculopathy,plexopathy, and distal entrapment neuropathies at the wrist. Imaging of the elbow should also be considered to rule out a compressive lesion. Clinical examination will assist in ruling out other sources of soft-tissue pain at the medial elbow including medial epicondyritis, ulnar collateral ligament strain, and olecranon bursitis.

Mild cases of ulnar neuropathy can be effectively managed with anti-inflammatory medication, curtailing activities that result in direct pressure or frictional stress to the medial elbow, and if necessary, an elbow pad. Surgical management, consisting of decompression and transposition of the ulnar nerve, is generally reserved for more severe cases manifesting with persistent pain and neurological compromise.

**Radial Neuropathy**

Above the elbow, the radial nerve is most susceptible to injury as a result of compression at the level of the spiral groove or following mid humerus shaft fractures. In both instances, the triceps muscle is generally spared, and motor/sensory deficits can be observed in a predictable distribution below the level of the lesion. As the radial nerve emerges under the extensor carpi radialis longus at the lateral elbow, it divides into the superficial radial and posterior interosseous nerves. Radial tunnel syndrome, or compression of the posterior interosseous nerve, has been best described as the nerve that enters through a fibrous arch (arcade of Frosh) between the two heads of the supinator muscle (62,63,83,84). This is generally believed to be a repetitive motion disorder, involving extension of the elbow and pronation of the forearm, and should be considered as a cause of persistent lateral elbow pain, although diffuse forearm pain is common. While patients will often demonstrate tenderness at the origin of the common extensor tendon, maximal pain can be elicited with palpation over the supinator, and unlike tennis elbow, resisted forearm supination is more painful than resisted wrist and middle finger extension. Sensory symptoms should be absent as the posterior interosseous nerve is a pure motor nerve and weakness in wrist and finger extension can be observed, particularly in chronic cases.

Electrodiagnostic studies are commonly unrevealing in radial tunnel syndrome. Radial motor conduction slowing may be observed across the lateral elbow, and electromyography may reveal denervation and polyphasic motor unit potentials in the distribution of the nerve, generally sparing the extensor carpi radialis longus and brevis, as well as the supinator. Diagnostic lidocaine injections can also be used to attempt to distinguish lateral epicondyritis from radial tunnel syndrome. Surgical exploration and decompression of the nerve should be reserved for resistant cases unresponsive to conservative strategies including physical therapy, injection therapy, and acupuncture.

**Median Neuropathy**

The median nerve is also uncommonly injured in the upper arm, generally susceptible in cases of blunt trauma or fracture.
As the nerve emerges along the distal medial humerus, it passes under the ligament of Struthers, a fibrous band that attaches to an anomalous supracondylar spur. This represents a potential, albeit, rare site of median nerve compression. Lesions at this level may be associated with pain at the medial elbow, distal sensory symptoms in the distribution of the median nerve, and weakness below the level of the lesion, including the pronator teres. This feature often distinguishes it from pronator teres syndrome, in which median nerve entrapment occurs as the nerve penetrates the two heads of the pronator teres and is compressed against the flexor digitorum sublirum (85,86). Since the median nerve branches to the pronator teres arise proximal to the muscle, weakness in forearm pronation is not typically observed.

The anterior interosseous nerve, a pure motor branch of the median nerve, emerges below the two heads of the pronator teres and may also be susceptible to injury. Compression of the nerve against accessory muscle involving the flexor pollicis longus and the flexor digitorum superficialis has been reported. Distinguishing clinical features include the absence of sensory symptoms, preservation of forearm pronation strength, and weakness in the flexor pollicis longus and flexor digitorum profundus (digits 1 and 2), and pronator quadratus. The loss of tip-tip pinch between the index finger and the thumb (pinch sign) demonstrates this pattern of weakness. Nerve conduction studies are typically normal but may demonstrate prolonged latency of the anterior interosseous nerve in the forearm. Electromyography may be helpful in diagnosis, particularly distinguishing pronator teres and anterior interosseous nerve syndromes. As previously discussed, failure of conservative management may necessitate surgical decompression. Distal entrapment of the median nerve at the wrist will not be discussed here.

REFERENCES


Cumulative Trauma Disorders

Cumulative trauma disorders (CTDs) are injuries to the nervous and/or musculoskeletal system attributed to repetitive physical tasks, poor sustained posture/biomechanics, vibrations, and/or repetitive forceful exertions (1). It has also been called repetitive strain injury, overuse injury syndrome, cumulative movement disorder, and repetitive motion injury. CTD is an umbrella term that includes multiple specific diagnoses and nonspecific conditions affecting primarily the upper limbs, shoulder, neck, and lower back. Furthermore, within each of these body regions, there can be multiple types of CTDs that involve pathology of the tendons, joints, muscles, or nerves. Among the most common CTDs are lumbar back pain, carpal tunnel syndrome (CTS), epicondylitis, neck pain, and de Quervain’s tenosynovitis. Although the workplace is the most common setting for development of CTD, these disorders can arise in any setting with prolonged exposure to the aforementioned mechanical stressors.

The mainstay of CTD management involves ergonomic interventions to correct underlying postural and mechanical causes. Treating symptoms without modifying the workplace is the primary reason for recurrence of CTD and magnifies the overall economic burden caused by these disorders. In addition to ergonomic interventions, treatment can include physical therapy, orthotic devices, anti-inflammatory medication, muscle relaxants, pain medications, steroid injections, and surgical interventions in select cases. Psychosocial factors play a significant role in disability resulting from CTD and must be addressed as well.

EPIDEMIOLOGY AND ECONOMIC COSTS

According to the U.S. Bureau of Labor Statistics, there were 357,160 total recorded cases of musculoskeletal disorders in 2006 that resulted in time away from work. CTDs account for approximately 56% to 65% of all occupational injuries (2). There are several diagnoses that are due to cumulative trauma, with the most common being low back pain (LBP), CTS, and neck/shoulder pain. Repetitive motion injuries as a group had a median of 19 missed workdays (BLS, 2006; http://www.bls.gov/). Based upon these numbers, CTDs are commonplace and constitute a large portion of occupational injuries.

The economic costs of CTDs are enormous. Total U.S. annual cost estimates have ranged from $13 to $20 billion (3). Estimating cost is a challenge since CTDs can encompass varying diagnostic entities from one study to the next. Additionally, it can be difficult to attribute the cause of workplace injuries to repetitive stresses versus acute overexertion over a brief time period since there are no clear cut temporal definitions currently used in the literature.

PATHOPHYSIOLOGY

Although many hypotheses have been put forth, the pathophysiology of CTDs is not completely understood. Cumulative biomechanical stresses can result in tissue alteration in tendons, muscles, joints, and nerves over time. In soft tissues, damage is marked by inflammation, collagen deposition, and tissue contraction, which can in turn lead to pain or loss of motion. One hypothesis to explain the origins of myalgia is that prolonged muscle contractions lead to decreased local blood flow, deoxygenation, and metabolite buildup that manifest as muscle fatigue and soreness (4). The mechanism for tendon-related disorders may be related to inflammation and hypoxia. It has been demonstrated that repeated mechanical stresses release prostaglandin E2 (PGE2) from human fibroblasts in vivo (5). Exposure of rabbit patellar tendons to PGE2 resulted in degenerative changes. Therefore, tendon pathology may result from prolonged or recurring inflammation. Another possible precipitant is hypovascularity, which may predispose tendons to hypoxic degeneration and subsequent tendinopathy. Regardless of cause, tendinopathies have been shown to have disordered collagen arrangement and increased proteoglycan ground substance (6). The most common site of overuse injury is the osteotendinous junction (7). When referring to the general term tendinopathy, it is important to make the distinction between tendinitis, an acute inflammatory process, versus tendinosis, a chronic, degenerative process of the tendon.

PSYCHOSOCIAL

Several studies show that psychosocial factors contribute to disability in CTD, especially in cases involving cervical or lumbar spine pain. CTDs are costly to both employers and employees. They result not only in time lost from work but also in decreased productivity and poor employee morale, which in turn leads to further disability (8). Studies have identified job dissatisfaction (9,10) as a significant risk factor for the
development of LBP. One study analyzed industrial insurance data to demonstrate a positive correlation between low income and divorce with chronicity of LBP (11). Dersh et al. found that 56% of patients with occupational LBP had major depression, while 11% had an anxiety disorder (12).

Although less studied, there is evidence to suggest that psychosocial factors play an important role in the severity or chronicity of CTDs that are unrelated to spine pain. One study used patients’ self-report of upper extremity symptoms to show a positive correlation between depression and severity of carpal tunnel, de Quervain’s tenosynovitis, lateral elbow pain, and trigger finger (13). Another study found job strain to adversely affect successful return to employment in CTS patients (14).

The correlation between significant psychosocial stressors and disability from CTD introduces the question of which issue arises first. A recent epidemiologic study suggested that most patients with occupational LBP found that most psychiatric disorders appear only after the onset of work injury (15). Nonetheless, these factors have been shown in several pain conditions as contributing to chronicity even if they are a secondary process. Overall, there is increasing evidence that these issues should be addressed appropriately for optimal recovery. In evaluating new CTD patients, clinicians should screen for psychosocial factors and psychiatric comorbidities and consult relevant mental health practitioners if indicated.

ERGONOMICS

According to the International Ergonomics Association Council of 2000, ergonomics is “the scientific discipline concerned with the understanding of interactions among humans and other elements of a system, and the profession that applies theory, principles, data and methods to design in order to optimize human well-being and overall system performance.” It requires an understanding of human abilities and the limitations imposed by the work environment, machines, tools, and specific job tasks (4).

Several elements contribute to occupational CTD, including forceful exertions, repetition of a work task, awkward biomechanical postures, and vibration (1). The National Institute for Occupational Safety and Health (NIOSH), a division of the U.S. Department of Health, reviewed the epidemiological evidence for each of these physical work factors for the major CTD subtypes. A brief summary of these findings is presented below. It is important to note that many CTDs will involve a combination of these factors.

Repetition or Prolonged Activities

Repetitiveness is commonly cited as an occupational factor leading to CTD of the upper extremity (8). It can be defined as cyclical or repeated motions requiring contraction of the same muscles to mobilize joints. The NIOSH conducted a 1997 review that found moderate evidence for repetition as a cause of work-related musculoskeletal disorders in neck and shoulder pain, CTS, and hand/wrist tendonitis. It did not find sufficient evidence in LBP or epicondylitis (1).

Forceful Exertions

There is epidemiologic evidence to support the role of forceful exertions in the workplace in contributing to neck, lumbar back, elbow, and hand/wrist pain. Ergonomic design plays a critical role as force requirements may increase based upon several factors including poor body mechanics, high torque or speed of power tools, and friction between objects and the worker (16).

Posture

Improper postural mechanics during the performance of a task appears to play a significant role in the development of CTD, especially those involving the spine and upper extremities. Sustained wrist and forearm flexion-extension or radial-ulnar deviation may induce friction between tendons and adjacent anatomic surfaces. A good example of this is frequent radial deviation of the wrist and subsequent development of de Quervain’s tenosynovitis.

Vibration

Another frequently mentioned risk factor in CTD is vibration (17). Exposure to vibration occurs in a variety of contexts: using power tools, holding an object as it is processed in a machine (e.g., wood in a power saw), or using percussion tools (e.g., hammering a nail). Vibration has been implicated as the cause of Hand-Arm Vibration Syndrome (HAVS).

CUMULATIVE TRAUMA DISORDERS

As the various CTDs of the upper extremity and spine are discussed, emphasis is placed on anatomy, pathophysiology, incidence, clinical diagnosis, and management. There are some common trends in CTDs. In general, the pathophysiologic mechanisms of these disorders are not completely understood. The evidence for cumulative trauma in the pathogenesis of these conditions is discussed, and summary is provided in Table 36-1. Diagnosis of these conditions is typically made primarily based upon clinical history and exam, with supportive studies such as imaging or electrodiagnosis used in equivocal cases. Table 36-2 provides a summary of specific exam maneuvers suggestive of a certain CTD. Management begins with relative rest and workplace modification including ergonomic interventions. Conservative treatments include nonsteroidal anti-inflammatory drugs (NSAIDs), therapeutic heat/cold, and physiotherapy consisting of stretching and strengthening of relevant muscle groups. Other modalities such as myofascial release, deep friction massage, transcutaneous electrical nerve stimulation (TENS), acupuncture, ultrasound, and iontophoresis are not well described in the literature. Steroid injections can be utilized in some conditions when the above treatments fail. Finally, surgery can be performed in refractory cases.
Neck

The cervical spine is a complex structure consisting of eight individual motion segments, beginning with the articulation of the occiput on C1 (atlas) and ending with the articulation between C7 and T1. Cervical nerve roots exit bilaterally through intervertebral foramen to innervate the myotomes and dermatomes of the head, neck, and arms. The cervical musculature controls head movement and provides stability. Pain in the neck area can arise from overwork of the musculature, impingement of cervical nerve roots, or degenerative arthritis of the spine. Neck pain is commonly encountered in jobs requiring prolonged posturing of the neck, forceful exertions, and highly repetitive tasks with static postural loads. Because major neck muscles extend to the shoulder or base of the skull, shoulder pain and headache are commonly associated. Tension neck syndrome involves persistently stiff neck muscles resulting in aching discomfort at the base of the neck, upper back, and suboccipital region.

Poor neck posture has been strongly associated with occupational neck pain and appears to be the most significant factor in work-related neck disorders. Specifically, prolonged neck flexion or extension in combination with arm elevation is problematic. Mayoux-Benhamou and Revel used electromyography (EMG) to demonstrate improved neck muscle efficiency with neutral head position as compared to the flexed or extended position (18). Other studies have found associations between forward head posture and headaches (19), overhead arm tasks and radiating neck pain (20), and simultaneous head extension and arm elevation with neck/shoulder pain (21). The literature has also identified certain occupations involving overhead work with a higher incidence of neck pain. Neck symptoms were reported in 62% of dental hygienists and 66% of sewing machine operators with more than 15 years of experience (22,23). Significant symptoms related to the neck have also been reported in dentists, meat carriers, miners, heavy labor workers, iron foundry workers, and civil servants (24).

Repetition and high exertional loads are important factors in neck pain as well. A cross-sectional study of 82 female assembly line workers who were exposed to repeated intermittent...
neck flexion and shoulder abduction revealed an odd ratio (OR) of 4.6 for any neck/shoulder diagnosis and OR of 3.6 for diagnosis of neck tension syndrome (25). A smaller, but more quantitative, study measured trapezius activity on surface EMG in order to demonstrate that increased frequency of sustained, low muscle activity was positively correlated with neck discomfort in workers (26). For forceful exertions, the trapezius is the major muscle of the neck utilized to carry out work involving high forces. One longitudinal study found that workers with a reduced static trapezius load had less shoulder pain at 2-years follow-up (27). Another study found that postal workers, who carry heavy shoulder bags, consistently reported greater shoulder/neck disability compared to gas meter readers who perform a similar amount of walking as postal workers (28). This was the case after controlling for age, work years, and previous lifting work.

Management of neck pain involves avoiding prolonged static postures and introducing dynamic and varied work tasks (29). Exercise therapy for mechanical neck disorders has short-term benefits in terms of pain and function (30). Although no specific exercises can be clearly recommended over others based upon evidence, a reasonable regimen would include neck retraction exercises, stretching of the anterior shoulder capsule and pectoralis muscles, and strengthening of the trapezius and rhomboids. In general, treatments for neck pain arising from cumulative trauma have not been well studied, and management principles have been extrapolated from general mechanical neck pain studies. Pharmacotherapy can include NSAIDs, muscle relaxants, and opioids for severe pain. Manipulation alone is not helpful but may have benefit in conjunction with exercise therapy (31). There is moderate evidence for a lack of efficacy of botulinum toxin injections in chronic mechanical neck pain according to a recent systematic review (31). The therapeutic value of TENS, biofeedback, and acupuncture remain unclear. For further details regarding cervical spine pain, refer to Chapter 32.

Shoulder
The shoulder is a complex structure that affords great mobility at the expense of stability. Its stability can be divided into static and dynamic components. Statical, the bony glenoid, cartilaginous labrum, glenohumeral ligaments, and joint capsule provide moderate stability. Dynamically, the rotator cuff muscles and biceps tendon function to assist with shoulder stability. The most frequent CTDs of the shoulder are rotator cuff tendinopathy (especially involving the supraspinatus), biceps tendinopathy, and shoulder impingement.

The rotator cuff consists of the supraspinatus, infraspinatus, subscapularis, and teres minor and resides in the subacromial space, which is defined by the acromion, subacromial bursa, and coracoacromial ligament above; the coracoid process at the medial border; and the humeral head below. The rotator cuff plays an especially important role in the case of overhead elevation of the arm, which requires tonic contraction to keep the humeral head anchored in the shallow glenoid fossa (32). This explains why rotator cuff tendinopathy is particularly common in laborers who work with their arms overhead or athletes who throw repeatedly (7).

A disorder that often accompanies rotator cuff tendinopathy is impingement syndrome, in which there is progressive encroachment of the rotator cuff tendons in the subacromial space due to intrinsic or extrinsic sources (33). Intrinsic causes include trauma or degeneration of the rotator cuff with instability or laxity of the shoulder complex. The laxity results in cephalad migration of the humeral head resulting in impingement. Extrinsic causes include bony changes to the acromion, coracoid, acromioclavicular joint or greater tuberosity, cervical nerve root compression, and other systemic conditions, including rheumatic disorders. Acromial morphology (type I flat, type II curved, type III hooked) is another factor that can predispose one to impingement (34). Individuals with a hooked acromion are most likely to develop rotator cuff abnormalities. The positioning of the rotator cuff has also been reported as having a relationship to rotator cuff pathology. Rathbun and Macnab report the “wringing out” phenomenon, whereby a hypovascular region in the supraspinatus tendon is created when the arm is held in adduction (35). Overall, shoulder impingement, a mechanical compression process, often leads to rotator cuff tendinopathy, which can lead to further impingement.

The individual with rotator cuff tendinopathy or impingement syndrome will report pain deep within the shoulder or posteriorly, with referral to the deltoid muscle insertion region. There may also be loss of strength and motion secondary to the pain. The discomfort is worsened by activities at shoulder level or above. Pain will occasionally occur at night while resting on the involved shoulder, perhaps from a concomitant subacromial bursitis. On physical examination, a positive Hawkins’ or Neer’s sign is suggestive of impingement. Hawkins’ test involves passive internal rotation and abduction of the shoulder while the elbow is flexed at 90 degrees. Neer’s test involves forward shoulder flexion until the arm is overhead while simultaneously maintaining light pressure on the acromion. Other tests include diagnostic injection of local anesthetic into the subacromial space, which should provide temporary pain relief and improved ROM in these conditions. The differentiation of rotator cuff tendinopathy and impingement is often difficult. A helpful clinical maneuver to differentiate these conditions is to assess isometric strength testing of the rotator cuff muscles (i.e., resisted external rotation of shoulder or abduction), which should produce moderate to severe pain in rotator cuff tendinitis, but minimal discomfort in a pure impingement syndrome. Further evaluation can include imaging such as x-rays or MRI/MR arthrography, which may show narrowed subacromial space, calcified tendons, or partial tears.

Bicipital tendinopathy is another condition that can result from repetitive strain and overhead reaching. Patients typically complain of anterior shoulder pain localized to the region of the bicipital groove. On examination, a positive Yergason’s or Speed’s maneuver (36) along with pain during active elbow flexion may be important clues. Diagnostic injection with local anesthetic or MRI may confirm this clinical diagnosis. Since other shoulder pathology frequently occurs in concert, the
clinician should also investigate for rotator cuff tendinopathy, impingement, or shoulder instability.

There have been numerous studies illustrating the association between repetitive movements and/or exertional force and shoulder pathology. Chiang et al. found an OR of 1.6 (95% CI 1.1 to 2.5) among fish-processing workers’ repetitive upper limb movement and shoulder girdle pain (37). Another cross-sectional study of fish industry workers revealed that repeated arm elevations and shoulder abductions as documented on videotape were significantly associated with shoulder and neck pain (25). Welch et al. noted a prevalence of 32% incidence of rotator cuff injury in sheet metal workers, with most occurring from overhead duct work (38). Other at-risk occupations for shoulder pain include electricians (39), garment workers (40), hospital workers (41), and construction workers (42). Herberts et al. reported 18% of shipyard welders and 16% of steel plate workers had rotator cuff pathology (especially supraspinatus tendinitis) (43). Overall, rotator cuff injury has been reported as the third most common diagnosis encountered in workers, accounting for 8.3% of cases (44).

In the industrial setting, treatment of shoulder impingement and tendinopathy emphasizes avoiding awkward postures and decreasing overhead work, especially tasks involving shoulder internal rotation. One study found that postural alterations could improve the range of shoulder motion at which pain was experienced, in spite of having no overall effect on pain intensity in symptomatic impingement patients (45). Acute interventions include NSAIDs, ice, and pain management. If persistent, corticosteroid injections into the subacromial space can be utilized. Range-of-motion exercises involving both the glenohumeral and scapulothoracic articulations will decrease the likelihood of asynchronous motion leading to impingement. Strengthening exercises should focus on the rotator cuff muscles and scapular stabilizers.

Elbow

Epicondylitis

The lateral and medial epicondyles serve as the origins of the musculature for the forearm and wrist. Lateral epicondylitis (also termed tennis elbow) and medial epicondylitis (also termed golfer’s elbow) are overuse syndromes of these muscles. Of these two disorders, lateral epicondylitis is more common and usually involves the extensor carpi radialis brevis and less commonly the extensor carpi radialis longus, extensor digitorum communis, or extensor carpi ulnaris. Medial epicondylitis typically involves the pronator teres and flexors carpi radialis and ulnaris.

Epicondylitis begins as an inflammatory reaction within the tendinous origin and progresses to microtears that heal through fibrosis and granulation tissue production. It is common in both sports as well as in occupations such as gardening, dentistry, meat processing (46), cashiers, and carpentry. It is an overuse syndrome precipitated by repetitive contraction of the wrist flexors/extensors or pronators/supinators, such as in hand gripping, twisting or screwing motions, hammering, assembling small parts (47), or swinging sporting equipment. Any contraction of muscle can be either concentric, which results in muscle shortening, or eccentric, which results in muscle lengthening. Between these two contractions, eccentric muscle overload results in higher muscle tension and appears to be more likely to cause injury than concentric contractions (48).

The diagnosis of epicondylitis is based on history and physical examination. Patients with lateral epicondylitis typically report focal lateral epicondylar pain, pain with resisted wrist extension and radial deviation, and possibly weakness of hand grip due to muscle fatigue and discomfort. Medial epicondylitis presents with focal pain at the medial epicondylo, pain with resisted wrist flexion and radial deviation, and possibly decreased hand grip strength. In both lateral and medial epicondylitis, there should be normal elbow range of motion; otherwise, elbow pathology should be investigated. In equivocal cases, further diagnostic workup may include local anesthetic injection or imaging. An anesthetic block in either the lateral or the medial epicondyle will confirm the diagnosis. MRI evaluation showing distinct signal intensity changes and contrast enhancement can be correlated with fibrovascular proliferation and fatty degeneration of the common extensor tendon. Minor signal intensity changes without contrast enhancement correlate with fibrosclerotic degeneration and intratendinous cartilage formation (49). MRI will also be helpful in distinguishing epicondylitis from other conditions such as stress fractures, ligament and tendon injury, and nerve entrapment (entrapment of the posterior interosseous nerve may mimic lateral epicondylitis) (50). Diagnostic ultrasound is an emerging diagnostic tool in epicondylitis that may also provide information about disease severity (51).

Treatment of epicondylitis begins with relative rest and ergonomic improvements in tool or sporting equipment design. Medical care for epicondylitis often begins with activity restriction, icing, and anti-inflammatory medication. In terms of anti-inflammatory medications, both oral and topical medications have shown short-term pain improvement (52). Prior to progressive strengthening exercise, the patient should work toward obtaining pain-free movements in wrist flexion/extension and pronation/supination. Strengthening regimens begin with light concentric exercises for all motions about the wrist and forearm, including grip strength. The program is then advanced to include eccentric exercises as well as more work-specific activities. In a small pilot study, an eccentric exercise program demonstrated significantly reduced symptoms compared with concentric strengthening in lateral epicondylar pain (53). Return to limited work activities can be attempted when strength is 80% of that in the noninvolved side. Work volume should not increase by more than 5% per day (54).

Studies involving local corticosteroids have yielded conflicting results. Many studies have shown improvement of symptoms in the acute period but no differences in outcomes at 1 year. One study actually found 1-year outcomes to be better in patients treated with physical therapy or a “wait and see” approach (55).

Use of a counterforce brace can reduce torque about the extensor digitorum brevis muscle and may be an effective
Trigger Finger (Flexor Tenosynovitis)

The finger flexor tendons of the hand are enclosed by a synovial sheath that extends from the metacarpal bones to the distal interphalangeal (DIP) joints. There is a series of fibrous “pulleys” that tether the flexor tendons to the IP joints, allowing frictionless finger flexion. The pulleys are arranged in a series along the flexor surface of the digits, starting with the A-1 pulley at the metacarpophalangeal (MCP) joint area to the A-5 pulley at the volar plate of the DIP joint. The flexor tendons are surrounded by tenosynovium, which allows smooth movement underneath the pulleys during finger flexion and extension. In trigger finger, there is inflammation and thickening of the tenosynovium or pulleys, leading to tendon nodule formation. These changes compromise the normal smooth gliding motion of the tenosynovium under the pulleys and the involved digit may even become locked in a flexed position. Histologically, fibrocartilaginous metaplasia and synovial proliferation are noted in and about the tendon and pulley (63).

The pathophysiology of trigger finger is thought to be a multimodal process, involving a combination of degenerative changes, inflammation, local trauma, and heredity. Certain hand motions have been implicated. These include prolonged hand gripping, especially with an excessively wide grip, use of pistol-gripped power tools that require frequent finger flexion, and repetitive flexion of the distal interphalangeal joint (DIP) with extension of the proximal interphalangeal joint (PIP) and metacarpophalangeal joints (MTP). Occupational trigger finger has been described in homemakers, manual laborers, sewing machine operators, welders, and processing plant workers (63,64). However, more recent studies have questioned the work-related nature of this disorder, including a British study that suggested that the vast majority of cases are nonwork related (65).

The presenting symptoms of trigger finger vary from a slight clicking to occasional frank locking. “Triggering” refers to a catching of the tendon and occurs more commonly in flexion. Pain, if present, occurs with finger extension and is most pronounced at the site of the nodule or thickened pulley, which is typically at the MTP joint. On physical examination, palpation over the volar aspect of the metacarpal head results in local tenderness. The examiner can feel a prominence or thickening of the flexor tendon. Pain is aggravated by passive extension or active resisted flexion of the involved digit.

The diagnosis of trigger finger is usually quite obvious from physical examination. No further imaging studies are needed. The goal of treatment of the trigger finger is to restore the normal gliding of the tendon through the pulley system within the synovial sheath. The management of this syndrome is threefold: reduce inflammation, decrease local pressure, and make appropriate worksite changes. First-line treatments include rest from provocative activities and immobilization of the affected digit for 4 to 6 weeks using a splint or buddy taping to an adjacent finger. Splinting is encouraged to reduce pressure and tension on the flexor tendon, maintaining the MCP joint at 10 degrees of flexion while permitting active IP joint motion. Evans et al. reported improvement in 73% of patients with splinting alone (66).

If symptoms persist beyond 6 weeks, local steroid injection into the synovial sheath at the level of the A-1 pulley may decrease inflammation and edema within the pulley and the flexor tendon. There are a few small randomized controlled trials demonstrating the efficacy of steroid injections (67,68). There are not clear guidelines as to when these injections can be repeated, but the majority of patients respond to a single injection. Marks and Gunther demonstrated an 84% cure rate with one injection, which improved to 91% with a second steroid injection (69). Surgery is recommended for refractory cases of trigger finger and is generally a simple outpatient procedure.

De Quervain’s Tenosynovitis

The extensor tendons of the wrist are grouped into six compartments. de Quervain’s tenosynovitis occurs when there is inflammation and/or entrapment of the tendons in the first dorsal compartment, namely the extensor pollicis brevis (EPB) and abductor pollicis longus (APL) tendons. The tendons of the EPB and APL run down the radial aspect of the radius coursing over the prominence of the radial styloid before insertion into the thumb. The tendons share a common synovial-lined sheath within an osseofibrous canal.

Disease within the first compartment is precipitated by repetitive wrist motion, especially thumb flexion combined with ulnar deviation of the wrist (70). Typical activities that may precipitate the condition include buffing, wringing, grinding, polishing, sanding, sawing, cutting, and screwdriver use (8). de Quervain’s has been described in waitresses, hospital workers, garment workers, machine operators, and domestic cleaners (71,72).
The patient with de Quervain’s typically complains of pain and swelling about the region of the radial styloid, especially during grasping movements with the thumb. Swelling and crepitus may be palpated along the radial forearm if significant fluid is present in the tendon sheath. Finkelstein’s test involves actively flexing the thumb into the palm while making a closed fist, while having the examiner passively deviate the wrist in an ulnar direction. A positive test is marked by significant discomfort in the region of the radial styloid and maybe more suggestive of EPB than APL dysfunction (73). A positive Finkelstein’s test is suggestive of de Quervain’s but may be positive in carpometacarpal osteoarthritis of the thumb as well. Another useful clinical test is resisted thumb extension at the MCP with the wrist maintained in radial deviation.

The mainstay of treatment involves curtailting activities that place great shear force on the first dorsal compartment. Acute therapies can involve ice application for 15 minutes every 6 hours or NSAIDs. Orthotic devices such as a thumb spica splint or buddy taping the thumb to the index finger may be needed to further limit motion and rest the APL and EPB tendons. Steroid injection into the tendon sheath of the first dorsal compartment may be a useful adjunct during the acute inflammatory stage to quiet the tenosynovitis. McKenzie reported an improvement in 90% of symptomatic patients treated with one to three injections of hydrocortisone into the tendon sheath at 18-month follow-up (74). Later trials have demonstrated the efficacy of steroid injections with few complications and return to activities in 1 to 3 weeks (75,76). In a small randomized trial, the addition of an oral NSAID did not have an additive benefit to steroid injection therapy for de Quervain’s tenosynovitis (77). In chronic cases, steroid injection may be particularly difficult because of stenosis in the tendon sheath. In refractory cases, surgical decompression of the first dorsal compartment may be curative in approximately 90% of cases, with relatively few complications (78).

**Hand-Arm Vibration Syndrome (Vibration White Finger)**

HAVS is a constellation of vascular and neuromuscular symptoms associated with high levels of exposure to vibration. This is typically an occupational disorder that results from both increased duration and intensity of exposure to handling vibrating tools or objects, such as jackhammers, drills, or chain saws. Although it is difficult to quantify the threshold intensity or the amount of operating time needed to cause HAVS, a positive correlation of symptoms with cumulative exposure has been shown (79). Onset of HAVS can occur as rapidly as within 3 months of full-time, daily vibration exposure (80). This disorder is most prevalent in forestry workers, construction workers, stone cutters, and shipyard workers (79,81).

The clinical presentation of HAVS includes vascular, sensorineural, and musculoskeletal symptoms. The vascular component consists of Raynaud’s phenomenon and may be triggered by a cold environment. There is initial digital blanching (associated with vasoconstriction), followed by cyanotic discoloration that eventually resolves into erythema when blood vessels dilate allowing reperfusion. During this bright red phase of reactive hyperemia, there may be significant pain in the wrist, hands, or fingers. Not all patients will experience all three phases, as some only have the initial blanching phase. The duration of these symptoms varies greatly from minutes to hours. Neurologically, symptoms include tingling paresthesias or numbness. Patients may also note intermittent aching pain in their hands. Grip strength may be reduced due to weakness of the finger flexors or intrinsic hand muscles. In chronic cases, skin ulceration may develop, usually at the fingertips.

There are various tests for the diagnosis of HAVS. Physical exam should include two-point discrimination and vibration perception testing, which can uncover sensorineural dysfunction (82–84). Duplex ultrasonography checks the patency of arteries of the upper limbs. Cold provocation testing utilizes the fact that patients with Raynaud’s secondary to HAVS often have prolonged coolness of digits after cold water immersion. Coughlin et al. demonstrated high sensitivity and specificity for cold provocation thermography, in which thermal measurements of the digits are made before and minutes after cold immersion of hands (85). Nerve conduction studies have found sensory nerve conduction velocity slowing in the digits in 36% and across the wrist in 20% of patients in advanced stages of HAVS (86). HAVS patients often have both median and ulnar neuropathies (87). Another objective measure involves recording finger systolic blood pressures (FSBP) at baseline and again after exposure to cold. The cold temperatures result in greater vasoconstriction, which is measured by the FSBP decrease in HAVS patients (88). Poole et al. found a specificity of 90% to 95% and sensitivity of 43.5% to 60% for this technique (89).

Treatment of HAVS begins with limiting exposure to hand vibration, using antivibration gloves and coated tool handles, and ensuring proper technique in handling equipment (90). Avoidance of contact with cold objects or maintaining warm hands in cold environment can decrease bouts of Raynaud’s. Smoking cessation is highly recommended because of its vasoconstrictive effects on peripheral arteries (80). Medications such as calcium channel blockers, nitrates, pentoxifylline, and NSAIDs may help reduce symptoms especially if used in combination (91,92).

**NERVE ENTRAPMENT SYNDROMES**

**Carpal Tunnel Syndrome**

The carpal tunnel is an enclosed space formed by the sheath of the flexor retinaculum and the carpal bones. Within the tunnel lie the median nerve and tendons of the flexor digitorum superficialis, flexor digitorum profundus, and flexor pollicis longus (FPL). When the median nerve becomes entrapped in this space, the resulting numbness, paresthesias, weakness, and atrophy of intrinsic hand muscles are known as carpal tunnel syndrome (CTS).

There are numerous epidemiologic studies associating repetitive movement and/or forceful exertions with CTS. CTS
has also been associated with vibrations and awkward postures, but there is less epidemiologic support for this. Chiang et al. studied fish processing workers and categorized them into three groups based upon levels of repetitiveness and force exerted (37). Prevalence rates of CTS in workers performing tasks with high force and high repetitiveness was 28.6%, those with tasks with either high force or high repetitiveness had a rate of 15.3%, and those with low force and low repetitiveness had a rate of 8.2%. Another study evaluated 56 workers at a single supermarket and found an OR of 8.3 between CTS development in workers with highly repetitive and forceful wrist motions versus those with low exposure (93). In this study, both clinical symptoms and electrodiagnostic criteria established a diagnosis of CTS. Silverstein et al. conducted a broader study of 652 workers across multiple occupations and found an OR of greater than 15 between CTS and highly repetitive and forceful work (defined as force >6 kg) versus low force and repetitiveness (94). Aside from epidemiologic evidence, some studies have analyzed carpal tunnel pressures. Gelberman et al. compared the carpal pressures of 15 CTS patients with 12 controls. This study found normal subjects to have carpal pressures of 2.5 mm Hg at neutral wrist position, which increased to about 30 mm Hg with 90 degrees of wrist flexion or extension. In those with CTS, baseline pressure was 32 mm Hg in neutral position and increased to 94 mm Hg at 90 degrees of wrist extension (95). Another study suggested that compressive forces may lead to persistent impairment of intraneural microcirculation due to mechanical injury to blood vessels (96).

In spite of numerous studies implicating CTS as a CTD, there is some evidence to the contrary. One common flaw to many CTS studies is the failure to control for concurrent underlying medical conditions associated with CTS, including diabetes, obesity, rheumatoid arthritis, or thyroid problems. Acheson et al. studied 297 patients diagnosed with occupational CTS and found a high prevalence of concurrent medical conditions that may have contributed to their carpal tunnel symptoms (97). Finally, there are some studies that have failed to demonstrate a positive correlation between cumulative trauma and CTS. Schottland et al. measured the median nerve latencies in poultry processing workers and found no significant association between slowing of either motor or sensory latencies and repetitive work (98).

The initial symptoms of CTS are, typically, sensory dysfunction in the form of numbness and paresthesias (often described as tingling “pins and needles”). These symptoms occur in the median nerve distribution, affecting the palmar aspect of the thumb, index finger, middle finger, and radial half of the ring finger. Involvement of the entire hand, however, is not uncommon. Symptoms classically worsen during the night, which may be related to increased edema secondary to a nocturnal increase in tissue pressures. A “flick sign,” in which the patient vigorously shakes the hand in an attempt to relieve symptoms, has been reported in as high as 93% of those with CTS (99). Weakness may follow sensory symptoms after a variable amount of time and may lead to thenar muscle atrophy.

The diagnosis of CTS is based largely upon history and physical examination and can be further supported by electrodiagnostic studies. On exam, the Tinel’s and Phalen’s tests are provocative tests that are suggestive of CTS. In Tinel’s test, the median nerve is percussed at the wrist, and the subject is asked if this results in pain or paresthesias in a median nerve distribution in the hand. For Phalen’s test, a subject maintains wrist flexion for 60 seconds after which typical CTS symptoms are reproduced in a positive test. One study found the positive predictive value (PPV) of the Tinel’s sign alone to be 55% whereas the PPV of a positive Tinel’s plus a classic diagram of hand pain to be 71% (100). Electrodiagnostic testing offers additional objective data of CTS, but one study has demonstrated that median neuropathy is present in 18% of 125 randomly selected individuals who had no symptoms of CTS (101).

Management of CTS begins with appropriate workplace changes including reducing the repetitiveness of tasks and limiting “wrist-intensive” duties. Nonsurgical management often consists of nighttime splinting, NSAIDs, tendon/nerve gliding exercises, and local corticosteroid injections. One randomized trial found nocturnal wrist splinting to be beneficial in symptom reduction at 4 weeks compared to no treatment (102). Goodman and Gilliatt reported that 34 of 51 patients treated with splinting alone were symptom free after 6 to 30 months of follow-up (103). Giannini et al. reported relief of symptoms in 90% of CTS patients within 45 days and in 93% after 6 months, in those who received steroid injection (104). A 2007 systematic review concluded that steroid injections have greater symptom improvement than placebo at 1 month (105). Finally, for patient refractory to these nonsurgical measures, surgical decompression can be employed and offers many patients sustained symptom relief.

**Pronator Syndrome**

The median nerve exits the upper arm and enters the forearm by passing under the bicapital aponeurosis between the two heads of the pronator teres. Entrapment of this nerve can occur at four points: under the ligament of Struthers in the medial distal arm, under the lacertus fibrosus (a thickened aponeurosis from the biceps tendon to the flexor forearm mass), under the pronator teres, or under a thickened or fibrous portion of the flexor digitorum superficialis. These four entrapment areas can be differentiated based on history and physical examination. The pronator syndrome may occur in those who perform jobs requiring repeated pronation, supination, lifting, or handling of heavy objects (106).

Clinically, the patient with pronator syndrome often presents with forearm pain, numbness, and weakness. In distinction from CTS, weakness in proximal median innervated muscles such as the flexor carpi radialis or pronator teres may be present. In addition, the distribution of sensory abnormality will include the thenar eminence in pronator syndrome. This area is spared in CTS (107). Entrapment at the ligament of Struthers and at the lacertus fibrosis can be accentuated by resisted full elbow flexion or resisted supination. Entrapment...
at the pronator teres and flexor digitorum superficialis can be accentuated by resisted forearm pronation and resisted middle finger flexion, respectively (108).

Treatment should begin with duty modification and relative rest. NSAIDs can be used to decrease pain. Simultaneously, one should begin physiotherapy to stretch and strengthen the involved area. If conservative methods fail, surgical resection of fascial bands, tendinous arches, or the ligament of Struthers may be considered. These authors recommend electrophysiologic evaluation prior to surgery to confirm the precise location of the nerve entrapment. In addition, peripheral nerve visualization with MR imaging or musculoskeletal ultrasound (in the hands of an experienced radiologist) can be very useful to exclude space-occupying lesions as an etiology of the compressive neuropathy (109).

**Ulnar Nerve Entrapment**

The ulnar nerve arises from the medial cord of the brachial plexus and travels along the medial aspect of the arm and then behind the medial epicondyle in the ulnar groove. It enters the forearm between the two heads of the flexor carpi ulnaris (FCU) and travels in the cubital tunnel. The cubital tunnel is formed just distal to the ulnar groove as the ulnar nerve passes underneath the FCU. This median nerve is covered by the arcuate ligament of Osborne that bridges across the two heads of the FCU. The ulnar nerve then gives off the dorsal and palmar cutaneous branches before passing through Guyon’s canal, which is defined by the pisiform and hamate bones and the overlying transverse carpal ligament.

Entrapment or compression of the ulnar nerve within the cubital tunnel is the second most common compression neuropathy in the upper limb after CTS (110). Elbow flexion causes the arcuate ligament to become taut, the FCU to tighten, and the ulnar collateral ligament of the elbow to buckle and encroach into the tunnel (111,112). Apfelberg and Larson demonstrated a 55% narrowing of cubital tunnel volume during elbow flexion (113). The combined movements of shoulder abduction, elbow flexion, and wrist extension result in the greatest increase in cubital tunnel pressure. Repetitive flexion and extension of the elbow itself may chronically irritate the nerve or result in hypertrophy of the arcuate ligament, leading to further compression (114). Cubital tunnel syndrome is commonly seen in occupations involving repetitive elbow flexion, prolonged flexed elbow posture, and repeated trauma or pressure to this area (115). Specifically, these repetitive stresses occur in shoveling, hammering, lifting, and digging (115). Cubital tunnel syndrome has also been noted in assembly line workers because of repetitive elbow flexion and extension (116). Problems arising from leaning on the elbows for prolonged periods have been identified in computer operators at workstations, as well as in truck and taxicab drivers. There has also been an association between cubital tunnel syndrome and sewing machine operators, who also frequently rest their elbows on their workstations (117).

The second most common site for ulnar nerve compression is in the region of Guyon’s canal. Chronic focal compression, ganglion cysts within the canal, ulnar artery thrombosis, or acute blunt trauma to the region may lead to a distal ulnar neuropathy (118). Ulnar nerve entrapment at Guyon’s canal was first identified in gold polishers in 1896 and has been subsequently reported in boot makers, secretaries/typists, machine operators, meat packers, pipe cutters, mechanics, and long-distance cyclists (119,120).

Cubital tunnel syndrome initially presents with numbness, tingling, or paresthesias of the fourth and fifth digits and can progress to weakness of grip strength, wasting of ulnar-innervated hand intrinsics, and, finally, a “clawing” of the hand. The symptoms may be worse at night due to flexion of the elbow during sleep (115). There are several provocative examination maneuvers that aid with diagnosis. The elbow flexion test, in which one maintains the elbow in full flexion for one minute, can provoke paresthesias in an ulnar distribution (121). A positive Tinel’s sign, in which tapping of the ulnar groove reproduces paresthesias or pain in the fourth or fifth fingers, is suggestive as well. Novak et al. found that the sensitivity of the flexion test at 60 seconds was 0.75 and the specificity 0.99, while the sensitivity of the Tinel’s test was 0.70 and specificity 0.98 (122). There are additional tests that assess muscle weakness that occurs in more severe cases of entrapment. Froment’s test involves asking the subject to make a fist and grasp a piece of paper firmly between the thumb with an extended IP joint and the index finger. This test assesses for adductor pollicis weakness, which may be present in ulnar neuropathy. Subjects will compensate by firing the median-innervated FPL to grasp the paper between thumb and index finger, but this will result in flexion of the thumb IP joint (Froment’s sign). Wartenberg’s sign occurs when a subject has difficulty adducting 5th digit when it is extended.

Compression of the ulnar nerve at Guyon’s canal may cause variable clinical findings, depending on where the nerve is injured within the canal. Loss of sensory function, motor function, or a combination thereof may be noted. If compression occurs in the proximal aspect of Guyon’s canal, the patient will have paresthesias or numbness along the lateral hand and possibly motor weakness of the hand intrinsics. If the lesion occurs at the lower wrist, there can be sparing of the sensory branch to the dorsal aspect of the hand. Finally, if the lesion is at the distal aspect of Guyon’s canal, there is usually only sensory loss along the distribution of the superficial branch of the ulnar nerve. Focal pressure applied between the pisiform and the hook of the hamate may induce symptoms.

To confirm the diagnosis of ulnar neuropathy, electromyography should be employed. Nerve conduction studies and EMG can precisely localize the area of entrapment and suggest an early diagnosis when symptoms are subtle (123,124). Furthermore, it can clarify the nature of pathology by distinguishing axonal degeneration, segmental demyelination, and abnormal nerve irritability from one another (125).

In the industrial setting, the first step in management is to reduce extrinsic compression of the ulnar nerve via relative rest, activity modification, and/or padding. Additional measures include anti-inflammatory medication and nocturnal
Thoracic Outlet Syndrome

Thoracic outlet syndrome (TOS) comprises several subsets of conditions that all lead to compression of the lower trunk of the brachial plexus or subclavian vessels along the costoclavicular passages. Categories of TOS include scalenus anticus, costoclavicular, and the hyperabduction syndromes. In the neck, the subclavian artery travels along with the lower trunk of the brachial plexus between the anterior and middle scalene muscles. The subclavian vein typically passes anteriorly to the anterior scalene muscle, but in some cases may also travel along with the subclavian artery between the anterior and middle scalene muscles. The neurovascular bundle then travels between the clavicle and the first rib and underneath the pectoralis minor muscle as it reaches the arm. Compression is theorized to occur at any of these anatomic sites.

Physiologic narrowing of the costoclavicular space can occur with sustained activities requiring hyperabduction of the arms, drooping of the shoulders, and contractions of the scalene muscles. This narrowing can occur in the context of weight lifting, frequent overhead work, or overhead lifting. Additionally, patients with an accessory cervical rib are particularly vulnerable to brachial plexopathy of the lower trunk. Tubbs et al. demonstrated histologic changes in nerves distal to the site of compression by cervical ribs (129).

Clinically, TOS can have vascular or neurogenic symptoms. The classic symptoms consist of paresthesias in the fourth and fifth digits with possible medial forearm involvement. Vascular changes including Raynaud’s phenomenon may be present. Hand weakness is seen in true “neurogenic” thoracic outlet compression. Often, the patient may have only a sense of upper extremity fatigue with sustained overhead postures. On physical examination, sensory abnormalities may be apparent in the distribution of the C8/T1 dermatome or lower trunk of the brachial plexus. Provocative maneuvers, such as Adson’s, Roo’s, costoclavicular, or Wright’s hyperabduction test, have low sensitivity and specificity. One study suggested a mean sensitivity of 72% and specificity of 53% of these tests (130). This same study also found the Adson’s maneuver to have a PPV of 85% and the hyperabduction test to have a PPV of 92%. However, another more recent study found these maneuvers to have a high false-positive rate in normal subjects and an even higher false-positive rate in CTS patients (131). More sensitive diagnostic testing includes Doppler studies, MRA, angiography, or nerve conduction studies. One small series of 13 TOS patients (as confirmed by angiography) suggested that Doppler studies are an effective noninvasive diagnostic tool and correctly made the diagnosis in 10 of these 13 patients (132).

Initial management consists of a coordinated program to ameliorate occupationally induced factors, along with physical therapy to stretch implicated muscles (scalenes, pectoralis minor, trapezius, and levator scapulae). In addition, strengthening of scapular elevators and stabilizers, such as the upper/middle trapezius and rhomboids, is used to improve posture and increase the area within the costoclavicular space (133). Postural biomechanics must be emphasized, as this likely underlies most cases of mild to moderate TOS. A study from Kuopio University Hospital, which stressed normalization of the motion of the cervical spine and upper thoracic aperture via conservative therapies, demonstrated 88% patient satisfaction. In this study, satisfaction was predicted by normalized grip strength, negative Tinel’s sign, and return to work (134).

Surgery for resection of anomalous fibrous bands, thickened musculature, accessory cervical rib, or first rib is reserved for refractory cases thought to be caused by those anatomic factors. Studies suggest that transaxillary first rib resection may provide better results than supraclavicular neuroplasty without rib resection (135). The outcome of TOS surgery among injured workers may be worse than the general population. The strongest predictors for remaining disabled were significant work disability prior to surgery, longer time between injury and TOS diagnosis, and older age at injury (136). Botulinum chemodenervation of the scalene muscles may be helpful in alleviating symptoms in TOS patients awaiting definitive surgical decompression (137).

Low Back Pain

LBP is one of the most common CTDs and is encountered in myriad occupations ranging from sedentary jobs, such as truck driving or secretarial work, to those requiring frequent lifting, such as nursing or construction. A prospective, longitudinal study in South Manchester revealed that lifting objects weighing more than 25 lb, pushing/pulling heavy objects, and standing or walking for prolonged periods were associated with a higher incidence of LBP, especially among female workers (138). Vibration to the spine has been tied to LBP in forklift drivers and crane operators (139,140). Psychosocial factors play a significant role in many cases of LBP. One multivariate analysis reported an OR of 2.52 for LBP incidence in those with psychological distress (141). In an initially asymptomatic VA cohort, depression was the single greatest baseline predictor of incident LBP (hazard ratio = 2.3) (142). A thorough discussion of the pathophysiology, diagnosis, and management of LBP is beyond the scope of this chapter and can be found in Chapters 33 and 68.
Ergonomic Interventions

Ergonomic workplace modifications are increasingly utilized to prevent CTD. The Occupational Safety and Health Administration, a division of the U.S. Department of Labor, has published workplace guidelines for different industries including health care, retail groceries stores, shipyards, etc. The Center for Disease Control also has a concise summary of ergonomic considerations in a variety of workplace settings available at http://www.cdc.gov/od/ohs/ergonomics/ergohome. While an in-depth discussion of ergonomic evaluation and design is beyond the scope of this chapter, some general considerations are introduced below.

Positions/Postures

- The worker should avoid extreme flexion of the elbow, extreme flexion or extension of the wrist, and sustained supination or pronation of the forearm.
- During repetitive work (e.g., typing), the upper extremity should be relaxed at the sides, with the elbows flexed no more than 70 to 90 degrees.
- Avoid prolonged maintenance of the elbow in the overhead position during repetitive or sustained activities.
- Avoid activities with sustained grip or pinch.
- Avoid repetitive radial or ulnar deviations of the wrist.
- Avoid sitting in the same position for extended periods of time. Take breaks to stretch and walk around.

Seated Computer Workstation

- The height of the workstation is a very important component in the overall design, as small changes in height alone may influence the positioning of the neck, shoulders, elbows, and wrists. If a desk area is too high, the worker is forced either to abduct or elevate the shoulders, forcing increased sustained muscle contraction along with compensatory ulnar deviation at the wrists. If the desk area is too low, the worker may be forced to slouch in the chair, with rounding of the upper back, flattening of the lumbar lordosis, and an increase in forward head posture.
- Telephone headsets could be used if extensive or frequent telephone calls are expected.
- The top of the computer monitor should be in line with the worker’s eyes, allowing for a natural downward gaze angle of 15 degrees toward the middle of the screen. This position also allows the worker to maintain the normal lordotic curve of the cervical spine, avoiding the forward head posture.
- The optimum distance from the monitor is 12 to 18 in. If the monitor is too close, the worker will experience eye strain; whereas if it is too far away, the worker is forced to slouch forward, increasing stress throughout the entire spine. Also avoid variation in the distance of eyes to the monitor to avoid frequent changes in focus.
- While seated, the worker’s feet should rest on the floor with the knee flexion maintained between 90 and 105 degrees. If the feet are not resting on the ground, a footrest should be used.
- The ideal seat depth should allow 1 to 4 in. between the popliteal fossa and the front edge of the seat. A shorter depth seat will require the buttocks and upper thighs to sustain excessive pressures, whereas a longer seat depth may impede circulation behind the knees.
- The back rest should support the lumbar curve while maintaining the thoracic spine in contact with the chair back. It may be necessary to add a rolled towel or lumbar support if the back rest does not sufficiently support the natural lordosis of the lumbar spine.
- Adjust keyboard and mouse height so that elbows can be flexed at a 90-degree angle and forearms are parallel to the ground.

Many workplace environments have already implemented ergonomic design changes. Further research and studies are needed in this area to determine optimal positioning and ways to avoid repetitive strain.

CONCLUSION

CTDs are common in our society and account for a large fraction of all occupational injuries. Clinicians managing patients with CTDs should not only have knowledge of the pathophysiology and biomechanics at the core of the problem but should be familiar with common workplace and recreational scenarios that would lead to these disorders. This allows the examiner to perform a focused history and physical exam. Next, an appropriate diagnostic workup consisting of imaging, electrophysiologic studies, and labs can be selectively ordered. The clinician should then take the lead in coordinating the different aspects of treatment. The mainstay of treatment should focus on correction of underlying causative factors. This involves activity restriction, if possible, and modifying equipment and posturing techniques in accordance with ergonomic principles. Further conservative treatment involves a regimen of targeted strengthening or stretching exercises, use of orthotic devices, modalities, and anti-inflammatory medications. Appropriate management of pain should be a priority at all stages of treatment. While the majority of CTDs respond to conservative therapy, surgical consultation should be sought in refractory cases.

Although our knowledge of the prevention and management of CTDs has been enhanced by numerous research studies, further controlled studies are needed in this area. Future areas of clinical research include evidence-based ergonomic interventions for primary prevention, specific physiotherapy regimens, the role of newer therapeutic modalities (i.e., acupuncture), and long-term outcomes research for nonsurgical and surgical interventions. Additionally, basic science research of these disorders will lead to a better understanding of their pathophysiology, which may lead to new therapeutic targets. Overall, the management of CTDs continues to be an important and challenging field.


Hand Disorders

Our hands are our primary tools for interacting with our environment. One may argue that the most significant function of the shoulder, elbow, and wrist is to position the hand in space. Once the hand is positioned, it is able to proceed with its functional task. Therefore, any abnormalities of hand function are important to evaluate and when possible treat, because of the critical role that our hands play in our daily routines.

The anatomy of the human hand is extremely complex, with many intricate mechanisms that allow us to grip a tennis racket, hold a pencil, or play the piano. Perhaps because of its intricacy, the care of the hand, both surgical and nonsurgical, has been relegated to specialists in hand surgery. Physiatrists have stayed on the sidelines. However, with a thorough understanding of the hand anatomy and pathophysiology, a well-trained physiatrist is in a strong position to manage all medical aspects of hand care.

The goal of this chapter is to provide a brief overview of common hand problems, their pathophysiology, and their treatment. Most of the treatment discussed is nonsurgical in nature. However, a brief overview of surgical options and their indications is also provided. A more detailed account of these clinical entities and their treatment is available in hand surgical texts and in this chapter’s reference section.

Carpal tunnel syndrome (CTS) is the most common compressive neuropathy in the upper extremity (1) (Tables 37-1 and 37-2). This condition typically causes paresthesias in the median nerve distribution. Fifteen percent of the general population has symptoms consistent with CTS. The prevalence of electromyogram-confirmed CTS is 3% in women and 2% in men. Prevalence is greatest in women older than 55 years (2). It is more frequently found in people who are obese or smoke and in those with diabetes mellitus (3,4). Cervical root compression and thoracic outlet abnormalities can also be associated with CTS (5). No hand preference, dominant or nondominant, has been identified, and 60% of patients have bilateral symptoms (2,6). Many factors contribute to the development of CTS, including vibration, local pressure over the carpal tunnel, awkward wrist positions, and forceful hand positions (7–9). However, there remains considerable debate about whether CTS is a result of repetitive stress without other factors being present (10,11). Food processing, construction, and manufacturing are occupations that have a higher incidence of CTS (12). Debate remains as to the association of CTS and computer keyboard work (13–15). It has even been hypothesized that keyboard work may be associated with less CTS (16).

The diagnosis of CTS is usually straightforward and is based on history and clinical evaluation. Electromyogram (EMG) is often used for diagnostic confirmation of CTS. It can measure the extent of damage and demyelination of the median nerve (15). In mild cases, there may be an absence of electromyographic and nerve conduction changes. As symptoms progress, sensory distal latency is usually the first abnormal EMG finding. Therefore, the diagnosis of CTS is first established on history and clinical findings and then may be confirmed by EMG evaluation. Recently, there have been multiple reports on using ultrasound evaluation of the median nerve to diagnose CTS. These studies have shown that there is a change in the cross-sectional area of the median nerve when CTS is present (17–19).

The carpal tunnel is located just distal to the palmar wrist crease. It is surrounded on three sides by the carpal bones, creating an arch. The radial wall is bordered by the scaphoid and trapezium, the dorsal is bordered by the lunate and capitate, and the ulnar wall is composed of the hamate. The arch is covered by a thick fibrocartilaginous band called the flexor retinaculum (or transverse carpal ligament.) Nine tendons—two tendons that flex each finger, flexor digitorum superficialis (FDS) and flexor digitorum profundus (FDP), and one to the thumb, flexor pollicis longus (FPL)—course through the carpal tunnel (20). The median nerve traverses with these tendons through the tunnel on its way to provide innervation to the thenar muscles and to provide sensation to the radial three and one half digits (Fig. 37-1).

CTS is thought to be the result of increased pressure within the fibroosseous tunnel. Normal pressure within the carpal tunnel is 7 to 8 mm Hg with the wrist in neutral. Increased pressure of 30 mm Hg can result in symptoms of CTS and 90 mm Hg can be observed with wrist flexion and extension (21,22). This pressure increase causes relative ischemia and impaired nerve conduction of the median nerve (23,24). If abnormally increased pressure continues, segmental demyelination occurs. The median nerve sensory fibers are the first to be affected owing to their extensive myelination and high metabolic demands. Prolonged pressure causes injury to the motor fibers, and weakness ensues. The prevalence of CTS increases with pregnancy, inflammatory arthritis, distal wrist fracture, amyloidosis, hypothyroidism, diabetes, acromegaly, and in individuals who use corticosteroids and estrogens (13,20,25). One third of all cases of carpal tunnel are associated with these conditions and in this chapter’s reference section.
medical conditions (26); diabetes is the most common association (13,26).

The typical symptoms of CTS are numbness, tingling, pain, burning, or a combination of these (27). These symptoms occur in the radial three and one half digits: the thumb, index, middle, and half of the ring finger. CTS often causes nocturnal awakening secondary to the hand paresthesias. These nocturnal symptoms are 51% to 77% sensitive and 27% to 68% specific for CTS (28). Gripping, driving, holding vibrating objects, or prolonged pinching, such as holding a book, may result in increased paresthesias. Many patients describe relief of their symptoms with shaking of the hands (29). With progression, patients may describe an awkward feeling or weakness of the hand and begin dropping objects.

Physical examination (Table 37-3) usually begins with the exclusion of any cervical, shoulder, or elbow pathology, which may produce similar symptoms. C6 radiculopathies are often confused with CTS because the sensory symptoms involve the radial aspect of the hand. Strength testing should include wrist flexion-extension, grip, and thumb opposition. Specific CTS provocative tests include Phalen’s test, in which the wrist is held in full passive wrist flexion. This position increases pressure within the carpal tunnel and may reproduce paresthesias.

<table>
<thead>
<tr>
<th>Physical Exam</th>
<th>Electromyogram</th>
<th>Causes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tinel’s at wrist</td>
<td>Prolonged distal latency of median sensory, possible motor</td>
<td>Space-occupying lesions (trauma, tumor, inflammation)</td>
</tr>
<tr>
<td>Grip test</td>
<td>Needle exam findings in thenar muscles only</td>
<td>Direct trauma</td>
</tr>
<tr>
<td>Carpal tunnel compression, positive Phalen’s</td>
<td>Physiologic compression</td>
<td>Physiologic compression</td>
</tr>
<tr>
<td>Two-point discrimination, &gt;6 mm</td>
<td></td>
<td>Possibly vibration</td>
</tr>
<tr>
<td>Opposition strength</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**TABLE 37.1 Carpal Tunnel Syndrome**

**FIGURE 37-1.** Carpal tunnel syndrome. Median nerve traversing with the flexor tendons through the carpal tunnel providing sensory innervation to the radial three and one half fingers and motor innervation to the thenar muscles.

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>Symptoms and Signs</th>
<th>Prescriptions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trigger finger</td>
<td>Locking of digits in flexion, snapping, popping, tender nodules, crepitus</td>
<td>Rest, activity modification, modalities, cortisone injection, A-1 pulley release</td>
</tr>
<tr>
<td>de Quervain’s</td>
<td>Radial wrist pain, positive Finkelstein’s</td>
<td>Splint, modalities, cortisone injection, first dorsal compartment release</td>
</tr>
<tr>
<td>Dupuytren’s</td>
<td>Superficial palmar nodules, painless flexion contractures</td>
<td>Possibly collagenase injection, surgical release</td>
</tr>
<tr>
<td>Carpal tunnel syndrome</td>
<td>Nighttime hand paresthesias, positive Tinel’s, carpal tunnel compression, Phalen’s</td>
<td>Splint, activity modification, cortisone injection, surgical release</td>
</tr>
<tr>
<td>Cubital tunnel syndrome</td>
<td>Ulnar hand pain and paresthesias, positive Tinel’s at elbow, possible interosseous weakness</td>
<td>Elbow protection, surgical release, nerve transposition</td>
</tr>
<tr>
<td>Basilar thumb degenerative joint disease</td>
<td>Pain at base of thumb, positive grind maneuver</td>
<td>Splint, activity modification, NSAIDs, cortisone injection, surgical reconstruction</td>
</tr>
</tbody>
</table>
in individuals with CTS. This test has a wide reported range of sensitivity and specificity (40% to 80%) (28,30,31). The time to the development of paresthesias should be noted because it can be used to monitor change with treatment. Tinel’s test involves tapping the median nerve just proximal to the transverse carpal ligament (32). Reproduction of the paresthesias into the hand by the Tinel’s test is 20% to 60% sensitive and 67% to 87% specific for CTS (30,31). Carpal tunnel compression involves pressure placed with the examiner’s thumb or index and long fingers over the carpal tunnel. This pressure is maintained for 30 seconds to 1 minute and if positive will reproduce paresthesias. Durkan (33) believes that this test is more sensitive and specific for CTS than Tinel’s or Phalen’s test.

Treatment of CTS begins with modification of repetitive or awkward activities that precipitate paresthesias. Splinting the wrist in a neutral position at night has been demonstrated to reduce symptoms in 80% of patients (34,35). Nonsteroidal anti-inflammatory drugs (NSAIDs), diuretics, vitamin B6, and oral steroids have been tested, but no specific recommendations have been given for their prolonged usage (27,36). Therapeutic interventions such as ultrasound, iontophoresis, gentle stretching and strengthening exercises, ice, and carpal tunnel protection principles may be employed. Their application is clinically accepted but not scientifically determined to be efficacious. Protection principles stress avoidance of positions or activities that increase pressure within the carpal tunnel. Nerve and tendon gliding exercises have been described and are thought to be useful (37–39). Acupuncture and yoga have also been demonstrated to decrease symptoms (40).

Corticosteroid injections into the carpal tunnel are recommended if splinting and other conservative measures fail to reduce the symptoms. They have been shown to decrease symptoms in 75% of patients and improve nerve conduction (41–44). These injections are performed in a sterile fashion with needle placement ulnar to the palmaris longus. The needle is directed dorsally, distally, and radially at a 45-degree angle. In patients with severe CTS, 80% have return of symptoms in 1 year despite appropriate conservative care (21). If the patient has signs or symptoms of constant numbness, loss of sensation, or thenar muscle atrophy lasting longer than 1 year, serious consideration of surgery is recommended (20). There are many approaches to a CTS release. Traditionally, an open incision has been performed and the transverse carpal ligament resected. Endoscopic CTS release can be performed but may pose a slightly greater risk for injury to the median nerve (45). Relief of symptoms is similar in the two approaches, but patients typically return to work quicker with the endoscopic technique (46,47). Postoperative rehabilitation versus home exercises seem to have the same outcomes, except that it has been shown that rehabilitation hastens the time to return to work (48).

### Ulnar Neuropathy

The ulnar nerve may be compressed at a number of points along its course. There are four areas of common compression around the elbow and one area at the wrist at Guyon’s canal. The areas of compression about the elbow include the medial intermuscular septum, the arcade of Struthers, the cubital tunnel, and the deep flexor aponeurosis (49) (Fig. 37-2). Compression of the ulnar nerve at Guyon’s canal may be secondary to thrombus at the ulnar artery (hypothenar hammer syndrome) or to ganglion cyst from the pisotriquetral joint or the triquetral hamate joint.

Ulnar neuropathy at the elbow produces weakness of the wrist ulnar deviators, the lumbricals of the ring and small fingers,

**TABLE 37.3** Evaluation: History and Physical Findings

<table>
<thead>
<tr>
<th>Duration/severity</th>
<th>Presence of numbness: location along thumb, index, middle and half of ring finger, constancy. Use of Semmes-Weinstein monofilaments to determine sensory loss</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nocturnal paresthesias</td>
<td>Nocturnal paresthesias</td>
</tr>
<tr>
<td>Weakness of grip and/or pinch</td>
<td>Weakness of grip and/or pinch</td>
</tr>
<tr>
<td>Muscle wasting of the thenar muscles in absence of subluxation of the first MCP joint</td>
<td>Muscle wasting of the thenar muscles in absence of subluxation of the first MCP joint</td>
</tr>
<tr>
<td>Provocative testing; Tinel’s sign: paresthesias upon tapping the median nerve; Phalen’s sign: the reproduction of symptoms in the median nerve distribution (e.g., thumb, index, middle and half of ring finger), after 1 min of forcful flexion of the wrist while the thumb is flexed into the palm</td>
<td>Provocative testing; Tinel’s sign: paresthesias upon tapping the median nerve; Phalen’s sign: the reproduction of symptoms in the median nerve distribution (e.g., thumb, index, middle and half of ring finger), after 1 min of forcful flexion of the wrist while the thumb is flexed into the palm</td>
</tr>
</tbody>
</table>

**FIGURE 37-2.** Ulnar neuropathy: points of compression in the elbow region.
and all of the interosseus muscles. Sensory loss is at the ulnar side of the hand extending from the wrist crease to the tips of the small finger and ulnar half of the ring finger both dorsally and in a palmar direction. Entrapment at Guyon’s canal spares the wrist muscles as well as the dorsal hand cutaneous sensation. A small sensory branch of the ulnar nerve, the superficial cutaneous branch, is also spared, thus preserving a small area of sensory innervation to the base of the hypothenar eminence. Patients may complain of pain at the medial elbow and anywhere along the course of the ulnar nerve. Numbness and paresthesias may also be present. Weakness is usually a later sign. Patients may develop atrophy of the interosseus and hypothenar muscles and show posturing of their hand. Because the extrinsic finger extensors are spared, a mismatch of strength is seen between the extrinsic and intrinsic finger muscles. This disparity produces the classic “claw” deformity of the fingers. Because the lumbricales of the index and long fingers are median nerve innervated, these two digits may show normal posturing. This “clawing” of the ring and small fingers is called the “benediction sign” (Fig. 37-3).

The conservative treatment of ulnar neuropathy at the elbow is based largely on protection principles. The ulnar nerve in the cubital tunnel lies in a position that is vulnerable to trauma. Therefore, counseling the patient to avoid leaning the elbow on tables, counters, or arm rests may prevent repetitive trauma to the nerve. The patient may also wear various forms of elbow pads or protective elbow sleeves. These not only cushion the cubital tunnel area but also serve as a constant reminder to the patient to avoid pressure on the elbow.

Even though the cubital tunnel resides at the center of rotation for the elbow joint, flexing the elbow may still stretch the ulnar nerve. This is possible because the ulnar nerve may show as much as 21.9 mm of excursion at the elbow and 23.2 mm at the wrist with combined motion of the shoulder, elbow, wrist, and fingers (50). This excursion may be inhibited by fibrosis or entrapment of the nerve at one or more of the aforementioned compression sites. Thus, the normal gliding of the ulnar nerve may be limited and cause tension on the nerve with movement either proximal or distal to the elbow as well as with elbow flexion and extension. The patient should be counseled to avoid prolonged elbow flexion or posturing of the upper extremity in any position that reproduces their symptoms.

Surgical decompression or transposition of the ulnar nerve should be considered when the site of compression is clear either by EMG findings or by imaging and when the symptoms are progressive or debilitating. Many types of idiopathic ulnar neuropathies may be unresponsive to surgical treatment or may improve spontaneously. Nonetheless, when invasive treatment is considered, there are three major categories of surgical treatments for ulnar neuropathy at the elbow (51). The first is simple decompression of the nerve usually combined with medial epicondylectomy. The second and third are subcutaneous and submuscular transpositions of the ulnar nerve. In general, greater amounts of dissection around the elbow lead to greater chances of postoperative elbow stiffness. Moreover, the amount of manipulation of the ulnar nerve is directly related to the chances of damage or continued neuropathy postoperatively. Therefore, the simpler procedures of decompression tend to be the safest. However, placing the ulnar nerve anterior to the medial epicondyle protects it from direct trauma and from stretching. In addition, the deeper the nerve is buried, the less chance for further trauma to the nerve.

**THUMB CARPOMETACARPAL OSTEARTHROSIS**

Basilar thumb osteoarthritis is the most common symptomatic arthropathy of the hand. Thirty-three percent of postmenopausal women older than 50 years have radiographic evidence of osteoarthritis of the thumb carpometacarpal (CMC) joint (52). The etiology of CMC arthritis is likely to be multifactorial, including genetic, environmental, and physiological contributions (53). The thumb CMC joint is a double-saddle configuration. This configuration allows for movement in multiple planes such as flexion-extension, adduction-abduction, and pronation-supination. Sixteen ligaments have been described stabilizing the trapezium and the CMC joint (54). The most important of these is thought to be the so-called beak ligament or deep anterior oblique ligament. Ligamentous laxity plus axial loading of the joint are thought to be major contributors to the development of osteoarthritis at the base of the thumb. However, there may also be a link between hand osteoarthritis and obesity (55).

Classification of the thumb OA is radiographic and is graded according to the change in the trapeziometacarpal space, degree of synovitis, and subluxation. Stage 1 is associated with a normal joint space, whereas stages 2 to 4 have a decreased joint space. Stage 3 has obligatory osteophyte changes and sclerosis of the joint. Stage 4 has all of the observations plus involvement of the scaphotrapezial joint (56). Symptoms vary with each stage.

In the treatment of basilar thumb degenerative joint disease (DJD), one tries to manage the symptoms. No cure is known. We are not even sure that we can slow the progression of the disease. Ideally, the patient acquires new prehensile patterns and adaptive equipment in order to minimize symptoms and maximize function. No single intervention is necessarily superior to others. There is no specific time frame or window of opportunity to follow. The CMC joint biomechanics is such that any pinch force generated at the thumb and
The index finger is greatly magnified at the CMC joint interface. In fact, 1 kg of pinch force translates to 12 kg of intra-articular pressure; thus, one strives to minimize fingertip-to-fingertip pinch activities (57). When pinching cannot be avoided, one tries to increase the size of the objects being pinched. In other words, the greater the distance between the thumb and the fingertip, the less pressure on the CMC joint; therefore, enlarging the grip of tools or objects that are being gripped decreases the pressure at the base of the thumb.

Twisting activities also stress the CMC joint by causing a torque or twisting force on the joint. Therefore, various gadgets such as key holders and electric can openers may be helpful. In addition, one may use pens, kitchen utensils, or gardening implements with built-up grips. Many companies manufacture such items.

Splints are used to help stabilize the thumb, reducing pain and enabling more symptom-free function. There are two commonly used splints for this problem: a short opponens splint which is hand-based and crosses the first metacarpophalangeal (MCP) joint; and a long opponens splint which is forearm-based, supports the wrist, and crosses the interphalangeal joint. Patients typically best tolerate the long opponens, forearm-based splint at night and prefer the less obstructive hand-based splint for daytime activities. These splints usually put the thumb in palmar abduction with the MCP in 30 degrees of flexion. Several authors have documented benefits from splinting in stabilizing the joint and providing pain relief (53,58). Others have found studies of the efficacy of splints to be methodologically weak and unconvincing (58a). Splints usually are not successful in symptom reduction when there is fixed deformity of the joint. A review of the evidence for splinting the CMC joint in arthritis is available (59).

Many patients use pain-relieving modalities such as contrast baths, hot-water soaks, or paraffin baths. Anti-inflammatory medications are often used for pain control. In addition, one may consider cortisone injections into the CMC joint (60). If the patient has involvement of the scapho-trapezio-trapezoidal (STT) joint, this may be injected at the same time with a single needle stick.

As a last resort, some patients may consider surgical options for basilar thumb DJD. The most common procedure used to treat this problem is ligamentous reconstruction of the joint (61). The trapezium is wholly or partially excised. The base of the first metacarpal is then reattached to the carpus using a tendon slip from the flexor carpi radialis (FCR), the abductor pollicis longus (APL), or the extensor pollicis brevis (EPB). Various techniques are used to weave the tendon between the bases of the first and second metacarpals with the distal scaphoid. Some variations also employ the remaining tendon slip as a cushion to fill in the gap left by the trapezium. This is also called the ligamentous reconstruction tendon interposition (LRTI) or the “anchovy” procedure. Although the tendon reconstruction procedures help restore mobility of the thumb and reduce pain, they are not suited for heavy activities or manual labor. The patients requiring a large amount of grip strength and durability would be better served by arthrodesis of the joint (62,63). This allows for quicker recovery and preservation of grip strength. The tendon reconstruction method requires 6 months of recovery and often only results in 60% to 70% of normal grip and pinch strength (64). Various interposition arthroplasty procedures are less common and are not as well established as tendon or fusion procedures.

Other sequelae of osteoarthritis, such as instability or pain of the first MCP joint, may contribute to pain and prehensile abnormalities.

**de QUERVAIN’S TENOSYNOVITIS**

de Quervain’s tenosynovitis is characterized by pain on the radial aspect of the wrist at the first dorsal compartment. This insidious condition most commonly affects women between the ages of 35 and 55 years (65,66). Women are affected 10 times more frequently than men. There does not appear to be a predilection for dominant or nondominant hand nor for race. Repetitive, prolonged unaccustomed posturing of the thumb or nonneutral wrist movements usually provoke symptoms (67). Waitresses, nurses, garment workers, maids, assembly line workers, and machine operators are at greater risk for development of this condition (68–70).

The extensor tendons to the fingers and wrist travel through six dorsal compartments of the wrist. The first (most radial) dorsal compartment contains the EPB and adductor pollicis longus (Fig. 37-4). These tendons course through the wrist and insert into the base of the thumb. The second dorsal compartment contains the APL, EPB, and FCR. The third dorsal compartment contains the extensor digiti minimi (EDM) and extensor carpi ulnaris (ECU). The fourth dorsal compartment contains the extensor indicis proprius (EIP) and extensor carpi radialis longus (ECRL). The fifth dorsal compartment contains the extensor pollicis longus (EPL) and extensor carpi radialis brevis (ECRB). The sixth dorsal compartment contains the extensor digitorum communis (EDC) and extensor digiti minimi accessorius (EDMA). The EPB and APL are the primary tendons involved in de Quervain’s tenosynovitis. The EPB is the most commonly involved tendon, followed by the APL. The tendons are tethered by the retinaculum, which is a set of ligaments that run over the wrist and form a tunnel-like structure for the tendons to pass through. The retinaculum is located on the radial side of the wrist, and the tendons are tethered to it by the extensor retinaculum. The tendons pass through the retinaculum and then insert into the base of the thumb. The EPB tethers the thumb to the carpus, and the APL tethers the thumb to the metacarpal bones of the hand. The tendons are compressed against the retinaculum, and this results in increased tension and increased friction, which leads to inflammation and pain. The pain is usually located on the radial aspect of the wrist, and it can radiate down the thumb or up the forearm. The pain is usually worse with activities that involve gripping or holding objects, such as lifting, holding a coffee cup, or turning a doorknob. The pain is usually exacerbated by gripping or holding objects, and it can be relieved by resting the thumb or wrist.

**FIGURE 37-4.** de Quervain’s tenosynovitis. The first dorsal compartment contains the EPB and adductor pollicis longus tendons.
an osteofibrous canal to their insertion on the metacarpal and proximal phalanx of the thumb. A significant angulation is present as these tendons traverse over the radial styloid, placing the tendons at risk for repetitive injury (71). The function of these muscles is to position the thumb in extension and abduction in preparation for gripping and pinching.

de Quervain’s tenosynovitis is initially the result of inflammation within the first dorsal compartment. Frequent reoccurrence of this condition results in a reactionary degenerative thickening of the extensor retinaculum and synovial tendon sheath (72,73). In these chronic states, inflammation is absent (74,75). The thickening results in a mechanical stenosis within the first dorsal compartment, causing impingement of the two tendons (76).

The primary symptoms are pain and swelling over the radial styloid process with progression up into the radial aspect of the forearm or distally into the thumb (73). Pain increases with grasping, adduction of the thumb, or ulnar deviation of the wrist. The symptom complex is usually gradual in onset, but traumatic etiologies have been described (72,73).

On physical examination, patients usually have tenderness with palpation over the fibroosseous first dorsal compartment. Pain is commonly elicited with resisted thumb extension and abduction. A positive Finkelstein’s test is pathognomonic for de Quervain’s tenosynovitis (73,77). This test is performed by flexing the thumb into the palm and making a fist around the thumb. The wrist is then passively deviated ulnarly. Increased pain in the region of the radial styloid with this maneuver is considered positive.

de Quervain’s tenosynovitis is a clinical diagnosis. X-rays have not been found to be beneficial. Other conditions with a similar presentation include peripheral neuritis, collagen-vascular diseases, sprains of the CMC joint, arthritis of the CMC joint, fracture of the distal radius, ganglions of the wrist, acute calcific tendinosis, and aberrant CTS (78,79).

Activity modification is often the most important consideration in conservative treatment. Elimination of highly repetitive activities that include pinching or gripping is beneficial (66). Immobilization of the thumb in a forearm-based thumb spica splint offers protection and rest. Heat modalities, stretching of the first dorsal compartment muscles, and ice may offer relief of symptoms. To date, there has not been an outcome study on the use of modalities and exercise for this condition.

Injection of local steroids has been shown to be of benefit (80–82). Anderson et al. (80) reported that 81% of individuals undergoing injections for this condition described symptom relief at 6 weeks. At 4-year follow-up, 58% remained asymptomatic, and 33% had complete reoccurrence. If conservative treatment is not effective, surgical release of the first dorsal compartment can be performed (83,84).

TRIGGER FINGER “STENOSING TENOSYNOVITIS”

Trigger finger or “stenosing tenosynovitis” is a disorder characterized by snapping or locking of the finger (85). This locking sensation is usually felt at the level of the MCP joint (86). Progression can result in the inability to extend the MCP joint fully. The digit can become mechanically locked in flexion, requiring passive assistance to obtain extension. A similar problem can affect the thumb.

The flexor tendons of the digits (superficialis and profundus) are held in place by five annular and three cruciate ligaments (87) (Fig. 37-5). These pulleys maintain the position of the tendon relative to the MCP, posterior interphalangeal (PIP), and distal interphalangeal (DIP) joints, as well as prevent bow stringing of the tendon when the finger is flexed.
Trigger finger or stenosing tenosynovitis is a very common condition seen by hand specialists. It is thought to be secondary to a thickening of the A-1 pulley or flexor tendon owing to shear or compression forces (89,90). Acute conditions demonstrate inflammatory changes in the region of the A-1 pulley. In chronic conditions, no inflammatory changes are noted (91). For this reason, the nomenclature of “stenosing tenosynovitis” has lost favor. Chronic conditions result in degenerative changes consistent with fibrocartilaginous proliferation of the A-1 pulley or tendon. The pathologic thickening results in a disparity of the tendon pulley configuration (92). This size differentiation causes a mechanical locking of the tendon proximal to the A-1 pulley or tendon. The symptoms usually consist of a snapping or locking sensation with full flexion of the digit. This sensation is usually painful, but nonpainful conditions have been described. The onset is usually gradual, over several months, but in certain situations can be due to trauma or carpal tunnel release (99,100). The symptoms of locking or clicking phenomena are usually worse in the morning and after repetitive gripping or pinching-type activities (101).

Examination of the finger is usually unremarkable unless reproduction of the locking phenomena can be observed (101,102). Most often, a tendon nodule or crepitus can be felt over the palmar aspect of the MCP joint in the region of the A-1 pulley (86). Grip strength can be diminished secondary to pain. Ligament and neurovascular integrity is normal. No diagnostic tests are confirmatory for this condition. X-rays have not been found to show any abnormality correlated with trigger finger (103). Serologic testing should be done to check for the presence of underlying conditions such as diabetes mellitus, hypertension, and inflammatory arthritis, which are risk factors for trigger finger (101).

Trigger finger is often nothing more than a nuisance. Depending on one’s hand usage and occupation, however, a trigger finger can lead to disabling pain. Physical modalities such as ultrasound, iontophoresis, and ice may relieve symptoms (104). Evans et al. (105) reported 73% success in using a flexion-blocking splint at the MCP for 3 weeks. Their protocol also included limiting activities requiring grasp, active flexion or repetitive stress, and hooked-fish exercises. Colbourn et al. (106) confirmed these findings but required 6 weeks of continuous splint usage.

Corticosteroid injections have been reported to be highly efficacious in the treatment of trigger finger (107,108). Newport et al. (109) reported that one to three injections of local anesthetic and cortisone were associated with resolution or improvement in 77% of 338 fingers. Marks and Gunther (110) reported that 84% of trigger fingers and 92% of trigger thumbs responded to a single injection. This increased to 91% and 97%, respectively, with a second injection. Serious complications such as infection and tendon laceration have not been reported (111).

Surgical intervention has been advocated if injection therapy does not offer benefit. There has been a plethora of surgical information regarding A-1 pulley releases for the treatment of trigger finger. Thorpe (112) reported that of 53 operations, 60.4% were completely successful, and 11.3% had incomplete resolution with persistence of clicking and pain within the first year after surgery.

**FINGER SPRAINS AND STRAINS**

The most common injuries of the hand are finger sprains and strain. Typically, these are described as “jammed fingers” because an axial load is placed on the digit when trying to catch a ball, striking the digit against a surface, or falling onto an outstretched hand.

**DIP Joint**

To understand the ligamentous and tendonous injuries of the digits, one must first understand the anatomy of the finger joints. The DIP joint is a classic hinged joint with very little radial-ulnar deviation or pronation-supination (Fig. 37-6A, B). This is partly due to the bony configuration, which is bicondylar in shape. Additionally, the radial and ulnar collateral ligaments originate at the center of rotation for DIP flexion and extension and, therefore, remain fairly tight throughout the range of motion of the joint. The volar surface is covered by a thick volar plate, which has a small amount of laxity allowing 40 to 50 degrees of hyperextension. Two tendons cross the DIP joint and terminate at the base of the distal phalanx. On the volar surface lies the termination of the A-1 pulley. On the dorsal surface lies the terminal slip of the extensor tendon. These two tendinous structures are the most commonly injured of the DIP joint.

Rupture or fracture avulsion of the terminal extensor tendon causes the distal phalanx to droop in a classic “mallet finger” deformity (113). In this situation, there is unopposed flexion by the FDP and a complete lack of an extensor force. Therefore, the joint is held in flexion. The DIP may still be extended passively and may flex actively against resistance. Typically, this injury occurs with an axial load on the digit directed slightly in a flexion moment. Aside from the characteristic posturing of the joint, x-rays may show a bony avulsion fragment off the proximal dorsal rim of the distal phalanx. At times, the still intact extensor tendon may draw this fragment off the dorsal surface and into the joint itself. Assuming the joint remains stable (indicated by lack of subluxation on the lateral view x-ray), the treatment is always that of splinting the
DIP in slight extension. Although surgical repair is possible, most would argue that this leads inevitably to a stiff DIP joint. Instead, religious splinting of the joint for 6 to 8 weeks most often restores active extension, even with proximal migration of the avulsion fragment. Delayed treatment may require longer splinting periods, and it is possible that conservative treatment may fail. At such times, surgical treatment may be considered, although this rarely results in a mobile joint.

On the flexor surface of the DIP joint, the most common injury is that of rupture of the FDP (114). This is typically called a “jersey finger” because a common mechanism of injury is that of a football player catching the tip of his finger inside the neck of an opponent’s jersey and forcefully flexing the DIP joint against a stronger extension force. X-rays are usually unremarkable and rarely show an avulsion fragment off the proximal volar rim of the distal phalanx. Unlike the common extensor tendon, there is little to stop proximal migration of such an avulsion fragment and of the free end of the profundus tendon from migrating proximally even into the palm. Treatment is surgical, and little is regained by conservative care.

Proximal Interphalangeal Joint

The ligamentous structure of the PIP joint is almost an exact replica of the DIP joint. The lone exception is that the volar plate has less redundancy than the DIP joint with stout check rein ligaments that extend proximally and prevent hyperextension (Fig. 37-7). Superficial to the PIP joint lies the extensor tendons dorsally, the FDP and the FDS volarly, a number of retinacular ligaments, and the extensor mechanism that coordinates PIP and DIP movements. Because of the multiple tendonous and ligamentous structures that cross the PIP joint and the tight stout volar plate, the PIP joint is much more likely to develop stiffness after trauma or surgery than the other digital joints.

The most common injury to the PIP joint is a ligamentous sprain or strain of the volar plate, collateral ligaments, or accessory ligaments. The most common fracture of the digit is a chip fracture off the proximal volar lip of the middle phalanx (115). These injuries usually result from an axial load to the digit (or “jam”) causing dorsal subluxation of the middle phalanx on the proximal phalanx or more simply,
CHAPTER 37 | HAND DISORDERS

FIGURE 37-7. PIP, crl, check reign ligament.

hyperextension and radial or ulnar deviation. In almost all cases, the treatment is that of buddy taping the involved digit with the adjacent digit. Tape is applied proximally and distally to the PIP joint. A small amount of cotton is placed between the digits to prevent skin maceration. Buddy taping allows for protective active-assisted range of motion of the involved joint while preventing hyperextension as well as radial or ulnar deviation.

The exception to this rule is when there is evidence of instability. Instability would be indicated by subluxation of the PIP joint seen on the lateral view of the digit. Some authors would also argue that an intra-articular fracture involving greater than one third to one half of the joint space would also indicate an unstable joint and necessitate surgical intervention (115). On examination, instability may also be evident by an inability to extend the joint actively or by the joint “opening up” or showing an unexpected amount of passive movement on hyperextension plus either radial or ulnar deviation.

Assuming there is no large fracture fragment, these unstable PIP sprains are treated with a dorsal extension–blocking splint (116). Typically, the joint will sublux on extension and reduce in flexion. Ideally, the joint is examined under fluoroscopy or with lateral x-rays in various degrees of flexion. One must establish the point of extension at which the joint subluxes. A dorsal-blocking splint is then fabricated to prevent extension past this point. Therefore, the joint is allowed to move within the arc of motion where it remains in reduction. This prevents the severe stiffness of static splinting while it allows healing. The joint is then examined radiographically every 1 to 2 weeks, and the blocking splint is repositioned to allow greater extension as the joint stays reduced in greater degrees of extension. Often, the joint is also buddy taped to the adjacent digit within the blocking splint to promote range of motion.

There is almost never an indication for static postinjury splinting of the PIP joint except after surgery. The PIP joint has such a tendency toward stiffness that it is almost better to

Kienböck’s disease is avascular necrosis of the lunate. According to early definitions, this is an atraumatic event. However, the etiology may consist of a combination of factors, including vascular and skeletal variations as well as trauma and repetitive loading (117). The typical symptom is that of mtdorsal wrist pain with point tenderness to palpation over the dorsum of the lunate. In the early stage of the disease, x-rays may be normal. Three-phase bone scans at this stage show an area of decreased uptake at the lunate in all phases of the scan, indicating a lack of blood flow. Magnetic resonance imaging (MRI) shows a decreased T1 and T2 signal at the lunate or what appears to be a black lunate (118). As the condition progresses, the x-rays will first show cystic changes followed by progressive collapse and sclerosis. At these later stages, the bone scan shows increased uptake at the lunate, indicating further bony healing of the fractures of collapse.

Most authors suggest surgery for Kienböck’s disease; however, conservative treatment with cast immobilization has been used in the acute or early stages (119). Many surgical procedures have been described, including joint leveling (either shortening the radius or lengthening the ulna) (120), various intercarpal fusions (121), silastic arthroplasty (122), or vascularized bone grafts (123). Most importantly, Kienböck’s should be treated immediately to limit further compression of the lunate and collapse of the carpus.

DUPUYTREN’S DISEASE

Dupuytren’s contracture is a thickening of the palmar fascia. It affects men 7 to 15 times more than women and is thought to be an autosomal dominant inheritance, with variable penetrance (124,125). Usually, Dupuytren’s affects whites of northern European descent, and its prevalence increases with age (126,127). The fascial thickening progresses ulnarly to radially and can result in flexion deformities of the MCP, PIP, and, occasionally, the DIP joint.

Dupuytren’s contracture is a benign hypertrophy of the fascia. It usually begins insidiously as small imperceptible nodules in the area of the palmar crease. It can progress to thick cords that form along the fascia tension lines of the palm (128). The underlying tendons, synovial sheaths, and skin layers are not affected (129).

The pathophysiology of Dupuytren’s is not fully understood; there is, however, a higher incidence in alcoholic, diabetic, and epileptic patients, and there is thought to be a correlation with tobacco use (130). A clear association between Dupuytren’s and trauma or work activities has not been identified (131). The palmar fascia thickening is caused by an abnormal proliferation of fibroblasts (132). This proliferation is closely correlated with that observed in scar formation and healing. Three stages in the nodule and cord formation have been described. The first stage is proliferation. During this stage, the numbers of myofibroblasts within the palmar fascia spontaneously increase. The second stage is involution, when the myofibroblasts align along the tension lines of the palm and digits. The fascia enlarges owing to contraction of the myofibroblastic activity. In the third phase, the myofibroblasts resolve, leaving contracting collagen, which is perceived as nodules and matures into cords (133).

Many individuals with this condition are unaware of its presence. Usually, it is first noted when the palmar nodules and cords become tender with pressure (128). Activities that require frequent or strenuous gripping often result in tenderness near the nodules or cords. The palmar fascia of the small finger is the most commonly affected in 70% of individuals with this condition (134). This is typically followed in succession by the fascia of the ring, middle, thumb, and index fingers. Other pathologic processes with a similar
Dupuytren’s contracture is a clinical diagnosis that is made upon exclusion of the above conditions. The hallmark of clinical evaluation is the palpable nodules and cords in the palmar fascia, most notably in the small finger. Hypertrophic changes of the skin are also appreciated (Fig. 37-9). Joint deformity, including flexion contractures of the MCP, PIP, and DIP, is usually present in advanced conditions. Transverse or webspace contractures may also occur. These contractures can result in significant functional limitations necessitating treatment.

Multiple treatment regimes for this condition have been described, including splinting, radiation, and use of disulfides, vitamin E, anti-gout medications, physical therapy, and ultrasonic therapy (127). These therapies have shown minimal effectiveness (135). Treatment of advanced Dupuytren’s is surgical fasciectomy. This is recommended when the patient can no longer perform a “tabletop test” (136). In this test, the individual places the palm on a flat surface and attempts to extend the involved finger actively. A positive test is noted if the MCP joint cannot be placed flat against the surface. This usually correlates with a greater than 30-degree fixed flexion contracture of the MCP joint. The goal of surgery is to restore function, not to cure the disease (136).

Recently, there has been a great deal of interest in percutaneous or enzymatic fasciotomies as an alternative to surgical fasciectomy. Hurst and Badalamente (137) have demonstrated that by injecting collagenase into the fibrous cords, joint contractures can be improved. They report that 90% enjoyed excellent results at an average of 9-month follow-up. Although no long-term studies have been completed, this procedure does offer promise.

Postoperative surgical rehabilitation is extremely important following fasciectomy, with concentration on maintaining skin integrity, restoration of joint range of motion, and overall improvement of function (138,139). Despite surgical treatment, this condition can be quite recalcitrant, and recurrence rates range from 28% to 80% (134).

COMPLEX REGIONAL PAIN SYNDROME

Complex regional pain syndrome (CRPS; previously known as reflex sympathetic dystrophy or RSD) is a neuropathic pain syndrome characterized by autonomic dysfunction and severe pain that may lead to crippling contractures of the limbs. Mitchell first described CRPS during the American Civil War when he observed wounded veterans who had burning pain in an injured limb (140). Since that time, many labels have applied to what is now called CRPS or some subset of CRPS. The most common of these are RSD, causalgia (minor and major), algodystrophy, shoulder-hand syndrome, and Sudeck’s atrophy. In 1993, at the meeting of the International Association for the Study of Pain (IASP), a task force proposed a unifying classification for these syndromes (141). The task force proposed a definition based on four criteria (Table 37-4).

The IASP proposed dividing the classification into types I and II, with type II being characterized also by injury to a peripheral nerve.

The primary treatment for CRPS is the use of physical therapeutics (142). In treating CRPS, one follows the classic order of rehabilitation beginning with pain and edema control, followed by range of motion, then strengthening, and then function. It is important to convey to the patient that immobilization is not an effective treatment for the pain and swelling; in fact, it may be instrumental in the pathogenesis and chronicity of the process.

Edema control entails elevation, decongestive massage, and various forms of compressive wrapping or garments. Pain control may be difficult using physical modalities alone. However, physical modalities should be the first line of defense. Contrast baths, Fluidotherapy, transcutaneous electrical nerve stimulation (TENS), and desensitization may be used before and after therapy session or exercise. If these are unsuccessful in adequately controlling the pain to the point at which therapy can be progressed, then one may consider further pain-relieving measures. Typical oral medications that may be used are tramadol, gabapentin, amitriptyline, and various α1-blockers. In about half of all cases, further augmentation of analgesia may be attained by injections such as stellate ganglion blocks. One may also use injections such as intravenous regional blocks, axillary blocks, and cervical epidural injections. These blocks may provide temporary pain relief, enabling the patient to begin more aggressive hand therapy. Once pain is controlled to the level that patients can tolerate therapy, then one may begin exercises.

The next goal of CRPS treatment is to restore normal range of motion. Often, the enduring disabilities resulting from CRPS are hand contractures. Gentle active or active-assisted range of motion should begin in a pain-free fashion. Any advancement in therapy should proceed slowly and carefully, keeping...
FOCAL DYSTONIA

Writer’s cramp and musician’s cramp are both focal dystonias that affect a discrete anatomical area. They are characterized by disabling cramps, contractions, or spasms during specific activities (145). When not so engaged, the hand appears and functions normally. The flexors are more commonly involved than the extensors. Among the flexors, the FDS, FDP, FPL, and the lumbricals may be involved. The extensor pollicis longus, extensor indicis, and digitorum communis may be involved among the extensors. Dystonia may occur sporadically in the population or may be genetically transmitted. The gene for early onset dystonia (DYT1) has been sequenced. Approximately 10% of people with dystonia have a family history of tremor or dystonia (146). Others report that a higher percentage of those affected have a family history of dystonia (147).

The incidence of writer’s cramp is reported to be 2.7 per million in Rochester, MN (148). It tends to affect male young adults. It is usually idiopathic and not a result of overt trauma, although it may follow a traumatic episode. Patients frequently have mirror dystonia, demonstrated by inducing the writer’s cramp in the dominant hand even when attempting to write with the nondominant (146). Focal dystonias tend to remain focal and do not become generalized dystonias over time.

The pathophysiology of dystonia is not entirely understood. However, there seems to be some evidence for abnormalities in the basal ganglia (145) or problems with cortical organization (149). Electrodiagnostic studies show a co-contraction of muscle and a loss of alternation of agonist/antagonist muscle contractions. There are prolonged bursts of muscle contractions and overflow contraction seen in those muscles not activated by the motor task (150). There is a neurophysiologic defect in the ability to produce and process neurotransmitters, including (a) gamma-aminobutyric acid (GABA), an inhibitory substance that helps the brain maintain muscle control; (b) dopamine, an inhibitory chemical that influences the brain’s control of movement; and (c) acetylcholine, norepinephrine, and serotonin.

Treatments have been aimed at correcting these possible pathological situations. GABA-regulating drugs such as the muscle relaxants diazepam, lorazepam, clonazepam, and baclofen have been used with modest success. Deep brain stimulation (151) and proprioceptive retraining (152,153) have anecdotally been reported as successful.
Currently, it appears that botulinum toxin offers the most reliable relief for focal dystonia, with minimal risk (154). Selection of patients for this treatment should include a careful neuromusculoskeletal examination. Once the diagnosis of focal dystonia is confirmed, the selection of the muscles that are to be injected is made. Often, these muscles can be identified by observing the patient performing the inciting tasks. However, it may be useful to use EMG to identify all those that are activated, including the deeper muscles that may not be obvious clinically. A starting dose of 2.5 U of botulinum toxin type A for small muscles and up to 50 U for larger muscles is usually injected. The botulinum toxin can be diluted into a larger volume for ease of administration. The use of EMG or ultrasound to guide injections increases the accuracy of correctly identifying the contracting muscles (155). The benefits are usually seen within 1 week and may last for a few months. Repeated injections are safe and are often needed to sustain the benefits (156). A secondary effect of botulinum type A injections is an associated muscle weakness, which may respond to muscle strengthening. A small proportion of patients receiving continued treatment may develop antibodies to type A. These patients may respond to botulinum type B or F (157,158).

REFERENCES


Foot Disorders

INTRODUCTION

Foot pain and problems are common among people of all ages. The prevalence has been reported as 15% to 40%, depending on the demographic distribution of the patient population (1–3). Foot disorders in the younger age group are frequently the result of trauma or congenital anomalies, and require a different approach to management than the elderly adult. The published literature on foot pain in the adult/aged population is quite extensive and indicates that it affects between 20% and 30% of community-dwelling elders (4). Some studies have shown that foot pain is associated with decreased ability to perform activities of daily living (5), problems with balance and gait (6–8) and increased risk of falls. Others have not found that to be the case (9). In fact, they reported that while there are a significant number of foot problems facing the elderly, only plantar fasciitis and a pes cavus foot profile are associated with functional foot problems. Another study has shown that elders with rheumatoid arthritis and who have significant foot pain are more likely to report back, hip, hand, and wrist pain. They also have pes planus and loss of motion in the ankle joint (10).

While it is useful to understand the issues of prevalence of foot pain and problems, and their associations with function, these must be understood within the context of the specific population (e.g., diagnostic group, age, etc.) of interest.

EXAMINATION

Physical Examination

The foot examination begins with an assessment of all the segments of the lower extremity. This should include measures of muscle strength, range of motion (ROM), alignment, stride characteristics, neurological examination (e.g., sensation, vibration and proprioception), vascular status, skin and nails, and review of shoe selection and wear patterns.

ANATOMY

Examination of the foot is challenging because of its complex anatomy, its biomechanical properties, and its function. This section attempts to describe important anatomic information that will enable the clinician to understand the affected anatomic site, but the reader is referred to books on anatomy, which present a more in-depth review (11–13).

Each foot is composed of 26 bones and 55 joints with muscle attachments that have their origin either within the foot or from the anatomical structures above the ankle. It is easier, therefore, to divide the foot into sections: the forefoot, midfoot, and rearfoot-ankle complex. This simplifies the examination process and provides the foundation necessary to understand biomechanical principles. The forefoot consists of 5 metatarsals and 14 phalanges constituting the five digits. The midfoot includes three cuneiforms, cuboid, and navicular. The rearfoot-ankle complex is composed of the talus, calcaneus, and distal ends of the tibia and fibula (Fig. 38-1). These three divisions form a tenuous interrelationship that can easily be disrupted, resulting in dysfunction and pathology.

The hallux contains two phalanges, and the remaining four digits contain three each. The hallux has one joint: the interphalangeal joint. The lesser digits form the proximal and distal interphalangeal joints. There is a tendency for the intermediate and distal phalanges of the fifth toe to be fused together. The first metatarsal head distally articulates on the plantar surface with the tibial and fibula sesamoid bones, which are within the flexor hallucis brevis tendon. There are three cuneiforms that articulate with the base of metatarsals one through three. The cuboid articulates with the fourth and fifth metatarsal bases. The five metatarsals, along with the three cuneiforms and cuboid, form Lisfranc’s joint. Injury to this joint often is missed at initial presentation with significant associated morbidity. The keystone of this joint is the second metatarsal cuneiform articulation. The base of the fifth metatarsal, along with the lateral styloid process, serves as the insertion site for the peroneal brevis tendon. Injuries to the base of the fifth metatarsal and the styloid process are common, especially after inversion injuries of the ankle. The configuration of the articulation between the metatarsals, cuneiforms, and cuboid helps form the transverse metatarsal arch. Proximally, the cuneiforms articulate with the navicular, and laterally, the cuboid articulates with the navicular and lateral cuneiform. Medially, on the navicular, there is an enlarged tuberosity that serves as part of the insertion points of the tibialis posterior tendon. The navicular and cuboid proximally articulate with the talus and calcaneus, respectively, forming the midtarsal joint. The calcaneus is the largest bone of the foot. It functions as a major weight-bearing bone and serves as the insertion site posteriorly for the Achilles tendon. Many of the intrinsic muscles
Motor and sensory nerves to the foot arise from branches of the sciatic nerve. The sciatic nerve divides into the tibial nerve, common peroneal nerve, and sural nerve. The common peroneal nerve divides into the superficial and deep peroneal nerves. They provide sensory input to the dorsal aspect of the foot and digits. The sural nerve innervates the lateral aspect of the foot. The saphenous nerve innervates the medial aspect of the foot and is a branch of the femoral nerve. The motor component of the peroneal nerve innervates the peroneus longus, peroneus brevis, extensor hallucis longus, extensor digitorum longus and brevis, tibialis anterior, and peroneus tertius. The tibial nerve divides into the medial and lateral plantar nerves. They provide sensory and motor innervation to the plantar aspect of the foot.

The feet are divided into three functional units or parts: the hindfoot (talus and calcaneus and its attachments) provides the foot with stability; the midfoot (navicular, cuboid, cuneiforms, Lisfranc’s and Chopart’s joints) provides both sagittal and frontal plane motion; and the forefoot (metatarsals and phalanges) permits push-off. The ligamentous structures establish and preserve the longitudinal arch, the intermetatarsal transverse arch, the position of the sesamoids, and stability of the ankle mortise.

Motor and sensory nerves to the foot arise from branches of the sciatic nerve. The sciatic nerve divides into the tibial nerve, common peroneal nerve, and sural nerve. The common peroneal nerve divides into the superficial and deep peroneal nerves. They provide sensory input to the dorsal aspect of the foot and digits. The sural nerve innervates the lateral aspect of the foot. The saphenous nerve innervates the medial aspect of the foot and is a branch of the femoral nerve. The motor component of the peroneal nerve innervates the peroneus longus, peroneus brevis, extensor hallucis longus, extensor digitorum longus and brevis, tibialis anterior, and peroneus tertius. The tibial nerve divides into the medial and lateral plantar nerves. They provide sensory and motor innervation to the plantar aspect of the foot.

The muscles of the foot are divided into extrinsic and intrinsic groups (Figs. 38-2 and 38-3). The extrinsic muscles arise in the leg and insert into the foot and are held in place by various retinacula as they enter the foot. They are divided into four compartments: superficial posterior, deep posterior, lateral, and anterior (Table 38-2). The intrinsic muscles of the foot are divided into dorsal and plantar groups. The plantar muscles are divided into four layers, with the first layer being the most posterior or superficial and the fourth layer the most anterior or deep (Table 38-3).

**Functional Anatomy**

The foot is divided into three functional units or parts: the hindfoot (talus and calcaneus and its attachments) provides the foot with stability; the midfoot (navicular, cuboid, cuneiforms, Lisfranc’s and Chopart’s joints) provides both sagittal and frontal plane motion; and the forefoot (metatarsals and phalanges) permits push-off. The ligamentous structures establish and preserve the longitudinal arch, the intermetatarsal transverse arch, the position of the sesamoids, and stability of the ankle mortise.
The subtalar joint is composed of the talus and calcaneus. The bones articulate at the posterior, middle, and anterior facets. The joint is responsible for triplane motion, primarily pronation and supination. There is 20 degrees of inversion to 10 degrees of eversion to the subtalar joint. In closed kinetic chain pronation, the calcaneus everts, with adduction

### Table 38.2 Extrinsic Muscles of the Foot

<table>
<thead>
<tr>
<th>Extrinsic Muscle Groups</th>
<th>Muscles</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Superficial posterior</td>
<td>Gastrocnemius</td>
<td>Plantar flexion</td>
</tr>
<tr>
<td></td>
<td>Soleus</td>
<td>Plantar flexion</td>
</tr>
<tr>
<td></td>
<td>Tibialis posterior</td>
<td>Plantar flexion</td>
</tr>
<tr>
<td></td>
<td>Flexor digitorum longus</td>
<td>Plantar flexion</td>
</tr>
<tr>
<td></td>
<td>Flexor hallucis longus</td>
<td>Eversion</td>
</tr>
<tr>
<td>Deep posterior</td>
<td>Peroneus longus</td>
<td>Eversion</td>
</tr>
<tr>
<td></td>
<td>Peroneus brevis</td>
<td>Eversion</td>
</tr>
<tr>
<td></td>
<td>Tibialis anterior</td>
<td>Eversion</td>
</tr>
<tr>
<td>Lateral</td>
<td>Extensor hallucis longus</td>
<td>Dorsiflexion</td>
</tr>
<tr>
<td>Anterior</td>
<td>Extensor digitorum longus</td>
<td>Dorsiflexion</td>
</tr>
<tr>
<td></td>
<td>Peroneus tertius</td>
<td>Dorsiflexion</td>
</tr>
</tbody>
</table>

### Table 38.3 Intrinsic Muscles of the Foot

<table>
<thead>
<tr>
<th>Intrinsic Muscle Groups</th>
<th>Muscle</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dorsal</td>
<td>Extensor digitorum brevis</td>
<td>Extension of toes</td>
</tr>
<tr>
<td>Plantar</td>
<td>Extensor hallucis brevis</td>
<td>Extension of hallux</td>
</tr>
<tr>
<td>Layer 1</td>
<td>Abductor hallucis</td>
<td>Abduction of hallux</td>
</tr>
<tr>
<td></td>
<td>Flexor digitorum brevis</td>
<td>Flexion of toes</td>
</tr>
<tr>
<td></td>
<td>Abductor digit minimi pedis</td>
<td>Abduction of fifth toe</td>
</tr>
<tr>
<td>Layer 2</td>
<td>Quadratus plantae</td>
<td>Helps in flexion of toes</td>
</tr>
<tr>
<td></td>
<td>Lumbricles (4)</td>
<td>Flexion MTPJ/extension IPJ</td>
</tr>
<tr>
<td></td>
<td>Tendons of long flexors</td>
<td>Flexion of hallux</td>
</tr>
<tr>
<td>Layer 3</td>
<td>Flexor hallucis brevis</td>
<td>Flexion of hallux</td>
</tr>
<tr>
<td></td>
<td>Adductor hallucis</td>
<td>Adduction of hallux</td>
</tr>
<tr>
<td></td>
<td>Flexor digit minimi brevis</td>
<td>Flexion of fifth toe</td>
</tr>
<tr>
<td>Layer 4</td>
<td>Plantar interosseis</td>
<td>Adduction of toes</td>
</tr>
<tr>
<td></td>
<td>Dorsal interosseis</td>
<td>Abduction of toes</td>
</tr>
<tr>
<td></td>
<td>Tendons of peroneal longus</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tibialis posterior</td>
<td></td>
</tr>
</tbody>
</table>

IPJ, interphalangeal joint; MTPJ, metatarsal phalangeal joint.
and plantar flexion of the talus. The opposite is true for closed kinetic supination (Fig. 38-4).

The position of the subtalar joint dictates the position of the midtarsal joint. When the subtalar joint supinates, the arch raises the forefoot, and the foot becomes more rigid. The midtarsal joint “locks.” When there is subtalar joint pronation, the midtarsal “unlocks” and becomes less rigid. The foot becomes a mobile adaptor, allowing for increased motion to take place to accommodate for varying terrain.

The midtarsal joint is made up of the calcaneocuboid and talonavicular joints. The midtarsal joint has two axes of motion: the longitudinal and oblique axes. The longitudinal axis inverts and everts the forefoot while pronation or supination occurs. The oblique axis allows for adduction, abduction, plantar flexion, and dorsiflexion of the foot.

**Physical Examination of the Foot**

ROM assessment of the lower extremity is important because absence of normal range has functionally important sequelae. For example, a tight Achilles tendon, which limits ankle dorsiflexion, creates a force anterior to the axis of the ankle, which results in pronation, or collapse of the medial column. Limitation of hallux dorsiflexion to less than 25 degrees creates an impediment to forefoot push-off.

When testing ROM for the ankle, there should be at least 10 degrees of dorsiflexion with 45 degrees of plantar flexion. When testing the range of dorsiflexion, the examiner should hold the foot medially. This keeps the midtarsal joint locked, and a more accurate measurement can be achieved. Dorsiflexion is measured with the knee extended and flexed. When the knee is flexed, there is more dorsiflexion of the ankle because of the “unlocking” of the gastrocsoleus complex. The ankle motion should be smooth and unrestricted. The subtalar joint should measure 20 degrees of inversion and 10 degrees of eversion. This can be measured with the patient in the prone position. The heel is bisected in half, and the examiner can then use a tractograph and invert and evert the foot for the measurement. The metatarsophalangeal joints, the distal and proximal interphalangeal joints of the lesser toes, and the interphalangeal joint of the hallux should have smooth and unrestricted dorsiflexion and plantar flexion.

Muscle strength testing provides important information about whether the foot is likely to be supported in stance (static) and dynamic phases of gait.

Support of the foot in stance comes from the ligaments and plantar fascia, which stabilize the longitudinal arch and the intrinsic muscles of the foot. The creation and establishment of the longitudinal arch in the dynamic phase depend on the posterior tibialis and intrinsic and extrinsic muscles of the foot. The gastrocsoleus provides propulsion through sagittal plane motion at the ankle. Weakness in these structures may help explain an atropulsive gait or instability during stance.

Determination of static alignment of the foot provides important information about how the foot will contact the ground. The preferred position is that the foot should be parallel to the floor. If ligamentous structures are intact, ROM abnormality or muscle imbalance is likely to explain alignment problems. Examination demonstrates possible forefoot varus or valgus deformity. This is performed with the patient in the prone position with the subtalar joint neutral.

The feet should be evaluated in the standing position, and the arch height is evaluated as flat, low, normal, or high, based on the navicular position. The angle of the talus measured radiographically may be used to assess arch height. Bisecting the Achilles tendon will demonstrate whether the calcaneus is everted or inverted with respect to the midline. A varus refoot generally causes pressure to the posterosuperior lateral aspect of the calcaneus. A bone spur can develop in this area and cause pain. This is known as a Haglund’s deformity. Severe calcaneoeverion can cause a lateral impingement and pain to palpation below the lateral malleolus.

The vascular exam will evaluate peripheral pulses in the foot. Dorsally, there is a dorsalis pedis pulse, and medially a posterior tibial pulse. If one cannot feel a pulse and there is questionable vascular compromise, the physician can order further vascular testing such as noninvasive arterial Doppler. This exam would be beneficial when dealing with lower extremity ulcers. Capillary filling time is tested and should be within 2 seconds. Press the distal tip of the digit and determine how fast the digit fills up with blood. Another area of testing would be for dependent rubor. Dangle the foot off the examining table and see if the foot becomes more red or purple. If this occurs, there may be pathology to the venous system of the lower extremity. Any varicosities should also be noted to the lower extremities, as well as chronic swelling. Patients with brawny edema from venous stasis disease may develop ulcerations to the medial lower leg or ankle.

The neurological exam will test deep tendon reflexes, ankle clonus, Babinski’s sign, sharp and dull discrimination, as well as vibratory sense and proprioception. Patients who are neuropathic have diminished sensation, vibratory sense,
and lack of proprioception. It is useful to test for protective sensation using Semmes-Weinstein filaments. Those who cannot feel the touch of a nylon filament 5 mm in diameter are at significant risk for ulceration. They are at risk for ulcerations because they lack sensation to the foot and do not feel the pressure caused by a foreign body, minor cut, abrasions from poorly fitted shoes, or thickened calluses.

Examination of the skin indicates where pressure points are. Callusites are in areas of high pressure and often need to be relieved.

**DIAGNOSTIC TESTING**

The proper selection of imaging techniques enhances the ability to diagnose foot disorders. When dealing with foot pain, one should be knowledgeable about selecting from among plain radiographs, computerized tomography, magnetic resonant imaging (MRI), and ultrasound. The application of diagnostic ultrasound has offered clinicians and clinical investigators opportunities to increase the sensitivity and specificity of patient evaluation and assessment of treatment outcomes.

A good review of which diagnostic tools are best applied to which clinical conditions is available for the reader (14). A summary table of recommendations (see Table 38-1) is given.

Ultrasound has recently received attention because of its usefulness in clinical settings. Sonography has been shown to provide valuable information about soft tissue, including muscle, fascia, tendon, blood vessels, and nerves. Ultrasound is convenient, low cost, and low risk and may be particularly useful in helping to diagnose and follow problems associated with plantar fascitis and posterior tibialis tendon dysfunction, significant clinical problems (14,15).

**BIOMECHANICS**

Foot motion depends on intrinsic and extrinsic muscles that are divided into four compartments (see Table 38-2). These muscles produce sagittal, frontal, and transverse plane motion.

To understand foot biomechanics, the physician must be proficient in the terms used to describe the motion and position of the foot. A list of terms commonly used follows:

- **Abduction**: movement of the foot away from the midline of the body
- **Adduction**: movement of the foot toward the midline of the body
- **Eversion**: turning inward of the feet with the soles pointing away from each other
- **Inversion**: turning inward of the feet with the soles pointing toward each other
- **Dorsiflexion**: motion whereby the distal end of the foot moves away from the tibia

**Plantar flexion**: motion occurring to the foot whereby the distal end of the foot moves away from the tibia

**Pronation**: triplane motion of the foot in the direction of abduction, eversion, and dorsiflexion

**Supination**: triplane motion of the foot in the direction of adduction, inversion, and plantar flexion

**Subtalar joint**: point at which the normal foot is neither inverted nor everted

**Neutral position**: from the floor during stance; clinically, whereby the posterior aspect of the calcaneus lines up with the lower one third of the leg

**Forefoot varus**: inversion to the plantar surface of the foot with the subtalar joint neutral and the midtarsal joint pronated; this is a congenital bony deformity

**Forefoot valgus**: eversion to the plantar surface of the foot when the subtalar joint is neutral and the midtarsal joint is pronated; the plane of the metatarsals are everted with the first metatarsal lower than the fifth metatarsal; the first metatarsal may be flexible or rigid; this is a congenital bony deformity

**Rearfoot varus**: inverted position of the sagittal plane and the plantar surface of the calcaneus to the weight-bearing surface with the subtalar joint in the neutral position and the body standing in normal base of gait (10 degrees to 15 degrees of abduction)

**Rearfoot valgus**: the sagittal plane of the posterior and plantar surface of the calcaneus are everted to the weight-bearing surface when the subtalar joint is in neutral position while the body is standing in base of gait

Often, with increased subtalar joint pronation, the foot becomes hypermobile, and increased stress and strain are placed on the lower extremities. A functional valgus occurs at the knee, causing increased stress to the medial aspect of the knee. Hyperpronation causes increased strain on the plantar fascia, causing a pull on the origin of the fascia of the calcaneus. Patients are predisposed to bunions, hammertoes, and postural fatigue with increased subtalar joint pronation. Posterior tibial and Achilles tendinitis can occur with a hypermobile foot because of the increased stretching to the tendon.

A supinated foot or high-arched foot is much more rigid. This type of foot structure predisposes the patient to lateral ankle sprains, lateral Achilles tendinitis, claw toes (flexion deformity at the proximal and distal interphalangeal joints), increased incidence of avulsion fractures to the base of the fifth metatarsal, peroneal subluxation, metatarsalgia, and sesamoid pain or fractures.

**MECHANICAL DYSFUNCTION**

**Posterior Tibialis Tendinitis**

**Anatomy and Pathomechanics**

The posterior tibial tendon has its origin from the lateral posterior aspect of the tibia, medial aspect of the body of the fibula, and interosseous membrane of the leg (16). The tendon courses under the medial malleolus and into the foot, with
most of the tendon inserting into the navicular tuberosity. The remaining smaller slips of the tendon insert into all tarsal bones except the talus. The posterior tibial tendon is a plantar flexor and inverter of the foot and a strong supinator of the subtalar and midtarsal joints. The tendon dynamically supports the medial longitudinal arch (17). The peroneus brevis tendon is the antagonist to the posterior tibial tendon (18), which is an adductor at the level of the midtarsal joint (19).

When there is weakness of the posterior tibial tendon, the foot becomes more abducted because of the overpowering effect of the peroneus brevis (20). When the posterior tibial tendon weakens, the midtarsal joint does not lock, causing increased pronation at the subtalar joint. In time, the medial longitudinal arch collapses, and there is increased shock at the rearfoot, causing a decrease in the propulsive activity of the foot (19).

Etiology
Posterior tibial tendon ruptures may be traumatic or degenerative in nature. Acute traumatic ruptures are not common, but may occur from puncture wounds, lacerations to the medial ankle (21), and severe eversion ankle sprains (22). Most cases of dysfunction are degenerative (16), such as inflammatory arthritis (23). Diabetes mellitus, obesity, and prior local steroid exposure have been known to cause a degenerative process to the tendon (14). Tendon hypovascularity at the midportion of the tendon just distal to the medial malleolus has been known to cause ruptures to the posterior tibial tendon (14). Abnormal biomechanical forces (equinus, pes valgus) may cause chronic tenosynovitis and weakening of the posterior tibial tendon (15).

A classification system for dysfunctional posterior tibial tendon etiologies has been formulated by Mueller and is based on four categories (19). Type I is a direct rupture due to direct injury to the tendon. Type II is a result of a systemic disease such as diabetes or inflammatory arthritis. Type III is idiopathic or possibly degenerative. Type IV is functional and may be related to severe pronation or stretching of the tendon.

Clinical Findings
The clinical concerns of the patient may depend on the chronicity of the deformity. In the acute stage, there is swelling and pain along the course of the posterior tibial tendon. The patient may experience the symptoms as an ankle injury (16). In the chronic stage, the patient notices a change in the appearance of the foot. There may be a collapse of the arch, or “too many toes” sign, whereby the toes are pointing outward as the patient is seen from behind (see Fig. 38-4). The patient’s gait becomes increasingly apropulsive, with limited heel lift and toe off. There is excessive wear on the medial aspect of the last of the shoe. The patient demonstrates little or no inversion power off. There is excessive wear on the medial aspect of the last. In the chronic patient, there might be crepitus, with painful ROM of the ankle and subtalar joint. There may be arthritic changes noted on x-ray to the ankle or subtalar joint.

Diagnostic Imaging
Standard radiographs do not demonstrate a tendon tear but distinguish any adaptive changes to the subtalar or ankle joints. Standard weight-bearing dorsoplantar, oblique, and lateral x-rays should be ordered in patients displaying pain and weakness in the posterior tibial tendon. In long-standing deformities, there are arthritic changes in the subtalar and ankle joints. MRI is one of the most accurate tests that can depict the extent of pathology of the tendon (25). CT can depict the condition of the tendon; however, the definition of tendon outline is not as clear. CT scans can be helpful in identifying osseous deformities such as degenerative arthritis and subtalar joint dislocations. MRI aids the practitioner with excellent resolution of the image regarding tendon outline, synovial fluid, and vertical splits in the tendon. Axial and sagittal images demonstrate tendon pathology well (26). The axial view helps delineate tendon girth and shape, fluid accumulation, and signal changes within the tendon itself (27). Sagital images allow the practitioner to view the extent of the pathology longitudinally. Coronal images are not as helpful in delineating posterior tibial pathology (16). Ultrasound is useful for describing the course of the tendon and can identify whether there is a tear, tendinitis, tendinosis, and peritendinitis (28,29).

In the axial view, the tendon is normally round or oval and is twice the size of the flexor digitorum longus tendon. In the normal tendon, there should be no accumulation of fluid surrounding the tendon (26). A tear on the axial image indicates mixed signal intensities, with the size of the tendon increasing or decreasing, depending on the extent of the tear.

Treatment
Conservative treatment is reserved in those patients who are older or who lead sedentary lives. Conservative modalities include removable walking casts (Cam walkers), custom foot orthotics (to control poor biomechanics), braces, ankle-foot orthotics, orthopedic shoes with medial heel wedges, and compression stockings (16). The goal for conservative treatment is to reduce the pain and inflammation and to delay the progression of the deformity. If conservative treatment fails, surgical options should be explored. Surgery in the early phase before significant tear may include a synovectomy to the tendon sheath. After a tear, tendon transfers (flexor digitorum longus) or an isolated fusion of the subtalar joint or triple arthrodesis (fusion of subtalar and midtarsal joints) may be required.
Achilles Tendinitis

Anatomy
The Achilles tendon is composed of the gastrocnemius and soleus muscles (triceps surae). It forms a spiral configuration that inserts into the central one third of the posterior calcaneus. A retrocalcaneal bursa can be found between the calcaneus and Achilles tendon. The Achilles tendon is covered by a paratenon. This paratenon is a loose elastic tissue that stretches with the Achilles tendon. The central layer of the paratenon (mesotenon) is responsible for the tendon’s blood supply. The action of the tendon is to decelerate and stabilize the foot, and then accelerate the foot during the gait cycle (16).

Mechanism of Injury
Poor foot and ankle mechanics and overuse can predispose the patient to injury. A fixed forefoot equinus results in compensation at the ankle joint. The ankle compensates with dorsiflexion, allowing the forefoot and rearfoot to remain on the ground during the midstance phase of gait. This may result in an overuse injury to the tendon. Compensation at the subtalar joint for any forefoot to rearfoot imbalance places strain on the Achilles tendon. If the patient has a high forefoot varus deformity, the subtalar joint may compensate by evertting the calcaneus (hyperpronation) and placing more load on the medial side of the Achilles tendon. A rigid planar-flexed first metatarsal or cavusvarus places strain on the lateral (hypersupination) side of the Achilles tendon (16).

Physical Findings and Treatment
There are three distinct areas of pathology to the Achilles tendon: the insertion site of the tendon into the posterior one third of the calcaneus with or without bone involvement, the midportion (vulnerable zone) with peri-tendinitis or partial or total rupture, and the myotendinous junction (30). Most injuries are related to overuse or poor foot biomechanics. The patients will give a history of increased activity or intensity to their workouts or beginning an exercise program after a period of inactivity.

Insertional Achilles tendonitis may result from a bony abnormality to the calcaneus. Patients describe their Achilles as stiff in the morning. A bony prominence to the posterior superior lateral aspect of the calcaneus (Haglund’s deformity) or insertional calcification may cause pain to the insertion site. This is quite evident in certain shoes rubbing on the area. Surgical management of this problem would be the last resort.

The patient may begin home treatment consisting of rest, ice, anti-inflammatory medication, night splints, deep transverse friction massage, and a ¼-in. heel lift to the shoes. Shoes with a lower heel counter or sneakers with a U-shaped Achilles notch to the back of the heel counter might help reduce the inflammation to the area. Stretching and strengthening exercises for the posterior leg and hamstring muscles will allow the Achilles tendon to function more efficiently. An ankle dorsiflexion stretch held for 30 seconds 10 times twice a day may help with the symptoms of pain. Home strengthening exercises using TheraBand or toe raises can be beneficial as well.

Physical therapy might be helpful in cases in which the patient is not improving on his or her own along with a removable walking cast (Cam walker). Therapy may include ultrasound, iontophoresis, deep transverse friction massage, and stretching and strengthening exercises. It is our recommendation not to inject cortisone into the Achilles tendon. The crystals of the cortisone may weaken and rupture the tendon.

Midportion pain to the tendon presents with fusiform swelling. This area is vulnerable to rupture because of the diminished blood supply. MRI is helpful in determining the extent of injury to the area. Treatment includes physical therapy (deep transverse friction massage, ultrasound, stretching exercises, and ice). Cast immobilization may be necessary in the beginning to help rest the area. We prefer removable walking casts such as a Cam walker. Home exercises for the patient should include stretching the Achilles, with ice and massage to the area.

Myotendinous junction injuries are not serious and heal readily with rest and physical therapy. In the more acute or chronic cases, a walking cast may be needed for pain relief. This type of injury heals within 4 to 6 weeks.

Peroneal Tendinitis

Anatomy
The peroneus brevis originates from the lateral distal two thirds of the fibula and inserts into the lateral base of the fifth metatarsal. The peroneus brevis is the strongest abductor of the foot as well as a flexor to the ankle and evertor of the foot (31). The muscle stabilizes the lateral segment of the foot, specifically at the calcaneocuboid joint.

The peroneus longus muscle originates at the proximal two thirds of the lateral fibula. The tendon passes inferior to the lateral malleolus and under the cuboid and courses medially under the foot to insert into the lateral plantar base of the first metatarsal and first cuneiform. The peroneus longus is a plantar flexor and evertor of the foot. The tendon at its insertion helps stabilize the first metatarsal.

Mechanism of Injury
Injuries to the tendon may occur traumatically or insidiously from poor foot biomechanics. Forced inversion sprains may cause pain to the peroneal tendons. A patient who has a supinated foot is predisposed to this type of injury because there is increased stress to the tendons on the lateral ankle.

Clinical Findings and Treatment
Pain to the peroneal tendons may be present inferior to the lateral malleolus, at the peroneal tubercle, or at the peroneal groove under the cuboid where the peroneus longus courses medially to insert at the lateral base of the first metatarsal and first cuneiform. Peroneus brevis tendonitis may also be present at its insertion point on the lateral base of the fifth metatarsal. Passive dorsiflexion of the first metatarsal may be helpful in determining tendonitis of the peroneus longus. Passive inversion of the calcaneus with adduction of the forefoot is a good way of testing for peroneus brevis tendinitis.
Peroneal tendinitis may occur after an inversion injury to the ankle. Patients relate pain to the ankle, with pain inferior to the lateral malleolus or at the insertion site of the peroneus brevis. There may be swelling over the course of the tendons. Athletic patients relate pain to the peroneus longus tendon with increased cutting or turning or when getting up on the ball of the foot. Pain is noted under the cuboid or on the lateral aspect of the calcaneus.

Treatment is directed at reducing the inflammation and strengthening the muscles. Rest, anti-inflammatory medication, and ice help in the acute stage of the problem. Ice pops can be made by filling a cup with water, placing a stick inside, and freezing it. Deep transverse friction massage may help break up some of the scar tissue in the area. Eversion with resistance exercises and TheraBand help strengthen the peroneal muscles.

If biomechanical abnormalities are involved, custom foot orthotics can be fabricated. Orthotics for walkers can be of three-quarter length. For the active patient, a full-length orthotic device with a long forefoot runner’s post helps control the foot not only at heel strike and midstance but also at push-off. Posting of the orthotic in the rearfoot and forefoot helps keep the foot more balanced.

If the patient does not respond to conservative treatment, surgical intervention might be considered. Surgery would entail a tenosynovectomy. If the peroneal tendons are subluxing out of the groove under the lateral malleolus, the groove is usually deepened.

**Plantar Fasciitis**

**Anatomy**
The plantar fascia is a strong aponeurosis that originates on the medial process of the calcaneal tuberosity and fans out into a medial, central, and lateral slip. The fascia blends in with the flexor plate distally and has a connection with the plantar aspect of the toes. The central slip is the thickest, the lateral slip acts as a covering to the abductor digiti minimi muscle, and the medial slip acts as a covering to the abductor hallucis muscle (13).

The plantar fascia acts as a windlass and pulls the arch up with dorsiflexion of the metatarsophalangeal joints (32). Poor foot biomechanics can lead to increased tension on the fascia, causing inflammation and pain. This can occur with patients who have increased subtalar joint pronation, pes planus, and limited dorsiflexion. The increased pronatory forces stretch the plantar fascia and cause a pulling at the origin. This, in turn, causes pain and inflammation. Patients with a pes cavus foot type are also prone to this condition. These patients have a more rigid foot type and cannot absorb shock as well at heel strike, and that places more stress on the plantar fascia (33).

**Clinical Findings**
Patients typically describe their pain after getting out of bed in the morning or after a period of inactivity. They state that the pain decreases after walking on the foot for a while (34). Most patients tolerate the condition before seeking medical help.

On clinical exam, there is tenderness on the medial plantar aspect of the heel. Some patients may have discomfort under the midcalcaneal region of the heel. There may be tightness of the Achilles tendon insertion along with tightness to the medial slip of the plantar fascia. Poor foot biomechanics such as a pes cavus or planus deformity might be present, along with a compensated forefoot varus deformity. Upon x-ray exam, there may be a calcaneal heel spur forming parallel to the weight-bearing surface. The heel spur is not pathognomonic for this condition. The spur does not cause the pain to the heel. The heel spur is the by-product of the chronic pulling of the fascia off the calcaneus (35). Some patients might describe pain to the lateral ankle or dorsolateral aspect of the foot. This is probably related to the patient offloading the pressure to the heel and walking on the outside of the foot. This creates an overuse condition to the lateral foot and ankle structures.

Most patients respond favorably to conservative treatment. The goals of treatment are to reduce the pain and inflammation and to correct poor foot biomechanics. We prefer to treat this condition in a stepwise approach.

The first level of treatment is a simple stretching exercise for the Achilles tendon and plantar fascia. The patient stretches the Achilles tendon while still in bed with a large towel. The patient dorsiflexes the foot and holds that stretch for a minimum of 30 seconds 10 times. Stretch to the plantar fascia is done by dorsiflexing toes, holding the metatarsophalangeals, and stretching fascia in the arch region. During the day, the patient may stretch by leaning against a wall. Again, each stretch is held for a minimum of 30 seconds with the knees straight and the heels on the ground. The patient may roll the arch over a soup can or golf ball. Ice and deep transverse friction massage may also be beneficial, as may taping of the foot (35). Over-the-counter foot orthotics is also recommended. Night splints are beneficial to help keep a stretch on the posterior leg muscles and plantar fascia.

If the pain has not improved, physical therapy, cortisone injections, custom foot orthotics, and supportive shoes are discussed with the patient. Shoes should consist of adequate depth to accommodate an orthotic and should have good heel and medial countersupport. If a cortisone injection is indicated, we prefer injecting the heel with a 3-mL syringe and a 25-gauge, 1.5-in. needle. The steroid of choice is 0.5 mL of Celestone Soluspan or Depo-Medrol with 1.5 mL of 2% lidocaine (Xylocaine) plain. The practitioner should palpate for maximum tenderness and inject from a medial approach. The patient should not receive more than three steroid injections to the same area over the course of 1 year. Multiple injections to the same site can weaken the soft tissue to the area. Injections through the plantar aspect of the foot should be avoided because they are painful to the patient.

A fairly new method of treatment is extracorporeal shock wave therapy. This is a noncutting procedure performed under anesthesia (local anesthesia with intravenous sedation). It is not fully understood how this promotes healing; however, the use of electrohydraulic shock waves converts a chronic condition into an acute condition (36). Patients who do not want the traditional or endoscopic surgical procedure may consider this alternative because of the lower risks attached to the procedure.
The current standard before considering this type of procedure is 6 months of conservative care. Several studies report benefit in controlled, randomized trials without adverse effects up to 1 year following treatment (26,27).

**Hallux Abducto Valgus**

**Etiology**

Hallux abducto valgus occurs because of the positional and structural changes that take shape as a result of hypermobility of the first ray. It is our opinion that hallux abducto valgus is a hereditary condition caused by poor foot biomechanics. In the pronated foot, the stability of the peroneus longus is lost and the first ray becomes dorsiflexed and adducted. This creates an increase in the intermetatarsal angle between the first and second metatarsals. Finally, intrinsic instability occurs around the first metatarsophalangeal joint and the fibular sesamoid drifts laterally, creating lateral deviation of the great toe (37).

**Clinical Findings and Symptoms**

Hallux abducto valgus is a progressive deformity that worsens with time. Shoes contribute to the pain but not to the deformity. The patient may have pain on the dorsomedial aspect of the first metatarsophalangeal joint. There may be erythema and a small bursa noted at the medial aspect of the joint. The patient usually has a pronated foot with decreased arch height. Patients tell the practitioner that they recently noticed their “bump” has enlarged over a short period of time. X-rays reveal an increase in the intermetatarsal angle and lateral deviation of the sesamoids, with lateral deviation of the great toe. There may not be joint pain or decreased ROM to the joint.

**Treatment**

Treatment is designed to reduce the pain and to control the pronation that is causing the condition. Conservative treatment is the best initial approach. Although conservative treatment will not remove the enlarged bony prominence, the pain may decrease. Weight-bearing x-rays (anteroposterior, lateral, and medial oblique views) help the physician determine the extent of the deformity.

Ice and oral anti-inflammatory medications may help reduce inflammation. A local cortisone injection to the medial aspect of the joint may relieve pain. Over-the-counter or custom foot orthotics designed to control the biomechanical forces can help reduce pain and help realign the foot to move in a more efficient manner. If conservative therapy fails, surgical intervention is discussed. When the pain interferes with normal daily activities or if shoeing is a significant problem and conservative options have been exhausted, then surgery is considered.

There are a variety of surgical procedures, ranging from first metatarsal head and base osteotomies to fusions of the first metatarsophalangeal joint or first metatarsocuneiform joint. The goals in surgical intervention are to decrease pain, improve function, and establish a congruous joint, reduction in the intermetatarsal angle, realignment of the sesamoids, maintenance of first metatarsophalangeal joint ROM, and repositioning of the hallux to a rectus position (37). It is preferable to place patients in custom-made foot orthotics after surgery to control the factors that led to the development of this condition.

**Hallux Limitus and Hallux Rigidus**

Hallux limitus is mild to moderate decreased ROM to the first metatarsophalangeal joint. Hallux rigidus is severe limitation of motion. Either condition may or may not present problems to the patient. There have been cases (eliminate reference to personal practice) in which the patient has no motion to the joint but is pain free. The etiology of this condition may be traumatic or biomechanically induced.

**Joint Mechanics**

Sixty to seventy-five degrees of first metatarsophalangeal joint dorsiflexion is necessary during the gait cycle. During the first 20 degrees, the first metatarsophalangeal joint acts as a ginglymoarthrodial joint. The motion is hingelike in nature. As the heel lifts, the ground reactive forces dorsiflex the hallux and the first metatarsal head plantar-flexes. This allows for closed-chain dorsiflexion of the first metatarsophalangeal joint beyond the 20 degrees (16).

The distal end of the first metatarsal head articulates with the sesamoid apparatus, and the base of the proximal phalanx glides along the dorsal aspect of the first metatarsal head. If the sesamoid apparatus is bound down and immobile, there is restriction of dorsiflexion to the joint. Ultimately, there is impingement to the joint causing degenerative changes resulting in arthrosis and remodeling of the joint over a period of time (37).

**Etiology**

A variety of causes can contribute to limitation of motion to the first metatarsophalangeal joint. Biomechanical causes may be broken down into a long first metatarsal, long proximal phalanx, dorsiflexed metatarsal, or hypermobility of the first ray. Trauma or postsurgical bunion surgery may be another cause of this condition.

A long first metatarsal can cause jamming to the joint during the propulsive phase of gait. This causes degenerative changes to the joint. Hypermobility of the joint occurs with weakness to the peroneus longus. The peroneus longus inserts into the base of the first metatarsal and first cuneiform. Without proper first metatarsal plantar flexion, the head of the first metatarsal does not articulate with the sesamoids and the metatarsal dorsiflexes, causing limitation of motion to the joint. Congenital elevatus of the first metatarsal does not allow the first metatarsal to plantar-flex with propulsion, leading to degenerative changes (37). Acquired elevatus is seen after bunion surgery, specifically closing base wedge osteotomies of the first metatarsal. A complication to this type of bunion procedure is dorsiflexion of the first metatarsal.

Degenerative changes to the first metatarsophalangeal joint may occur with systemic illnesses such as gout, rheumatoid arthritis, psoriatic arthritis, and septic arthritis. Finally, trauma to the joint may cause degenerative changes to develop over time. An intraarticular fracture to the head of the
first metatarsal or base of the proximal phalanx may predispose the patient to early degenerative changes to the joint.

**Symptoms**
Limitation of motion and pain in the first metatarsophalangeal joint occurs gradually over time. The patient complains of increased pain with increased activities. Shoes with a more flexible sole cause increased pain to the joint. Activities that load the toes and squatting become painful. Activities that require dorsiflexion of the great toe are painful, and tennis players may have significant disability.

The pain often begins over the dorsolateral aspect of the joint. There may be subtle signs of increased bony spurring at the lateral aspect of the joint on x-ray. There may be swelling to the joint and, at times, a small bursa formation over the dorsal bone spur. Lateral radiographs demonstrate a dorsal flag sign to the first metatarsal head. Eventually, x-rays demonstrate hypertrophy and squaring to the joint, with bony osteophytes (Fig. 38-5) and possible autofusion. The interphalangeal joint may develop extensus deformity and fuse.

**Treatment**
The goal of conservative treatment is to reduce pain and to neutralize the deforming forces that are causing that pain. Anti-inflammatory medication and ice may give temporary relief to early symptoms. Cortisone injections in the area of irritation may also reduce the inflammation to the joint. We prefer to mix short-acting and long-acting corticosteroids with local anesthesia into a 3-mL syringe with a ½-in., 25-gauge needle. Not more than three injections per year are injected into the same site.

Custom-made orthotics with a Morton’s extension are useful in controlling the first metatarsophalangeal joint. Morton’s extension is a long post made out of rubber and is placed past the first metatarsophalangeal joint distally to the great toe. Stiffer-soled shoes and rocker-bottom soled shoes with the breakpoint of the rocker distal to the sulcus of the first proximal phalanx have also helped in reducing pain.

If conservative treatment fails, surgical intervention may be warranted. The goal of surgery is to reduce tension in the joint with mobilization of the sesamoid apparatus. This is important in procedures that are not designed to destroy the joint, such as in decompressive osteotomies of the first metatarsal.

Surgical techniques include cheilectomy, hemi-implants, total first metatarsal joint implant systems, joint destructive procedures such as Keller’s bunionectomy (removing base of proximal phalanx), plantar-flexor wedge osteotomies, and arthrodesis of the first metatarsophalangeal joint. All the listed procedures have inherent risks.

**FRACTURES**

**Sesamoid Fractures**
There are two sesamoids under the first metatarsophalangeal joint. There are tibial and fibular sesamoids. Each sesamoid bone lies within the tendon of the flexor hallucis brevis. On occasion, there may be a congenital bifurcation of these bones, and the presence of two to four sesamoid bones may be found under the first metatarsal head.

Fractures to the sesamoids can be traumatic or insidious in nature. Traumatic injuries usually are a result of forced dorsiflexion of the first metatarsophalangeal joint. Activities such as ballet dancing, high-impact aerobics, long distance running, soccer, tennis, and football lead to such injuries. The tibial sesamoid is commonly fractured more often than the fibular sesamoid because of the increased weight placed under the metatarsal head medially (16). There are times when a patient will describe pain under the foot with no direct episode of injury. Certain biomechanical conditions that cause constant stress to the sesamoid bones, such as a cavus foot, plantar-flexed first metatarsal, or peroneus longus spasm, can predispose a patient to fractures (38).

Patients typically present to the physician with pain under the first metatarsal head. There is usually swelling associated with the pain, and there may or not be a specific event triggering this pain. The patient may be seen guarding the motion to the first metatarsophalangeal joint, especially with dorsiflexion of the great toe.

X-rays such as dorsoplantar, lateral, oblique, and axial sesamoid views are typically ordered. Contralateral views should be ordered to rule out a bipartite or tripartite sesamoid. Sesamoid fractures are usually seen on x-ray as comminuted or transverse
orthotics with a cutout under the first metatarsophalangeal joint is helpful in reducing the pressure under the sesamoid bones. This is particularly important in patients who have a plantar-flexed first metatarsal or cavus foot.

Sesamoid fractures are prone to delayed unions and nonunions (40). Fractures usually do not heal with bony union but rather by a fibrous bridge. Many nonunion fractures are asymptomatic (16). Those sesamoid fractures that are painful after 6 weeks may need an injection of steroid to reduce the inflammation. If pain persists and a loss of function is present, surgical excision is recommended. The clinician must be aware that if the tibial sesamoid is removed, a bunion can develop. If the fibular sesamoid is removed, a hallux varus can develop. If both sesamoids are removed, a hallux malleus can occur.

**Fifth Metatarsal Base Fractures**

Two types of fractures are described at the base of the fifth metatarsal: avulsion fracture and Jones’ fracture (Fig. 38-7). The base of the fifth metatarsal has a styloid process where the peroneus brevis tendon inserts. The amount of displacement and location of the fracture dictate its course of treatment.

(Fig. 38-6). These fractures may be jagged and irregular with variable degrees of displacement. Congenital bipartite sesamoids appear smooth on x-ray. If there is no change on x-ray and the clinician is still suspicious of a fracture, a bone scan can be performed (39). MRI is sensitive in ruling out a fracture or other disease states of the sesamoid bone.

Treatment is directed at reducing the level of pain to the area. Rest, ice, anti-inflammatory medication, and a surgical shoe with a flat rigid sole worn for 4 to 6 weeks that prohibits push-off are initiated as conservative treatment for this condition. Once the patient is back in shoes or sneakers, custom foot orthotics.
When there is a sudden pull on the peroneus brevis tendon, such as in a lateral inversion ankle sprain, an avulsion fracture can occur. This is a common injury in athletes when their foot lands in a supinated position. Avulsion fractures can also occur with excessive weight bearing on the lateral aspect of the foot. Patients with a supinated gait or cavus foot may be predisposed to this type of injury.

Treatment of nondisplaced avulsion fractures is usually conservative. The patient is fitted with a Cam walker or surgical shoe for 6 to 8 weeks. The patient is permitted weight bearing to tolerance. If there is significant displacement to the fracture, open reduction and internal fixation should be performed. The most popular forms of fixation are tension band and intermedullary screw fixation.

Jones’ fracture is a transverse fracture in the diaphyseal region of bone 1.5 to 3 cm distal to the fifth metatarsal tuberosity. This type of fracture occurs from a vertical ground reactive force with no inversion of the foot (41). The patient pivot-shifts the foot with the ball of the foot on the ground and the rearfoot off the ground. This type of injury is less stable than the avulsion fracture because of the long lever arm of the metatarsal. The metatarsal places stress on the fracture site, making this type of fracture unstable. Healing is more difficult because of the instability and poor vascularity at the fracture site.

Treatment of choice in fractures that are moderately to severely displaced is usually open reduction and internal fixation. This is especially true for athletes and active patients. The repair is amenable to screw or plate fixation. Bone grafting is necessary, along with internal fixation, for those fractures that went on to delayed union or nonunion.

**Lesser Metatarsal Fractures**

**Etiology**

Metatarsal fractures are usually related to direct or indirect trauma (Fig. 38-8). The second, third, and fourth metatarsals are the most frequently fractured (16). Indirect trauma is a stress fracture caused by indirect repetitive trauma. Direct trauma is a result of a sudden impact or crush injury.

**Clinical Findings and Treatment**

Clinical signs of metatarsal fractures include edema, ecchymosis, and pain over the area. It is important to evaluate the digital vascular status to ensure that there is no vascular compromise. Standard radiographs, including an anteroposterior, medial oblique, and lateral view, should be taken to assess the injury. If sagittal plane alignment is questioned, then an axial metatarsal view should also be performed. This enables one to evaluate whether the metatarsal is plantar-flexed or dorsiflexed.

Fractures to the lesser metatarsals may occur at the neck, body, or base. Metatarsal neck fractures are common because the injuring force is perpendicular to the long axis of the bone. The fracture pattern is usually oblique. If the fracture is not displaced, a short-leg non-weight-bearing cast may be used for 6 to 8 weeks. If there is displacement, then closed reduction should be attempted. Failure to realign the bone can cause shortening to the metatarsal, floating toe syndrome, or adjacent calluses under a metatarsal head (16). If closed reduction fails, the patient needs to undergo open reduction with internal fixation. Usually, intramedullary pinning with Kirschner wires (K wires) is performed.

Midshaft fractures to the lesser metatarsals have a variety of fracture patterns. These fractures are in diaphyseal bone and may be oblique, comminuted, spiral, or transverse. Nondisplaced fractures are treated with a removable walking cast such as a Cam walker for 4 to 6 weeks. If the fracture is displaced and there is significant shortening to the metatarsal, then open reduction with internal fixation may be necessary. Comminuted fractures are usually treated successfully with K wires, cerclage wiring, plates, and screws.

Metatarsal base fractures are usually transverse and are approximately 1 cm from the articular surface. An intraarticular fracture at the base is suspicious for a Lisfranc’s joint injury. If there is no joint involvement, then cast immobilization is adequate.

**Phalangeal Fractures**

**Etiology**

Fractures of the digits are a result of sagittal, transverse, or frontal plane forces. The most common plane of injury is in
CHAPTER 38 | FOOT DISORDERS


the sagittal plane. Sagittal plane injuries result in compaction of the toe (object dropped on toe) or hyperextension or hyperflexion to the joints of the digit (stubbed toe). A transverse plane deformity results in an adduction-abduction injury causing a spiral oblique fracture to the proximal, middle, or distal phalanx. This is the typical “bedroom fracture,” whereby the fifth toe strikes the bedpost while walking in the dark (16) (Fig. 38-9). The frontal plane deformity is the least common injury and occurs with inversion-eversion injuries of the digit. There is a high degree of torque so that the digit is prone to neurovascular injuries along with the potential for an open wound and soft-tissue damage.

Clinical Symptoms and Treatment
The symptoms are generally acute pain with ecchymosis and swelling to the toe. The patient has pain while wearing closed shoes. With crush injuries, there may be nail involvement causing loosening of the nail or a hematoma under the nail plate.

Radiographs are taken of any digital injury to rule out a fracture. Medial oblique, lateral, and anteroposterior views are taken. If there is displacement, then one should attempt closed reduction under local anesthesia. Closed reduction is performed by exaggerating and then reversing the injury pattern (53). Once there is satisfactory reduction, the digit is maintained in the correct alignment with tape. If closed reduction does not adequately align the toe, then open reduction will be necessary.

Immobilization is used for stable and undisplaced fractures. It is preferable to buddy-tape the injured and adjacent digit together. The patient wears a stiff-soled surgical shoe for about 4 weeks. Ice and anti-inflammatory medication usually control symptoms.

Orthotics
Functional orthotic devices control poor foot biomechanics. They are also beneficial in maintaining alignment after surgical intervention. A cast is made in subtalar joint neutral position while the patient is in the prone position. Two 5- by 30-in. fast-setting plaster-of-Paris splints are used for each foot. Orthotics are generally made for both feet, even if the patient has symptoms on only one.

There are a variety of lengths, materials, and posts for orthotics. The decision based on the patient’s age, activity, size, shoe style, and foot type (pronated versus supinated).

In the active patient, full-length running orthotics are designed to control the patient in heel strike, midstance, and the push-off phase of gait. If the orthotic ended just behind the metatarsal heads, there would be no control for the athlete going up on the ball of the foot. Orthotics that end behind the metatarsal heads are designed for patients who do not need control at push-off. This might be seen in a patient who needs an orthotic for daily walking.

The most common material used is the polypropylene shell orthotic. This is a firm device that has good support. Harder plastic devices are less well tolerated. Polypropylene orthotics are generally 3/16-in. thick. On lighter patients, or when less control is needed, the device can be made 1/8-in. thick. This gives more flexibility to the shell and fits into more shoes. Graphite is another type of material used for orthotics. This material is preferable for dress shoes because they are light and thin.

Posting to the rearfoot or forefoot is used to help balance and control forefoot and rearfoot biomechanical abnormalities such as forefoot varus, forefoot valgus, rearfoot varus, or rearfoot valgus. Posting may be intrinsic (within the shell of the orthotic) or extrinsic (outside the shell of the orthotic). When the forefoot is intrinsically posted, the distal edge of the shell is heat molded and curved to accommodate a forefoot varus or valgus deformity. Extrinsic posts are made out of acrylic or crepe and are placed under the distal medial aspect of the orthotic.

A rearfoot extrinsic post is made out of crepe or acrylic and is directly applied to the heel seat of the device. Posts are generally made to place the heel vertical to the ground or with a 4-degree varus post if the patient displays a rearfoot valgus deformity. A rearfoot intrinsic post is made within the shell of the heel. This type of post takes up less room in the shoe and is used for dress shoe orthotics.

Special accommodations can be applied to orthotics. If a patient has a repeated history of ankle sprains, a high lateral flange can be made on the lateral aspect of the orthotic. This aids in proprioception and keeps the ankle from inverting.
Decreased ROM to the first metatarsophalangeal joint is accommodated with a Morton’s extension. A cutout to the shell of the first metatarsophalangeal joint can help in patients with sesamoid pain. Extra padding in the heel may help patients with plantar fascial pain.

Diabetic patients may need an orthotic to help spread their weight more evenly. If a diabetic patient has a callus under a specific metatarsal head, a device should be made with a cutout to the area to reduce stress to that part of the foot. Diabetic orthotics are generally made of Aliplast and Plastazote.

**LISFRANC’S JOINT INJURIES**

The diagnosis and treatment of tarsometatarsal joint (Lisfranc’s joint) injuries can be quite challenging. Appropriate knowledge of the parameters of the injury, anatomy, and consequences is paramount to successful management (42,43). Injuries of the tarsometatarsal joint are usually a result of a high-energy force. This can be broken into direct or indirect trauma. Direct injuries are usually a result of a crushing force. Indirect injuries, the most common mechanism (44,45), result from a combination of twisting and axial forces with the metatarsals plantar-flexed on the ground and the foot in a slight plantar-flexed position (46). The automobile injury is typical of this type of mechanism, with the foot being jammed against the floorboard. Quenu and Küss (47) first described this injury in 1909. They classified the injury based on metatarsal displacement into three groups (Table 38-4). In 1982, Hardcastle et al. (48) modified the Quenu and Kuss classification to include the extent and type of incongruity based on radiographs (Table 38-5).

**Clinical Findings and Treatment**

The main emphasis in diagnosing these injuries should be placed on the clinical and radiographic examination along with high index of suspicion in injuries that are more occult or that may have been missed at the time of initial presentation. Physical examination reveals swelling and pain of the forefoot. An adduction or abduction deformity of the forefoot may be present. Ecchymosis at the plantar midfoot, termed the *plantar ecchymosis sign*, may provide insight into the presence of this injury (49). Severe swelling and hematoma formation can result in a compartment syndrome, and it is of the utmost importance that the examiner assesses the patient for this condition.

X-rays are necessary for diagnosis and should be weight bearing with multiple views. Anatomically, the medial border of the second metatarsal base should align with medial border of the middle cuneiform, the space between the first and second metatarsal bases should be less than 2 mm on an anteroposterior radiograph. An oblique view should reveal alignment of the lateral border of the third metatarsal with the lateral border of the lateral cuneiform, and the medial border of the fourth metatarsal base should align with the medial border of the cuboid. Attention should be directed to the cuboid, which may also fracture and collapse on itself, resulting in a significant forefoot abduction deformity. Injuries to the metatarsophalangeal joints, interphalangeal joints, and floating metatarsals have been reported with Lisfranc’s injuries (50,51). Stress views may also be required, especially in the more subtle injuries for which standard x-rays are negative but there is a high index of suspicion because of increased pain at the tarsometatarsal complex.

Pain can be elicited, with stress applied to the tarsometatarsal complex by applying an abduction-pronation stress.
and an adduction-supination stress during examination. The examiner should be cautious in the acutely injured patient because this could be intensely painful. Stress radiographic examination should be performed with a local nerve block to minimize discomfort to the patient and to allow an adequate amount of stress to be applied to the joint complex. CT scans and MRI studies may also be considered, especially in instances in which there is significant osseous involvement and ligamentous injury, respectively (Fig. 38-10).

Treatment must be directed at providing ligamentous and osseous stability. This usually comes in the form of surgical stabilization. It is paramount to understand that most of these injuries require surgical intervention with fixation for 3 to 4 months followed by intense rehabilitation. Fracture reduction is essential (Fig. 38-11). K-wire fixation, previously the standard procedure, has been found to be suboptimal because of loosening and fixation failure (52). Their use should be restricted to lateral column fixation, in which some mobility of the joint complex is advantageous, or as an adjunct to other types of fixation or in which the soft-tissue envelope has been severely compromised and temporary fixation is needed until soft-tissue healing has occurred and definitive stabilization performed (53). In situations in which the patient continues to have increased pain, inability to bear weight, and swelling, improper reduction and stabilization may be the cause (53).

Traditionally, postoperative patients have been maintained with no weight bearing for a period of 8 to 12 weeks in a cast. The emphasis is toward early ROM exercises in an attempt to reduce swelling and tissue fibrosis (54). Placement in a commercially removable cast after 3 weeks can provide for early ROM exercises. Progressive partial to full weight bearing is instituted. Fixation is usually removed 3 to 4 months before any unprotected weight bearing. In fact, complete ligamentous healing is believed to take 12 months and proceeds in four phases (52,55). Fabrication of orthotic devices and modification of shoe gear depending on patient activity may be necessary.

As a result of a broad spectrum of injuries associated with midfoot sprains that range from partial sprains to complete dislocation and the controversy in treatment, Nunley and Vertullo (56) have attempted to provide a classification system to guide investigation and management. The classification scheme is based on clinical findings, weight-bearing radiographs, and bone scintigrams. It is divided into three stages (Table 38-6). Frank displacement of Lisfranc’s complex was categorized by Myerson’s modification of Lisfranc’s injuries (45). Stage I injuries can be treated nonsurgically with cast immobilization for 6 weeks. Stage II and III injuries require surgical intervention with internal fixation following the same standards for high-velocity Lisfranc’s injuries. If the patient is an athlete, sport-specific rehabilitation is necessary for overall recovery. A recent review outlines the diagnostic approach and treatment options and emphasizes the importance of urgent care to maximize a good outcome (57).
FOOT PROBLEMS ASSOCIATED WITH DIABETES

Diabetes mellitus affects about 16 million individuals in the United States (58,59). Diabetic patients account for 50% or more of the nontraumatic amputations (60–63). This high number can be explained by the fact that the number of persons with diabetes is increasing, the diabetic population is aging, and more effective coding by diagnosis enables more accurate reporting (60). Advancing age, male gender, and nonwhite racial status increase the risk for amputation (64). Pecoraro et al. (62) reported that amputations are a result of several potential pathways. These include neuropathy, minor trauma, cutaneous ulceration, infection, ischemia, altered wound healing, and gangrene. The classic triad of peripheral arterial disease (PAD), peripheral neuropathy, and infection is the harbinger of the final pathologic event of gangrene and resultant amputation (65). It has been shown that the level of amputation in diabetic patients is more distal than that of nondiabetic patients (64). Poor glycemic control, presence of retinopathy, presence of proteinuria, diabetes education, and smoking contribute to the need for amputation (66,64). The incidence of subsequent amputation either of the ipsilateral or contralateral limb rises to 50% or more at 5 years (64,67,68). The need to preserve a limb becomes even more important owing to the increased oxygen demand and the work of walking that is required to maintain functional ambulation (69–71). The economic impact is high. The cost of diabetic ulcer care exceeds $5 billion per year. Prevention and education are necessary if these costs are to be contained.

Deformities occurring in the diabetic patient can range from hammertoes and bunions to the devastating collapse of the foot and ankle as a result of Charcot’s neuroarthropathy. This eventually leads to ulceration and a portal of entry for bacterial infection. The pathogenesis is multifaceted; however, the presence of sensory and motor neuropathy plays a significant role. Recent emphasis has been placed on collagen glycosylation resulting in decreased joint and skin mobility (72,73). Deformities associated with neuropathy and decreased joint mobility contribute to a gait pattern that produces high plantar pressure distribution resulting in ulceration. Cavanagh et al. (74) determined that foot deformities in diabetic patients with neuropathy account for increased plantar pressures and ulceration. Robertson et al. (75) have shown that there are structural differences between the forefeet of individuals with diabetes, neuropathy, and a previous ulcer and controls. They found decreased muscle density, increased metatarsophalangeal extension, and increased arthropathy. There was no difference in soft-tissue density between the groups. It is apparent that multiple risk factors can result in a cascade ultimately culminating in loss of limb (Table 38-7).

<table>
<thead>
<tr>
<th>TABLE 38.6 Midfoot Sprain Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage I</td>
</tr>
<tr>
<td>Stage II</td>
</tr>
<tr>
<td>Stage III</td>
</tr>
</tbody>
</table>


<table>
<thead>
<tr>
<th>TABLE 38.7 Risk Factors for Foot Ulceration in Diabetics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neuropathy</td>
</tr>
<tr>
<td>Sensorimotor (abnormal protective sensation)</td>
</tr>
<tr>
<td>Autonomic (dry, cracked skin)</td>
</tr>
<tr>
<td>Vascular disease</td>
</tr>
<tr>
<td>Abnormal plantar pressure (elevated in neuropathy even in the absence of deformity)</td>
</tr>
<tr>
<td>Abnormal gait in elderly living alone</td>
</tr>
<tr>
<td>Degenerative joint diseases of hip and knee</td>
</tr>
<tr>
<td>Muscle weakness</td>
</tr>
<tr>
<td>Heal cord tightness</td>
</tr>
<tr>
<td>Pronation, supination deformities of the foot</td>
</tr>
<tr>
<td>Toe contractures</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
Examination of the diabetic foot should include the same parameters for evaluation of the nondiabetic foot: patient history and neurologic, vascular, and musculoskeletal evaluation. Emphasis should be placed on determining the presence of neuropathy, vascular status, and the presence of biomechanical derangement. The presence of subjective complaints of numbness, tingling, or burning may be absent and may underestimate the presence of significant loss of sensation. Additionally, the loss of sensation may contribute to the lack of signs and symptoms that could be attributable to peripheral vascular disease. Neurologic examination should include evaluation of deep tendon reflexes, sensory testing, and motor examination. Sensory evaluation can be accomplished using the relatively inexpensive 5.07 Semmes-Weinstein monofilament. It is pressed against the skin until it buckles (Fig. 38-12). This corresponds to 10 g of linear pressure. If the patients cannot feel the monofilament at the pressure needed to cause it to buckle, they have lost protective sensation and could possibly be at risk for ulceration (76). This provides much more quantifiable information than using cotton swabs or pin prick. Loss of vibratory sensation using a 128-Hz tuning fork at the first and fifth metatarsal phalangeal joints carries the same significance as inability to feel the 5.07 monofilament (77).

Decrease or loss of pulses or signs and symptoms of ischemia, intermittent claudication, nocturnal and rest pain, shiny and cool skin, loss of hair, pallor with elevation, and rubor with dependency should be referred for vascular laboratory evaluation including segmental and digital pressures. The composition of atherosclerotic plaques in diabetic and nondiabetic patients is the same (60,65). However, PAD in diabetic patients occurs at an earlier age and progresses much more rapidly. Male and female patients are affected equally. The vessels involved in diabetic patients tend to include the tibial and peroneal arteries (66,78). Interestingly, microangiopathy, occlusive disease of the end arterial tree, does not appear to play a significant role in foot ischemia, although a thickened capillary basement membrane is present (66,78). McNeely et al. (76) have found the absence of an Achilles reflex, loss of protective sensation, and transcutaneous oxygen tension less than 30 mm Hg as independent and significant predictors of foot ulceration (76). The musculoskeletal examination should include evaluation for the presence of deformities located at the hip, knee, leg, and foot; decreased ROM; limb length discrepancies; presence of prior amputations; and gait evaluation. Abnormalities that affect the foot or superstructure, such as genu varum or valgum, hammer toes with prominent metatarsal heads, bunion formation, limited joint mobility (especially loss of dorsiflexion owing to contracture of the Achilles tendon), equinus deformities, and prior amputations can place abnormal pressure on the foot resulting in callus production and tissue breakdown. Deformities should be accommodated in adequately fitting shoes and may require prescription for extra depth or custom-molded shoes.

Surgical intervention may be needed in select individuals to rebalance the foot to minimize pressure, especially at prominent digital and metatarsal areas and in the presence of plantar ulcers due to excessive plantar pressures from equinus deformities as a result of tendoachilles contracture. It has been shown that long-term hyperglycemia results in marked increase in tendon-fibril density secondary to nonenzymatic cross-linking of fibrillar collagen (79). Gait evaluation and inspection for abnormal shoe wear should be performed. Gait abnormalities could result in unsteady gait leading to falling and abnormal pressure on the foot and causing tissue breakdown and injury (80). Defective proprioception due to neuropathy could result in a wide gait and constant watching of the floor and landmarks resulting in frequent falls. Evaluation at each office visit should assess any significant changes and appropriate action taken to minimize potential future problems.

Charcot's neuroarthropathy results in bone destruction collapse and fracture. It is commonly mistaken for an infectious process. Diabetes is the leading cause. The incidence of unilateral involvement has been reported to be between 0.08% and 7.7%, with bilateral involvement between 5.9% and 39.3% (81). Involvement of the contralateral limb due to increased stress should remain a high concern for the clinician. Patients on the average have had diabetes for 10 to 15 years and commonly have poor diabetes control. The development of Charcot’s arthropathy has been debated as a result of either a neurotraumatic origin or a neurovascular origin. It appears that a combination of factors from both theories is most likely the cause of this condition. Charcot develops as a result of a neurogenic neuropathy that causes an increase of blood flow to the extremities resulting in a profound osteopenia. Motor neuropathies due to imbalances create abnormal pressures. The presence of sensory loss makes the patient unaware of these pressures, and ultimate osseous destruction ensues (82,83). Armstrong and Lavery (82) have shown increased peak plantar pressures in

![Figure 38-12](image-url). Use of Semmes-Weinstein monofilament to determine the presence of protective sensation in the foot.
individuals with Charcot’s arthropathy and in those who have had a neuropathic ulcer. The peak pressures were noted to be in the forefoot even in individuals with midfoot collapse, where you would expect a concentration of forces. They theorized that because they identified patients with acute Charcot’s early, they eliminated the profound destruction seen at the midfoot. They further theorized that high peak pressures and neuropathy in the forefoot, coupled with an equinus deformity in the hindfoot before arthropathy develops, may predispose these individuals to the catastrophic cascade of events resulting in deformity. Armstrong et al. (84) found that performing a percutaneous lengthening of the Achilles tendon in diabetic patients at risk for ulceration due to increased peak plantar pressures resulted in decrease of these pressures. Misdiagnosis of this condition can result in significant deformity and disability.

Patients present with a red, hot, swollen foot and commonly relate the occurrence of a deformity, such as fallen arches or walking on the side of the ankles (Fig. 38-13). The clinician should always have a high index of suspicion of Charcot’s arthropathy in the presence of increased skin temperature without evidence of open wounds or lymphangitis. A history of trauma may not be present. The patient may relate a strange sensation or a vague pain that is not normally present. Armstrong et al. (85) reported that 76% of patients who presented with neuroarthropathy initially presented with pain.

Radiographs initially may be normal or show osteopenosis, fracture, dislocation, or bony fragmentation. An extensive deformity may be present without fracture. This is a result of chronic ligamentous failure resulting in loss of support because of weakening of ligaments with eventual acute deformity. This could be precipitated by minor trauma. If there is presence of ulceration, differentiation from an infection and osteomyelitis can prove difficult, especially if the ulcer probes to bone. It has been our experience that these patients may also have an increased erythrocyte sedimentation rate and white blood cell (WBC) count. This may require further diagnostic testing, including MRI, CT, and bone scans. The gold standard to differentiate Charcot’s arthropathy from osteomyelitis is the bone biopsy, which may be required in cases in which other diagnostic evaluations have been equivocal (81). Biopsy results of a Charcot’s deformity reveal synovium embedded with bone and cartilage fragments, as opposed to osteomyelitis, in which bone destruction with leukocyte infiltration would be seen.

Eichenholz has divided the radiographic destruction into three distinct phases: development, coalescence, and reconstructive (86).

**Stage of development:** acute destructive period characterized by joint effusion, soft-tissue edema, subluxation, fragmentation of bone, intraarticular fractures, and formation of bone and cartilage debris. Continued ambulation during this stage results in significant deformity (Fig. 38-14).

**Stage of coalescence:** period of healing with decreasing edema, absorption of debris, and healing fractures (Fig. 38-15).

**Stage of reconstruction:** further repair and remodeling of bone with increased bone density and sclerosis with improved joint stability (Fig. 38-16).

Shibata et al. (87) and Sella and Barrette (88) described a stage 0, during which there are no radiographic changes but the foot is warm with swelling. Schon and Marks (89) described a stage 0 that represents an at-risk patient who is neuropathic and has sustained an acute injury. Radiographs show a simple or comminuted fracture, possibly with joint space narrowing but no fragmentation, which is associated with stage 1.
Sanders and Frykberg developed an anatomic pattern classification system based on distribution of bone and joint involvement (90). This recognizes five types of anatomic deformity as the end result of the Charcot’s arthropathy process.

Pattern I: forefoot involvement of the interphalangeal joints, phalanges, metatarsophalangeal joints, or distal metatarsal bones; radiographic findings are typically atrophic and destructive

Pattern II: Tarsometatarsal joints (Lisfranc’s joint); ulceration seen at apex of collapsed cuneiforms or cuboid

Pattern III: naviculocuneiform, talonavicular, and calcaneocuboid characterized by dislocation of the navicular or disintegration of the naviculocuneiform joint

Pattern IV: ankle and subtalar joints with severe structural and functional instability

Pattern V: calcaneal fracture; typically avulsion of the posterior tubercle

Patterns I and II are associated with bony deformity and ulceration. The most frequent patterns are I, II, and III. The most severe deformity and instability are seen in patterns II and IV. Pattern V is the least common (81). Pattern IV is the most difficult type to treat with bracing owing to severe subluxation and ulceration (Fig. 38-17). Most commonly, they require surgical stabilization in the form of fusion of the tibial-talar or tibial-calcaneal segments (Fig. 38-18).
Intervention to prevent bony breakdown and deformity requires the clinician to be astute to the harbinger of the clinical presentation of Charcot’s arthropathy as mentioned previously. Strict non–weight-bearing cast immobilization is the gold standard of treatment. Prompt initiation of non–weight bearing may stop the destructive progression of this disease process, resulting in deformity. Immobilization could be obtained through total-contact casts (TCCs), standard below-knee casts, or prefabricated removable cast braces. Consideration should be given to the contralateral limb during this time of immobilization, which could be placed at increased risk for neuroarthropathy developing 4.5 months subsequently (81). This may require the use of prefabricated walking braces or orthosis to the contralateral limb. Wheelchair confinement may be necessary in situations in which both limbs are at risk (81). The non–weight-bearing status and cast immobilization should continue until all signs of inflammation resolve and serial radiographs reveal consolidation of bone. Erythema and edema will resolve first followed by warmth. Continued warmth indicates persistent inflammation that mandates continued immobilization. This process can take 2 to 6 months. Sanders and Frykberg recommend a minimum of 3 months of non–weight-bearing cast immobilization (90). After this inflammation has resolved, the Charcot’s foot has entered the stage of coalescence when protected weight bearing may begin. Patellar tendon-bearing (PTB) braces, prefabricated walking braces, custom braces, and a Charcot’s restraint orthotic walker (CROW) can be used. The CROW is a combination of a PTB and custom walking brace that enhances edema reduction, immobilization, and stability (Fig. 38-19). Progression to the final stage of treatment is predicated on the absence of any continued signs of inflammation, most importantly normalization of skin temperature and the presence of increased bone density and consolidation. The patient should maintain full weight bearing with the PTB brace, prefabricated walking brace, custom brace, or CROW for several weeks. Subsequently, a bracing prescription will depend on the deformity and may be an ankle-foot orthosis, double upright brace, or prescription footwear. Emphasis should be placed on proper footgear regardless of the magnitude of the deformity. Minimally, the patient should be placed in an extra-depth shoe with a Plastazote insole. If there is residual deformity or history of ulceration due to bony prominences, custom-molded shoes with extended steel shank and rocker bottom may be required. It should be noted that if shoe modifications are not placed properly, continued ambulation could result in development or progression of ulceration.

Surgery may be indicated for deformities associated with Charcot’s arthropathy resulting in significant instability that are not controlled by bracing or shoe gear, recurrent ulcerations, and selected acute fractures (91). Interestingly, in a review of 55 patients treated with conservative care, Armstrong et al. (85) found that 25% eventually required surgical correction in the form of either exostectomies or fusion. The goal of surgery is to create a plantigrade foot that is stable, realigned, and able to accommodate shoe gear and bracing in order to prevent ulceration and limb loss. Surgical intervention may include tendoachilles lengthening, exostectomies, and realignment arthrodesis.

**DIABETIC FOOT ULCERS**

Ulceration is usually the consequence of neuropathy and is compounded when deformity exists as a result of neuroarthropathy. In fact, more than 50% of patients with Charcot’s deformity in a study by Pinzur et al. (92) presented with an ulcer. It is therefore important to be able to record systematically the characteristics of ulcers in order to plan treatment, monitor progress, and communicate effectively with health care providers (93). The most widely accepted diabetic ulcer classification is that developed by Meggitt (94) and Wagner (95) and is referred to as the Wagner classification. The classification is based on the depth of penetration of the wound and extent of necrosis and is divided into six wound grades (Table 38-8). The Wagner classification does not adequately address the presence of ischemia or infection. Attempts to refine the classification have resulted in several newer

### TABLE 38.8 Ulcer Classification

<table>
<thead>
<tr>
<th>Grade</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Intact skin may have bony deformity or limited skin mobility</td>
</tr>
<tr>
<td>1</td>
<td>Localized superficial ulcer of the skin</td>
</tr>
<tr>
<td>2</td>
<td>Deep ulcer extending to tendon, bone, ligament, or joint</td>
</tr>
<tr>
<td>3</td>
<td>Deep ulcer with abscess or osteomyelitis</td>
</tr>
<tr>
<td>4</td>
<td>Forefoot gangrene</td>
</tr>
<tr>
<td>5</td>
<td>Gangrene of the entire foot</td>
</tr>
</tbody>
</table>

subtypes of the original classification, specifically the Brodsky Depth-Ischemia classification (96) and the University of Texas San Antonio Diabetic Wound Classification System (93). The latter addresses the depth of the ulcer, presence of infection, and presence of clinical signs of ischemia. It uses four grades of depth (0 to 3) and four stages (A to D) based on the presence of ischemia or infection separately or in combination. The complexity of these expanded classifications makes them more difficult to implement in a clinical setting. Regardless of the classification scheme used, it is important to record the circumference, depth, and size of the wound. Description of the base of the ulcer should include the presence of granulation tissue and its consistency and color. If necrotic tissue is present, it should be described in terms of either being a thick eschar or liquefied slough. The ulcer is probed looking for sinus tracts, abscess formation, and contact with bone. It has been shown that the ability to probe to bone is highly predictive of osteomyelitis (97). Drainage is described as to its amount and odor and whether it is serous, serosanguineous, or purulent. Is the soft-tissue fluctuant with crepitus indicative of the possibility of gas in the soft tissues? Radiographs should be performed routinely to rule out gas formation, osteomyelitis, foreign bodies, and fractures. Frank bony destruction with progressive periosteal resorption and osteolysis may indicate osteomyelitis in the proper clinical circumstances. However, osteomyelitis can appear radiographically identical to diabetic neuroarthropathy (98,99). The presence of intact skin may indicate a Charcot’s deformity, as described previously. If suspicion remains high for an infectious process, radionuclide imaging may be required. Technetium-99m is the most commonly ordered scan and is usually performed as a three-phase scan. Technetium localizes in areas of increased bone metabolism and vascularity, which is present in cases of osteomyelitis and neuroarthropathy. A fourth-phase scan 24 hours after injection has been used to increase specificity for osteomyelitis (98–100). Gallium-67 has been used as a scanning agent predominately because of its localization to transferrin and lactoferrin binding. Gallium also localizes to areas of increased bone activity, thereby lacking the same specificity as technetium in regard to determining an infectious process (101). WBC scans using technetium-99 HMPAO or indium-111 rely on accumulation of white cells within areas of leukocyte accumulation. Although indium is considered the most specific agent for inflammation, it is a labor-intensive procedure adding cost, and it requires a higher radiation dose. The benefit of indium is that it can be combined with a technetium bone scan to increase its specificity for detection of osteomyelitis because the energies of the two isotopes are different (98–100). The use of technetium sulfur colloid scan, which is taken up by the reticuloendothelial system, including bone marrow, has been advocated either singularly or in combination with a WBC scan. MRI appears to be gaining popularity as the modality of choice in the diagnosis of osteomyelitis; however, a rapidly progressive acute neuroarthropathy can exhibit bone marrow edema similar to osteomyelitis (98). A logical stepwise approach to identification of osteomyelitis in clinically suspect patients has been to order a three-phase bone scan if x-rays are negative (99,101). If the bone scan returns abnormal, proceed with a WBC scan in combination with MRI and possibly the addition of a sulfur colloid study. Proceeding initially with a WBC technetium-99 HMPAO scan may be more beneficial in the presence of osseous abnormalities on x-ray; this could be combined with a technetium sulfur colloid scan or MRI, especially for preoperative planning.

Aggressive sharp debridement of neuropathic ulcers is necessary to remove all necrotic tissue and to determine the true nature of the ulcer. What may appear to be superficial may just be the “tip of the iceberg” (65) (Fig. 38-20). The ulcer should be debrided down to healthy bleeding tissue. Clinically suspicious infected wounds should be cultured. Superficial swab cultures are unreliable owing to colonization from surface organisms. If necessary, cultures are obtained from the deeper parts of the wound after debridement (102,103). Non–limb-threatening infections may be treated on an outpatient basis empirically with cephalaxin, clindamycin, amoxicillin–clavulanic acid, or levofloxacin until culture results are available and changed if necessary (101). The patient should be monitored, and if the condition deteriorates, admitted for intravenous antibiotics and possible incision and drainage procedures. Limb-threatening infections require hospitalization, intravenous antibiotic therapy, and aggressive incision and drainage. Antibiotic selection of piperacillin-tazobactam, ticarcillin–clavulanic acid, and ampicillin–sulbactam are all excellent empiric choices (101). The decision to admit the patient because of severe infection should not be based on systemic signs alone. Diabetic patients are typically immunosuppressed with defective leukocytosis and impaired intracellular killing capacity. Typically, the diabetic patient will not mount a significant leukocytosis. Gibbons and Eliopoulos (104) have shown that two thirds of diabetic patients with severe infections do not have increased temperature, chills, or leukocytosis. Uncontrolled hyperglycemia may be the only evidence of systemic disarray, which may require hospitalization for metabolic control. The presence of a deep

**FIGURE 38-20.** Superficial-appearing ulcer with sinus tract extending to dorsal aspect of foot.
abscess, especially with extension from a plantar ulcer to the dorsum resulting in cellulitis, cellulitis extending 2 cm beyond the ulcer site, and ischemia are all signs of a limb-threatening situations requiring hospitalization (102,105).

Rest, elevation, and control of edema should all be included in any treatment regimen. Removal of improperly fitting shoes and placement into a surgical shoe or other pressure-reducing footwear are required (105). Persistence by the health care community to use whirlpool treatments or advocate foot soaks as a treatment to debride tissue mechanically is detrimental and should not be part of any treatment modality. These treatments only provide for limited superficial debridement. Their prolonged use can result in maceration, progression of infection, and increased edema. Their use only delays sharp debridement and wound healing (60,65,106,107). The mainstay of treating plantar ulcers has been with the use of TCCs (Fig. 38-21). It is believed to reduce plantar pressures at the ulcer site (108). Studies have found that ulcers healed on average 4 to 6 weeks after initiation of the TCC (109–112). It is labor intensive, requires weekly applications, and if improperly applied can result in significant complications, including increased joint stiffness, impaired mobility, abrasions, ulceration, osteoporosis, and muscle atrophy (113,114).

Unfortunately, ulcer recurrence rates can range from 20% to greater than 50% (114–117). Mueller et al. (117) performed a randomized study comparing TCC alone in 64 patients and in combination with percutaneous Achilles tendon lengthening in 31 patients. Eighty-eight percent of the TCC group and 100% of the Achilles tendon—lengthening group healed. In the first 7-months’ follow-up of those patients available, 59% of the TCC group and 15% of the Achilles tendon—lengthening group had a recurrence. At the time of a 2-year follow-up, 81% of the TCC group and 31% of the Achilles tendon—lengthening group had a recurrence. This study emphasizes the importance of addressing structural imbalances associated with ulcer formation, especially the destructive nature of an equinus deformity, which has been discussed previously. If TCC is to be used, experienced individuals should apply it and the patient monitored very closely. The patient should be counseled on the modality and the importance of compliance with treatment protocols. Effective alternatives to TCC include prefabricated total-contact posterior ankle-foot orthoses, removable prefabricated walking casts, wedged or half-soled shoes, and total-contact sandals (105,109,118) (Fig. 38-22). Prolonged topical use of povidone-iodine, Dakin’s solution, acetic acid, and hydrogen peroxide should be discouraged because they have been shown to disrupt the healing process. A moist environment is most conducive to wound healing. Numerous products are available, including alginites, semipermeable dressings, and collagen (119,120). Growth factors, living tissue equivalents, vacuum-assisted closure, skin grafts and flaps, electrical stimulation, hyperbaric oxygen, and ultrasound are among the available adjunctive treatment aids for wound closures and should be evaluated with regard to the individual patient’s need. Wounds that have been present for prolonged periods of time with no sign of healing or improvement should be considered for biopsy. Malignancy as a primary lesion or conversion to a malignancy as seen at areas of chronic drainage from sinus tracts resulting in squamous cell carcinoma (Marjolin’s ulcer) can occur and should be part of the differential in hard-to-heal wounds (Fig. 38-23). Additionally, wounds that are excised for

![Figure 38-21. Total-contact cast.](image1)

![Figure 38-22. Total-contact sandal.](image2)
FIGURE 38.23. Squamous cell carcinoma in situ at a site of a chronic nonhealing heel ulcer.

TABLE 38.9 Diabetic Wound Care Treatment Guidelines

<table>
<thead>
<tr>
<th>Evaluation</th>
<th>Location</th>
<th>Size</th>
<th>Depth</th>
<th>Presence of sinus tracts</th>
<th>Condition of tissue in wound bed</th>
<th>Exudate</th>
<th>Presence of odor</th>
<th>Amount</th>
<th>Type</th>
<th>Signs of infection</th>
<th>X-rays</th>
<th>Gas in tissue</th>
<th>Bony changes</th>
<th>Ancillary studies as needed</th>
<th>Vascular status</th>
<th>Need for vascular reconstruction</th>
<th>Treatment</th>
<th>Debridement</th>
<th>Cultures</th>
<th>Empiric antibiotic coverage</th>
<th>Oral</th>
<th>Parenteral</th>
<th>Rest, elevation control of edema</th>
<th>Off-loading of ulcer</th>
<th>Biopsy chronic or suspicious ulcers</th>
<th>Appropriate shoe gear/bracing to prevent reoccurrence</th>
</tr>
</thead>
</table>


Closure or debridement should be sent for histologic analysis (Table 38-9).

After an ulcer is healed, it is important to prevent its recurrence. The patient requires constant education and reinforcement. They must understand that they are constantly at risk for problems, and most will need a lifetime of care in the form of either treatment or preventive intervention. Placement of the patient into the appropriate footgear plays a pivotal role in prevention. The patient will most likely need a shoe with a pressure-relieving insole. The insoles redistribute plantar pressure over a larger surface area, thereby reducing abnormal peak pressures. Pressure areas, deformity, and type of insole determine the type of shoe. Generally, therapeutic shoes can be divided into depth shoes, which have $\frac{3}{16}$-in. additional depth when the insole is removed; extra-depth shoes, which provide $\frac{3}{8}$-in. additional depth; super-depth, which provide $\frac{1}{2}$-in. additional depth, and custom-molded shoes, whereby the shoe is molded to a cast of the patient’s foot (120). The extra-depth or super-depth shoe is needed to accommodate most custom-molded insoles. The shoes usually have a blucher style opening, are lightweight, and have soles that are shock absorbing. Additionally, various braces for more complicated deformities can attach to or fit into these shoes. Insoles can be modified with fillers to accommodate prior amputations. Digital amputations usually require only a spacer, whereas ray amputations and transmetatarsal amputations require a distal filler. More proximal amputations may require shortened molded shoes, rigid rocker bottoms, or high-tops with rigid counters (121,122). Various modifications to the shoe can be added to enhance ambulation and unload pressure areas. Lateral and medial heel wedges may be used to control varus and valgus deformities. Sole wedges can be added to improve gait patterns. Metatarsal bars can be placed on the outsole of the shoe to shift weight from the metatarsal heads. A rocker sole can be added to decrease pressure across the forefoot aspect of the foot. This can be incorporated with an extended steel shank, which can be used to decrease the magnitude of the toe break and facilitate roll off. When there is limited ankle motion, a solid ankle cushion heel (SACH) can be added to the heel to provide simulated plantar flexion and aid in shock absorption (121,123) (Figs. 38-24 and 38-25).
Proper treatment of diabetic foot ulcers begins with prevention. Patients must have a comprehensive treatment plan including adequate education. They should be trained not to self-treat and to seek professional care. Even the simple shaving of a callus must not be performed by the patient. A simple cut in an uncontrolled environment could result in loss of limb. Health care providers who have a comprehensive understanding of the pathology and feel comfortable in rendering care should provide treatment of the diabetic foot. No one person can provide all of the treatment so that it is necessary that management of the patient be undertaken in a comprehensive multidisciplinary approach. All diabetic patients should be considered at risk and their treatment tailored to fit their deformity and circumstances. The presence of the classic triad of neuropathy, peripheral vascular disease, and immunopathy must be understood in order to prevent loss of limb and provide effective treatment. It is imperative that ulcers and infections be recognized for what they are and treated aggressively but also to understand the significant ramifications of the destructive nature of neuroarthropathy. The ultimate goals should be to lower amputation rates and help maintain good function and quality of life of our patients.

REFERENCES

PART III | MAJOR CONDITIONS

The Prevention and Treatment of Osteoporosis

Bone health is important to overall health and quality of life. Bones provide a frame that permits mobility, and protects internal organs from injury, while being a storehouse for minerals vital to the self-sustaining functions of daily life. Although osteoporosis is the most common bone disease, because there are no warning signs before fracture, it continues to be underrecognized. It is often not treated, even after osteoporosis-related fractures have occurred. Disfigurement, chronic pain, depression, disability and increased mortality can result (1). Relatively few people are diagnosed in time for effective therapy to be administered to prevent a first or a second fracture. In the United States in 2007, the national average for osteoporosis screening after hip fracture was 20% (2). This research reflects similar worldwide data (3). In a recent Canadian study, of 2,075 women age 45 or older with fracture, 81% were from a standing height (designated fragility fracture); however, 6 to 8 months after fracture only 21% of patients had had osteoporosis screening or treatment (4).

Compared to performance measures for other diseases, in-hospital osteoporosis management of hip fracture patients in American hospitals has lagged considerably. For instance, after myocardial infarction, doctors prescribe a recommended β-blocker 93% of the time prior to hospital discharge; similarly, there is 81% compliance with lipid assessment post-stroke (2). Low bone mineral density (BMD) can be considered a precipitating factor for fractures, just as uncontrolled hypertension can lead to stroke. Although advances in many areas of osteoporosis research, patient care, and health care coordination have occurred in the last 10 to 15 years, we have far to go.

As osteoporotic fractures often result from a combination of reduced quantity and quality of bone, relative neuromuscular instability, and environmental hazards, physiatrists have the expertise and opportunity to apply the multidisciplinary rehabilitation model to successfully address medical, functional and communication factors that have contributed to the development of this serious public health issue. Shinchuk et al. found that 69.8% of community dwellers (median age: 60.2) admitted to a subacute rehabilitation facility after acute hospital stay had decreased BMD (5). Because many fractures, particularly those of the hip, pelvis, and lower extremity, require in-patient rehabilitation services, physiatrists are thus in a strategic position to implement a successful osteoporosis screening cascade.

DEFINITION

Fractures occur when applied loads are in excess of the capacity of the bone, which is dependent on the degree of bone mineralization and bone architecture. Osteoporosis is a disease characterized by low bone mass and deterioration of the micro-architecture of bone tissue, particularly trabecular bone; this leads to increased bone fragility and fracture risk (Fig. 39-1). Although bone mass is reduced in osteoporosis, the remaining bone demonstrates the normal composition of both organic (40%) and mineral components (60%).

The current procedure of greatest clinical utility for measuring bone density is the dual-energy x-ray absorptiometry or DXA. In 1994, the World Health Organization (WHO) established the term normal bone density to designate bone density within one standard deviation (SD) of the mean for normal young adults, and low bone density, or osteopenia, as the designation for those with bone density 1.0 to 2.5 SD below the mean for young adults. The number of SD from the normal young mean is designated as a T-score on the final DXA report. In the recent past, the diagnosis of osteoporosis was based solely on a bone mass measurement of 2.5 SD or more below this mean i.e., (T-score ≤ 2.5) (See Table 39-1) (6). Another means of expressing bone density is as a comparison to persons of the same age, gender, and ethnicity; this is reported as one’s Z-score and is clinically most useful in following bone mass in children and adults less than 50 years old. In the last decade, however, the financial burden and impact of osteoporotic fractures on quality of life (Table 39-2) became statistically more important in treatment guidelines. As a result, in 2008, the WHO expanded the designation of osteoporosis beyond only statistical determinants to include clinical criteria: those patients with osteopenia who have had a fragility fracture, particularly of the spine or hip, are also now defined as having osteoporosis (7). When osteoporosis occurs without association with other conditions, it is defined as primary osteoporosis; in women, it is referred to as postmenopausal osteoporosis, and senile osteoporosis in men. Osteoporosis is also commonly seen in association with other diseases or conditions (e.g., rheumatoid arthritis, vitamin D deficiency), medications (e.g., phenytoin, corticosteroids), or disuse with immobility. These cases can be called secondary osteoporosis. These conditions have generalized bone mass reduction and increased risk for fracture in common (Table 39-3).
PART III | MAJOR CONDITIONS

EPIDEMIOLOGY

General Population and Persons with Disability

The National Osteoporosis Foundation (NOF) estimated in 2008 that more than 12 million people in the United States would have osteoporosis by 2010, and 40 million would have low bone mass, or osteopenia (8). By the year 2020, these numbers are expected to rise to 14 and 47 million, respectively. Unchecked, these changes could double or triple the number of hip fractures in the United States by 2040 (9). It is important to remember that persons with osteopenia far outnumber those with osteoporosis; thus, more fractures occur in the osteopenic population (3). Fractures are also more likely to occur in trauma patients who have pre-existing low bone mass or

<table>
<thead>
<tr>
<th>Condition</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>BMD is within 1 SD of a “young normal” adult (T-score ≥ −1)</td>
</tr>
<tr>
<td>Osteopenia</td>
<td>BMD is between 1 and 2.5 SDs below that of a “young normal” adult (T-score between −1.0 and −2.5)</td>
</tr>
<tr>
<td>Osteoporosis</td>
<td>BMD is 2.5 SDs or more below that of a “young normal” adult (T-score ≤ 2.5)</td>
</tr>
</tbody>
</table>

Note: Definitions are based on WHO assessment of bone mass measurement at any skeletal site in white women (176). Abbreviation: SD, standard deviation.

TABLE 39.1 Defining Bone Loss by BMD

Table 39.2 Loss of Quality of Life Years in Osteoporosis

<table>
<thead>
<tr>
<th>Event</th>
<th>QALYs Lost Due to Event</th>
<th>Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hip fracture</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acute event</td>
<td>0.0833</td>
<td>Complete loss of quality for 1 mo (= 1/12)</td>
</tr>
<tr>
<td>Rehabilitation or short-stay hospital (8 days)</td>
<td>0.0237</td>
<td>Complete loss of quality for 9 days (= 9/365)</td>
</tr>
<tr>
<td>Readmitted (8 days)</td>
<td>0.0219</td>
<td>Complete loss of quality for 8 days (= 8/365)</td>
</tr>
<tr>
<td>Home care services (6 mo)</td>
<td>0.25</td>
<td>Quality of life reduced by 0.5 for 6 mo (= 0.5 × 6/12)</td>
</tr>
<tr>
<td>Nonmedical home care (6 mo)</td>
<td>0.25</td>
<td>Quality of life reduced by 0.5 for 6 mo (= 0.5 × 6/12)</td>
</tr>
<tr>
<td>Post-hospital physician visits</td>
<td>0.011</td>
<td>Quality of life reduced by 0.5 for 8 days (= 0.5 × 8/365)</td>
</tr>
<tr>
<td>ER, ambulance</td>
<td>0.0027</td>
<td>Complete loss of quality for 1 day (= 1/365)</td>
</tr>
<tr>
<td>Wrist fracture, acute event</td>
<td>0.0404</td>
<td>Quality reduced by 0.3 for 7 wk (= 0.3 × 7/52)</td>
</tr>
<tr>
<td>Vertebral fracture, acute event</td>
<td>0.0324</td>
<td>33%: clinically silent with no loss of quality 57%: quality of life reduced by 0.5 for 1 mo 10%: complete loss of quality for 1 wk, and then loss of quality by 0.5 for an additional 7 wk (= (0.57 × 0.5) + 0.1 × (1 × 1/52) + (0.5 × 7/25))</td>
</tr>
</tbody>
</table>

ER, emergency room.
Data from the National Osteoporosis Foundation.

osteoporosis (10). Of the 10 million people in the United States estimated to have osteoporosis, women account for 8 million of those affected, and men for 2 million; another 12 million men are estimated to be at risk for the disease (11). When compared with other ethnic/racial groups, risk is increasing most rapidly among Hispanic women (12).

In the more generalized population, more than 2.0 million fractures per year in the United States in 2005 were directly related to osteoporosis. Of those, approximately 25% were at the hip and pelvis (13) and 30% at the spine (11). One-half of American women and up to 25% of men will have an osteoporotic fracture in their lifetime. The prevalence of osteoporosis at the hip is 17% for white, 14% for Hispanic, and 6% for black postmenopausal women (14). As life expectancy increases, osteoporosis will become more prevalent in men and women. Its importance as a public health problem is underscored by the fact that the lifetime risk of hip fracture in women is larger than the sum of lifetime risks of having breast, endometrial, and ovarian cancer. At present, hip fractures constitute on average 77% of the cost burden of fragility fractures in the United States (15).

Because of lower bone mass accrual in youth, and higher rate of bone loss in mid- and late life, women are more susceptible than men to osteoporosis (16). In the past, osteoporosis has been largely neglected in men, but research shows it is an important clinical and public health problem. Based on current WHO diagnostic criteria for osteoporosis, its prevalence is 4%, 2%, and 3% among white, Mexican-American, and black men age 50 and older, respectively (17). It is now recognized that the prevalence of hip fracture in men is approximately one-third to one-half that of women of similar age. Mortality after fracture in men, however, is consistently higher than that in women (18). Although hip fracture rate in women is two to three times that of men, 1-year mortality after hip fracture for men is nearly twice that of women (19). Although women lose bone mass rapidly during menopause, by age 70, calcium absorption has decreased in both sexes, resulting in an equal rate of bone loss by age 65 in men and women. After the age of 75 years, osteoporosis affects half the population, men and women equally.

The epidemiology of the Women’s Health Initiative identified clinical risk factors and biomarkers for 5-year hip fracture risk in postmenopausal females, and enhanced our knowledge of race and ethnic differences in this population (20). The prevalence of osteoporosis in white women is similar to that of Hispanic and Asian women. However after age 50, 20% of white and Asian women and 7% of men are diagnosed with osteoporosis compared to only 5% of non-Hispanic black females, and 4% of males (14). While African-Americans are less likely to have osteoporosis, once diagnosed, they have the same increased risk of fracture (14). After hip fracture, black women have a higher mortality than white women, thought to be due to in part to more advanced age and differences in medical care (19).

Osteoporosis occurs commonly in the rehabilitation patient population in both its primary and secondary forms. In Greek patients 1-year post-stroke, Lazoura et al. demonstrated that bone loss at the paretic hip relative to the nonparetic hip in males (11.8% at the femoral neck, and 10.4% at the greater trochanter) was less than perimenopausal women of the same age (13% and 12.6%, respectively). However, contrary to U.S. trends, there was no statistical difference between male and female in prevalence of osteopenia (53.3% and 52.2%, respectively), but males were more likely to have osteoporosis (20% vs. 13%) (21).

Smeltzer et al. demonstrated that community-dwelling American women with disabilities (control group) had a higher incidence of osteoporosis (22.6% vs. 7%) and low BMD (53.1% vs. 40%) than nondisabled postmenopausal women; only 50.9% of the controls were postmenopausal, with mean age of 50.6 (22). This corroborates findings by Nosek et al.
that women with disabilities develop osteoporosis earlier (23). In Smeltzer’s DXA screening of subjects, the highest incidence of osteoporosis was seen in women with spina bifida (69.2%), spinal cord injury (65%), post-polio myelitis (44.2%), and cerebral palsy (40%). Risk factors included Caucasian race (87.6%), lack of exercise (64.6%), menopausal status (50.9%), and medication-associated risk (44.8%). Estrogen replacement therapy was the most common prescribed treatment (19.7%), with alendronate prescribed for 5.6%. Only one-quarter of these women with disabilities had been previously screened for BMD, and only one third were taking calcium supplementation (22).

Secondary osteoporosis is not restricted to disabled adults; children with disabilities, including cerebral palsy and juvenile idiopathic arthritis, are also susceptible (24,25). Although screening and treatment protocols are not as well studied as in adults, these children are also at increased risk for low bone mass, osteoporosis, and fractures compared to their peers (26–29). Pediatric patients with cerebral palsy are particularly susceptible to spontaneous fractures (30).

Loss of bone mass with immobilization is most dramatically demonstrated in spinal cord injury patients, in whom sublesional bone mineral loss occurs in the lower limbs and pelvis; in paraplegic patients, and also in the upper limbs in tetraplegic patients. Dauty et al. showed in 2000 that sublesional BMD in spinal cord injury patients decreased 41% relative to controls at 1-year post-injury. It is most prominent at the tibia (~70%) and distal femur (~52%), the most common fracture sites (31). The spine bone mass does not significantly decrease (32).

The morbidity associated with osteoporotic fractures is high (33). In 1995, there were greater than one-half million hospitalizations, and 800,000 emergency room encounters secondary to osteoporotic fractures. Of these, hip fractures are the most devastating (34). Hip fractures account for nearly 50% of all osteoporotic fracture hospitalizations in the United States, compared to 8% for vertebral fractures (35). In terms of resulting disability, WHO data show that, post-fracture, one hip fracture is the equivalent to four vertebral fractures, or twenty fractures at other sites (3). The direct cost of osteoporosis is estimated at $12.2 to 17.9 billion annually in the United States. Because of increased life expectancy, the number of hip fractures could increase three- to eightfold by 2040 (9). Therefore, early screening and implementation of exercise, diet, fall prevention, and pharmaceutical strategies will become even more crucial.

Most of the social and economic burden of osteoporosis relates directly to hip fracture as well (15). Although some fracture patients suffer only temporary disability, many patients face deformity, loss of function, dependence, or institutionalization. Hip fracture almost invariably results in hospitalization and is a strong risk factor for acute complications (36).

Fewer than 50% of hospitalized patients with hip fracture recover their prefracture competence in activities of daily living (ADLs); 80% are unable to perform at least one instrumental ADL, such as shopping or driving (37) and only 25% regain previous levels of social functioning (38). Nine of every 100 elderly, white female hip fracture patients will die because of that fracture within 5 years (39). Of those who are ambulatory before hip fracture, studies have shown that 20% require long-term care afterward (11). Hip fractures are now clearly associated with increased mortality as well; approximately 20% of these patients will die within 1 year of their fracture (40).

Although less debilitating than hip fractures, fractures at the spine, wrist, and other sites are more common, and result in considerably more morbidity than hip fractures. Fractures of the vertebrae are the most common, more than 700,000 per year in the United States (11), and are largely responsible for the “dowager’s hump” deformity. These fractures, when severe, may cause chronic back pain, loss of height, and disability (41), as well as balance dysfunction, and altered abdominal anatomy, leading to abdominal distention, constipation, pain, and reduced appetite. Multiple thoracic fractures can lead to restrictive lung disease as well (42–44). Because of adverse changes in the ability to perform ADLs, the resulting impairment may be equal to that seen after hip fracture (45). Wrist fractures more likely to result in short-term disability, such as pain, loss of function, nerve entrapment (particularly carpal tunnel syndrome), bone deformity, and arthritis. Past studies have demonstrated a 30% risk of complex regional pain syndrome with wrist fractures (46,47).

The WHO has established a quantitative value to the impact of osteoporotic fractures by determining the quality-adjusted life year (QALY) associated with these fractures, with perfect health for 1 year assigned a QALY of 1, and death assigned a QALY of 0. With this measure, a disability that reduces a person’s self-assessed quality of life by half, compared with perfect health, is assigned a QALY of 0.5. A focus group of postmenopausal women generally agreed with determinations made by the NOF Committee from the WHO QALY model. Similarly, the effectiveness of various pharmacologic treatments in preventing fractures and their consequences was based on available evidence from randomized controlled trials. Although effectiveness of rehabilitation strategies was not reviewed, the assumptions of effectiveness coupled with costs enable the calculation of the expected cost per QALY for any interventional strategy. With this methodology, a statistician can determine the likelihood of a particular strategy providing a favorable cost:benefit ratio for the health care system. Table 39-2 provides examples of QALYs assigned to various outcomes after NOF analysis (48).

ETIOLOGY AND RISK FACTORS

Fracture risk is dependent on the specifics of an individual’s genetic profile, peak bone mass, and strength of bone achieved in one’s lifetime and the subsequent rate of bone loss. There are identified risk factors and causes of low bone mass and osteoporosis that contribute to this fracture risk model. In primary osteoporosis, multiple etiologic factors may act independently or in combination in an individual patient to produce diminished bone mass. In secondary osteoporosis, specific
Risk factors in persons with disability can differ from the general population. For instance, reduced mobility, vitamin D deficiency, progressive time from onset of disability, and gender in stroke patients have been documented (52). Although the prevalence of osteoporosis in the traumatic brain injury population is not well studied, high risk for hypogonadism and immobility in this population is likely to predispose them to low bone mass (53). Time after onset of spinal cord injury, and associated decrease in mobility, also correlates with increased risk in the lower limbs for low bone mass and fractures (32).

### PATHOGENESIS

Osteoporosis is a heterogeneous disease with multiple causes. Although the pathogenesis of bone mass loss in secondary osteoporosis may be readily apparent (e.g., corticosteroid excess, or the lack of muscle effect on bone with subsequent negative bone remodeling in paraplegia), the exact pathogenesis of primary osteoporosis may be more difficult to define. Low bone mass may be attributable to failure to achieve adequate bone mass at skeletal maturity (age 13 to 25 for women) and/or subsequent age-related and postmenopausal bone loss. Although low bone mass is principally associated with fracture, other determinants of fracture include the quality of the bone (e.g., trabecular architecture), the ability to heal trabecular microfractures, and the propensity to fall (54,55). The pathogenic basis of inadequate bone mass, particularly in the elderly, also may be considered from the standpoint of tissue, cellular, and hormonal abnormalities.

### Tissue Abnormalities

Although cellular and hormonal abnormalities undoubtedly contribute to osteopenia and osteoporosis, the basic abnormality in all types of osteoporosis is a disturbance of the normal bone-remodeling sequence at the tissue level. Therefore, to fully understand the pathogenesis of osteoporosis, knowledge of bone remodeling is necessary.

Bone is constantly turning over (remodeling). The skeleton is a reservoir for 99% of the body’s calcium; remodeling provides calcium to the organism without sacrificing the skeleton. In addition, remodeling allows bone mass to respond to increased and decreased muscle activity (e.g., bone mass is increased in a tennis player’s dominant arm). The initial event in bone remodeling with normal bone turnover is an increase in bone resorption, as mediated by the osteoclast (the cells responsible for bone resorption). This event is typically followed within 40 to 60 days with an increase in bone formation, as mediated by the osteoblast (the cells responsible for bone formation). Bone resorption and formation are normally “coupled”: an increase or a decrease in resorption produces a corresponding increase or decrease in formation, so that the net change in bone mass is zero. In postmenopausal osteoporosis, and possibly in senile osteoporosis as well, bone resorption is increased without a corresponding increase in bone formation, thereby leading to a net loss in bone mass. In this case, bone remodeling is described

---

### TABLE 39.4 Risk Factors for Osteoporotic Fractures

| Personal history of low-impact fracture |
| Current low BMD |
| Hip fracture of either parent |
| Caucasian race |
| Advanced age |
| Female sex |
| Dementia |
| Recurrent falls |
| Inadequate physical activity |
| Poor health/ frailty |
| Current smoker |
| Low body weight |
| Estrogen deficiency |
| Corticosteroid use |
| Testosterone deficiency |
| Vitamin D deficiency |
| Low lifetime calcium intake |
| Alcoholism |
| Impaired eyesight despite correction |
as “negatively uncoupled.” In other forms of osteoporosis, particularly those associated with corticosteroid-induced osteoporosis, a primary decrease in bone formation may occur. The end result is the same—a net loss of bone mass, with concomitant increasing risk for fracture as bone density decreases. Abnormalities of bone remodeling at the tissue level will therefore contribute to the pathogenesis of osteoporosis.

**Cellular Abnormalities**

Conclusive evidence of cellular abnormalities contributing to the pathogenesis of primary osteoporosis is building. In its simplest terms, the rate of growth and activity of osteoblasts falls behind that of osteoclasts, resulting in a loss of bone density. These processes may be separate from abnormalities of bone cells that occur with aging alone. For instance, it may be a failure of the osteoblast, as a result of either decreased cell number, or decreased cell activity, which accompanies advancing age, but is not specific for osteoporosis. Recent investigations of the communication system between osteoblasts and osteoclasts hold promise; particularly the osteoblast CFU-M and RANKL proteins, which stimulate osteoclast activity (Fig. 39-2A), and its protein osteoprotegerin (OPG), which binds RANKL and inhibits osteoclast activity (Fig. 39-2B). The balance between RANKL and OPG production may determine the rate of bone resorption (56). In 2001, a study by Gong et al. associated patients with a dysfunctional receptor (lipoprotein receptor–related protein-5) with severe osteoporosis and pseudoglioma (57); overactivity in this receptor has been shown to form strong bones.

**Hormonal Abnormalities**

Many hormonal agents may affect bone cell function and bone mass. Although there are numerous age- and menopause-related alterations in the physiology of these hormones, a specific pathogenetic hormonal abnormality in osteoporosis (excluding the osteopenia associated with hypercorticism and hyperparathyroidism) has not been conclusively defined. Estrogen deficiency remains the most frequently incriminated factor in the pathogenesis of postmenopausal osteoporosis in women, while testosterone deficiency is considered a potential cause in younger men.

Indeed, estrogen deficiency of any etiology, including menopause, early oophorectomy (58), and functional hypogonadism associated with chronic strenuous exercise (female athlete triad [FAT]) (59), should be considered a prime risk factor for bone mass loss. Estrogen deficiency from any cause results in increased skeletal responsiveness to parathyroid hormone (PTH) and increased bone resorption, a transient increase in the

---

**FIGURE 39-2.** Osteoporosis is caused by disruption of normal bone remodeling sequences. (A) Role of RANK and RANKL in bone remodeling. (B) Regulation of RANK/RANKL binding by osteoprotegerin (OPG). (Delmas PD. J Clin Densitom. 2008;11:325–338, Fig. 39-4A. Adapted with permission from Boyle WJ, et al. Osteoclast differentiation and activation. Nature. 2003;423:337–342.)
increasing to year 19 after injury in the 40- to 59-year age group before plateau. This study also showed that BMD in the spine actually increased from weight bearing 1 to 9 years after injury in all study participants, though becoming significant only in those with paraplegia in the 20- to 39-year age group 10 to 19 years after injury (32).

Fractures occur in this population most frequently in the pelvis and lower extremities, particularly the tibia, and correlated with sites of most bone density loss. Similar to the general population, women with spinal cord injury tend to develop osteoporosis more frequently than men with spinal cord injury. Research is needed comparing risk of osteoporosis in patients with spinal cord injury of different races.

Genetic Abnormalities
The degree of peak bone mass achieved and the amount of bone loss as we age determine our risk of osteoporosis. Complex genetic and environmental factors determine these contributing factors. It is likely that a cohort of genes contribute to a predisposition to osteoporosis and that they differ among different ethnic backgrounds (65). Definitive identification of all candidate genes is not established, but numerous gene susceptibilities are implicated, including abnormalities in receptors for the active form of vitamin D₃ (calcitriol), estradiol, and PTH as well as genes coding for TGF-β and IL-6. These genes are associated with achieving peak bone mass and bone remodeling processes, and are active throughout the lifespan (66).

Spinal Cord Injury
Although disuse is thought to play a role in the development of osteoporosis in this population, neural factors are also implicated, as is an impaired PTH—vitamin D axis and calcium/phosphate metabolism (67). There is no demineralization in supraspinal areas following injury, and weightbearing in the spine is thought to limit loss of bone mass in vertebrae. It is likely that some of this increased spine bone density was due to development of osteophytes as suspected in all patients age. Spasticity, degree of injury, female sex, age, and duration after injury negatively influence bone mass. Reduced intestinal absorption and increased renal elimination of calcium, inhibition of sex steroids, pituitary suppression of Thyroid Stimulating Hormone (TSH), and insulin resistance and IGFs may also contribute (68), but the majority of patients with spinal cord injury have decreased bone mass below the lesion. Animal models have demonstrated increased osteoclast activity, with severe bone loss (48% trabecular and 35% cortical), decreased mineral apposition, and growth plate abnormalities consistent with osteoblast dysfunction (69), and elevated RANKL mRNA induction (70).

In his study of the spinal cord–injured population, Szollar et al. found that serum levels of calcium and calcitonin could not be correlated with changes in bone mass. However, in this study of patients with spinal cord injuries, the PTH level decreased in the first year after injury and started to increase 1 to 9 years after injury. Osteoblast activity decreased immediately after injury in these patients, with consequent dramatic decreases in bone degradation. In fact, loss of BMD in the proximal femur was measurable at 12 months following injury in all 176 persons with paraplegia or tetraplegia. Of significance, risk for fracture increased between years 1 and 9 after injury in the 20- to 39-year age group, but continued...
in all 176 persons with paraplegia or tetraplegia. Of significance, risk for fracture increased between years 1 and 9 after injury in the 20- to 39-year age group, but continued increasing to year 19 after injury in the 40- to 59-year age group before plateau. This study also showed that BMD in the spine actually increased from weight bearing 1 to 9 years after injury in all study participants, though becoming significant only in those with paraplegia in the 20- to 39-year age group 10 to 19 years after injury (32).

Fractures occur in this population most frequently in the pelvis and lower extremities, particularly the tibia, and correlated with sites of most bone density loss. Similar to the general population, women with spinal cord injury tend to develop osteoporosis more frequently than men with spinal cord injury. Research is needed comparing risk of osteoporosis in patients with spinal cord injury of different races.

**CLINICAL EVALUATION FOR OSTEOPOROSIS**

Although the first clinical indication of osteoporosis, either primary or secondary, is usually a fracture, it is optimal to screen and treat patients, both male and female, prior to first fracture.

**Quantitating Bone Mass**

A number of clinical parameters positively correlate with bone mass, such as paraspinal muscle strength in postmenopausal women and grip strength in premenopausal women and men. Although these are useful methods, they do not substitute for precise quantitation (71). Plain radiographs of the spine are relatively insensitive in quantitating bone mass because 30% to 35% of bone mass must be lost before demineralization is detected. They are, however, sensitive and reliable to assess for fractures of the spine when unexpected height loss is noted, or when fractures are suspected in axial or peripheral bone.

Several noninvasive procedures have been developed over the past 50 years to quantitate bone mass (bone density). The current procedure of greatest clinical utility, DXA is noninvasive, quantifies primarily trabecular (cancellous) bone at the spine and hip with an acceptable precision and accuracy, is reasonably simple to perform at a reasonable cost, and is associated with low radiation exposure. Most importantly, it predicts which patients are at risk for subsequent fracture and can be repeated to assess therapeutic response to treatment. The DXA technique also satisfies requirements for ionizing radiation safety to the greatest degree, as bone density measurements can be obtained within 30 seconds to 2 minutes with a radiation exposure of approximately 10 mrad—one-sixth the exposure of a chest x-ray—with 99% precision and approximately 97% accuracy. CT measurements of the spine provide an exclusive assessment of trabecular bone, and actual volumetric density, but the technique is compromised by high radiation exposure and lower precision.

WHO recommendations for baseline and follow-up testing are based on DXA measurements at axial sites (spine and hip). For patients in whom the spine cannot be measured, the forearm is substituted. The spine and the distal forearm contain primarily trabecular bone, metabolically more active than cortical bone, and preferentially altered in osteoporosis, thus most affected by medications used in the treatment of osteoporosis. The hip and parts of the forearm also have cortical bone. The vertebrae and hips are also the fracture sites likely to cause the most disability. WHO guidelines are adapted to the epidemiology of individual countries and are well described in special position papers by the NOF Guide Committee (72,73). As WHO DXA standards were established from research of white, postmenopausal women, some controversy exists in interpretation of DXA for men, premenopausal females, nonwhite populations and children. The International Society for Clinical Densitometry (ISCD), for example, recommends use of a single normative database for all women, and the use of a male normative database to calculate T-scores for men (74,75).

Revised 2008 WHO guidelines for DXA screening and treatment shifted from a T-score-only basis for treatment to evidence-based assessment of the 10-year risk for fracture at hip and other sites, based on age, gender, race, bone density values and known osteoporosis and fracture risks. Guidelines were then adapted to demographic and health care profiles of individual countries (72,73). Interpretation of these guidelines, though, can vary among diverse governmental and medical specialty organizations. All women 65 or older (men perhaps at age 70), and patients with known fragility fractures after age 45, particularly at the hip or spine, should be screened for osteoporosis via DXA. Screening should also be considered for younger adult women and men with at least two known risks, such as early menopause, tobacco abuse, corticosteroid dependence, or testosterone-blocking agents (see Fig. 39-3, Algorithm for Osteoporosis Management). Protocols for DXA and laboratory screening in children and young adults with chronic disability (i.e., cerebral palsy or spinal cord injury), and with medical conditions associated with osteoporosis such as eating disorders and FAT, are not well established, but the ISCD has published recommendations (www.iscd.org/Visitors/positions/OP-Index. cfm) (74). Fracture criteria in these populations warrant at least baseline DXA assessment. Screening, interpretation of results, and treatment for loss of bone density must be individualized until more research can be done in these populations.

The FRAX WHO Fracture Risk Assessment Tool (Fig. 39-4), developed at the University of Sheffield, in England by Kanis et al. assists the clinician in treatment guidelines for patients with low BMD (osteopenia, DXA T-score between −1.0 and −2.5) without previous pharmaceutical intervention (78). After inserting data reflecting patients’ gender, age, race, body weight, and height, known osteoporosis risk factors, and femoral neck bone density (in gm/cm²) with the manufacturer of DXA equipment, the FRAX tool establishes a treatment threshold in accordance with WHO guidelines. In the United States, if the
Etiologic Factors Contributing to the Risk of Osteopenia/Osteoporosis

1. Estrogen depletion
   a. Postmenopausal state (natural or artificial)
   b. Exercise-induced amenorrhea, anorexia nervosa
2. Calcium deficiency
   a. Inadequate calcium intake
   b. Malabsorption
   c. Lactose intolerance
3. Diminished peak bone mass at skeletal maturity; varies with sex (women > men), race (whites > blacks), and heredity
4. Diminished physical activity
5. Testosterone depletion
6. Aging
7. Low body weight (adipose tissue is the major source of extragonadal estrogen production postmenopause)
8. Alcoholism; smoking
9. Excessive coffee intake (>4–6 cups daily); excessive dietary protein or salt intake (increased calcium loss in the urine)
10. Medications: corticosteroids, thyroid hormone, phenytoin

Data from the National Osteoporosis Foundation, Prevention, NOF.org.

Bone Markers

Serum and urine markers of bone resorption and formation are diagnostic modalities used primarily to monitor efficacy of prescription therapy. In this sense, the markers may be complementary to bone density assessment. Current rate of bone resorption is most commonly assessed with a urinary marker n-telopeptide of type I collagen (NTX), and c-telopeptides of type I collagen (CTX) (80). The n-telopeptide (NTX) measurement results from 24-hour urine calcium collection, normalized for creatinine; urine is collected within 48 hours of routine serum osteoporosis laboratory studies to best define calcium metabolism pathway abnormalities (as are seen in Table 39-5, Laboratory Tests in Disorders of Calcium Metabolism). However, for convenience, a second-void fasting spot urine NTX can be preferred, due to calcium intake variance during a 24-hour period. Serum levels of NTX and CTX can also be tested. It should be noted, however, that all bone turnover markers have considerable day-to-day variability. Thus, many experts use them in only specific cases when knowing the rate of bone turnover is necessary.

Clinical Evaluation of the Patient at Risk for Osteoporosis

Patients at risk for osteoporosis require careful evaluation, consisting of the following elements. A thorough history is required to determine the presence of risk factors for osteoporosis such as menopausal status, family history of hip fracture, and certain medications (Table 39-6), and to exclude medical conditions leading to secondary osteoporoses (see Table 39-3). A separate intake questionnaire for osteoporosis to expand standard intake history intake can be helpful. History of previous fragility fractures and sites of persistent pain (i.e., atraumatic vertebral compression fracture) must be identified. Identify any history of falls or associated risk factors such as poor vision, bladder urgency, or peripheral neuropathy (Table 39-7). Document any loss of height from early adulthood (average loss of 2 to 3 inches from occiput to sacrum is expected between the ages of 40 and 80 (49)), and include in DXA prescription if greater than 1.5 inches. Assess current level of physical activity and exercise, and past history of eating disorders, in all patients. Social behaviors such as tobacco abuse or excessive alcohol are both positively correlated with bone loss and should be noted (81). The importance of the social history cannot be overemphasized, especially in the elderly, who may have history of frequent falls, require assistive devices or personnel in their living environment, or need consideration for transitional care unit placement. Inadequate daily calcium intake and exercise, vitamin D deficiency, corticosteroid use, diabetes mellitus, and multiple

<table>
<thead>
<tr>
<th>Disorder</th>
<th>Serum Calcium</th>
<th>Serum Phosphate</th>
<th>Vitamin D Hydroxy</th>
<th>PTH</th>
<th>1,25(OH)2 Vitamin D</th>
<th>Urine Calcium</th>
<th>Renal Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary hyperparathyroidism</td>
<td>↑</td>
<td>↓</td>
<td>Variable</td>
<td>↑</td>
<td>Normal</td>
<td>↓</td>
<td>Variable</td>
</tr>
<tr>
<td>Familial hypocalciuric hypercalcemia</td>
<td>↑</td>
<td>Variable</td>
<td></td>
<td></td>
<td>Normal</td>
<td>↑</td>
<td>Variable</td>
</tr>
<tr>
<td>Hypercalcemia of malignancy</td>
<td>↑</td>
<td>Variable</td>
<td>Variable</td>
<td></td>
<td>Normal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vitamin D deficiency</td>
<td>↓ or normal</td>
<td>↓ or normal</td>
<td>↓↓</td>
<td>↑ or normal</td>
<td>Usually normal</td>
<td>↓</td>
<td>Variable</td>
</tr>
<tr>
<td>Renal osteodystrophy</td>
<td>↓</td>
<td>↑</td>
<td>Variable</td>
<td>↑↑</td>
<td>↑</td>
<td></td>
<td>Variable</td>
</tr>
<tr>
<td>Primary hypoparathyroidism</td>
<td>↓</td>
<td>↑</td>
<td>Variable</td>
<td></td>
<td>↓</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: While 25 (OH) vitamin D levels are variable in many disorders, low levels are very common in general. Renal function is often normal in the various disorders, but it may be decreased. Patients may have more than one disorder. For example, some patients with primary hyperparathyroidism may also be vitamin D insufficient, leading to further elevation of serum PTH.
myeloma are common in the elderly. Inquire into pending dental procedures, as dental extractions could delay start-up of bisphosphonates, the most commonly prescribed medication class for osteoporosis, due to risk of osteonecrosis of the jaw (82,83). History of certain malignancies may be a relative contraindication for teriparatide, a potent anabolic treatment option.

A thorough physical examination establishes cognitive status, assesses oral hygiene and hydration status, and excludes causes of secondary osteoporosis (e.g., hemiplegia, rheumatoid arthritis, anorexia, spinal cord injury). One must assess for pre-existing fractures; for example, multiple vertebral fractures can induce severe kyphosis, with anterior-posterior widening of the rib cage, increasing its proximity to the iliac crests, and inducing abdominal protuberance (49). Document any risks for fall (i.e., visual disturbance, neurologic deficits, contractures, leg length discrepancy, poor balance with transfers, gait abnormalities, and improper use of assistive devices). Evaluate the potential for safe weight-bearing and resistive exercise (i.e., cognitive status, cardiopulmonary status, posture, degree of kyphosis, balance and pain with active and resisted motion).

Document height and weight, contracture limitations and leg length discrepancies after measurement. Assess the general mobility of spine and joints of the extremities, as well as abdominal, spinal, and extremity muscle strength. Verify sites of pain (i.e., vertebræ T8-L2 are associated with osteoporosis, whereas fractures at T6 or above are more likely associated with malignancy) (49). Assess for tibial tenderness in thin, female runners, especially with irregular or absent menses, as seen in FAT. Identify any risks for intolerance of prescription medications; for instance, poor dentition, history of gastric disorder such as GERD, gastritis or peptic ulcer, or diffuse myalgias can delay or preclude bisphosphonate use.

Proximal muscle weakness and chronic corticosteroid use will require special exercise focus. Proprioception, balance, transfers and gait must be evaluated. Proper use and design of assistive devices should also be addressed. Specific physical performance measures that correlate with higher bone mass density in hip and spine, wrist or whole body in postmenopausal women include longer step length, normal and brisk gait speeds and step length, longer single leg stance and grip strength (84).

**TABLE 39.7** Major Risk Factors for Falls

<table>
<thead>
<tr>
<th>Risk Reduction Strategies</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Demographic</strong></td>
</tr>
<tr>
<td><strong>Environmental</strong></td>
</tr>
<tr>
<td><strong>Neuromuscular</strong></td>
</tr>
<tr>
<td><strong>Medical</strong></td>
</tr>
</tbody>
</table>

| **Risk Reduction Strategies** | Light hallways, stairs, entrances, bathroom | Clear clutter/loose cords; move furniture | Anchor or eliminate rugs | Install grab bars; high commode seat | Sturdy shoes; assistive device | Nonskid mats; grab bars; tub bench/chair | Encourage use of sturdy, low-heeled shoes | Stair rails; cane/walker |
| | | | | | | | | |
| | High level balance challenge exercises | Cane or walker; tai chi | Resistive exercise; optimize vitamin D levels | Optimize myofascial release/postural training | Sturdy shoewear; balance training; cane/walker | Mobility training |
| | | | | | | | |
| | Annual visual examination | Medication; timed voids; avoid PM fluids | Hydrate; optimize medications | Annual medication review | Counseling support; medications | Counseling; abstinence | Nutritionist; Home Health Nursing Consult | Mobility training; counseling |

Adapted from data from the National Osteoporosis Foundation, 2002–2009.
TABLE 39.8 Basic Osteoporosis Laboratory Tests

<table>
<thead>
<tr>
<th>Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complete blood cell count</td>
</tr>
<tr>
<td>Serum chemistry (renal electrolytes, liver enzymes, BUN, creatinine, calcium, total protein/albumin, alkaline phosphatase, and phosphorus)</td>
</tr>
<tr>
<td>Vitamin D-25 hydroxy</td>
</tr>
<tr>
<td>Intact PTH</td>
</tr>
<tr>
<td>Serum protein electrophoresis</td>
</tr>
<tr>
<td>Thyroid function test</td>
</tr>
<tr>
<td>24-h urine calcium</td>
</tr>
<tr>
<td>Urine markers for bone resorption-urine NTX*</td>
</tr>
</tbody>
</table>

**Serum NTX can be substituted.**

Evaluation of these parameters and improving where possible with resistive exercises may improve bone mass parameters.

A basic laboratory evaluation is listed in Table 39-8. In primary osteoporosis, results of laboratory tests typically are normal (except 25-hydroxyvitamin D); the primary role of blood and urine tests is therefore to exclude other diseases and in a few cases of urinary NTX, to establish baseline bone turnover rate. For example, multiple myeloma should be suspected with anemia, abnormal serum protein electrophoresis (SPEP), and elevated B-cell population in complete blood count. Vitamin D deficiency is best evaluated by serum 25-hydroxyvitamin D. Serum IgA antitissue transglutaminase and IgA endomysial antibody, if positive, can be an indication of malabsorption (i.e., celiac disease). Low urinary calcium to creatinine ratio or a 24-hour urinary calcium are less specific measures of malabsorption in some cases.

Vitamin D deficiency may lead to secondary hyperparathyroidism. Severe vitamin D deficiency causes osteomalacia, with bone pain and poor mineralization of bone. More commonly, milder degrees of vitamin D deficiency lead to decreased gut absorption of calcium and in some cases secondary hyperparathyroidism, causing loss of bone mineral. Many osteoporosis patients have some degree of vitamin D inadequacy. Vitamin D deficiency also has an effect on muscle, leading to decreased lower body strength and increased propensity to fall. If 24-hour urine calcium value is low, inadequate calcium intake or absorption, or vitamin D deficiency, is likely. If urine calcium value is high, either dietary calcium excess or idiopathic hypercalciuria is a possibility. If the serum calcium is elevated, measurement of PTH is the most important test to do. Primary hyperparathyroidism leads to bone loss and must be differentiated from familial hypocalciuric hypercalcemia (FHH), a benign abnormality of the calcium receptor. Patients with FHH have mild elevations of serum calcium and PTH but very low urinary calcium excretion. It should be noted that patients with primary hyperparathyroidism may have vitamin D deficiency, leading to secondary hyperparathyroidism as well. See Table 39-5, Laboratory Tests in Disorders of Calcium Metabolism.

DXA screening is of value in the individual patient with history of spine or hip fragility fracture, or two or more risk factors for low bone mass. (See Fig. 39-3, Osteoporosis Indications for Prevention and Treatment of Osteoporosis. Each patient must be assessed individually. Clinical judgment remains very important in the assessment and management of osteoporosis.

![FIGURE 39-3. Osteoporosis Management Algorithm.](image-url)
FIGURE 39-4. FRAX WHO Fracture Risk Assessment Tool to assess 10-year probability of fracture (%). (A) 57 y.o. female (B) 77 y.o. female with same risk factors, BMI and T-score. Note increase in fracture risk with age as only factor. Current tobacco use would increase 57 y.o. female major osteoporotic fracture risk to 9.9% and hip fracture risk to 1.9%, and 77 y.o. female hip fracture risk to 5.3%, without increase in risk of other major osteoporotic fracture. Adapted with permission of Kanis, WHO Collaborative Centre for Metabolic Bone Diseases, University of Sheffield, UK. Available at: http://www.shef.ac.uk/FRAX/tool.jsp.
Management Algorithm for defining DXA screening candidates. If osteoporosis is found by DXA, treatment is indicated. For those patients with osteopenia, the WHO FRAX questionnaire (Fig. 39-4) can be used to determine if prescription medications are needed for improving bone mass. If DXA/FRAX analysis indicates that fracture risk is 3% or more at the hip, or 20% or more at other sites, more aggressive prophylactic therapy (e.g., bisphosphonate) is recommended. If bone mass measurements at the spine, hip, and wrist are normal by DXA, increased exercise, a diet rich in calcium, and calcium and vitamin D supplements may be sufficient fracture prophylaxis, in conjunction with fall reduction training. Such patients should have a repeat DXA in 2 years, with recalculation of the fracture risk via the FRAX tool at that time if osteopenia persists.

Iliac crest bone biopsy is used primarily to exclude osteomalacia or other metabolic bone diseases, such as is seen with late-stage renal failure. Although such biopsies can be used to define high and low bone turnover, this is not usual practice at this time for the typical patient with osteoporosis.

Educational materials can be provided to all patients to reinforce the importance of maintaining bone health, understanding the sequelae of untreated osteoporosis, identifying sources of dietary calcium, and improving fall risk in the home at the time of office visit (http://www.nof.org/).

### Preventive Strategies for Patients at Risk

In the last decade, efforts to improve bone health have focused on prevention and treatment protocols. Prevention programs focus on adequate nutritional status, including optimization of calcium, vitamin D and protein intake, and monitoring for excessive fats or carbohydrate intake. Avoidance of lifestyles known to result in bone loss, including cigarette smoking, excessive alcohol, and possibly carbonated beverage intake is also important. Weight-bearing and strengthening exercises, and fall prevention strategies in the home and community, are equally important.

Because bone mass is the principal, although not the only, determinant of fracture risk, preservation or improvement of bone mass via pharmacologic means is associated with a reduced risk of fracture. The rationale of the various therapeutic agents available for preserving or improving bone mass density or bone mass is based on knowledge of bone remodeling. In normal bone, there is no net change in the amount of bone mass present, as on-going bone remodeling is a balance between bone resorption, and the process of bone formation. In most forms of osteoporosis, however, a perturbation of bone remodeling occurs. Bone resorption increases over normal levels, and bone formation does not compensate for this increase, with a net loss of bone mass overall.

### Nutritional Adjuncts

#### Calcium

Dietary and supplemental calcium intake is a mainstay of osteoporosis prevention and treatment. The Surgeon General’s

---

**TABLE 39.9 Recommended Daily Calcium Intake**

<table>
<thead>
<tr>
<th>Age or Life Stage</th>
<th>Adequate Calcium Intake (mg/d)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–6 mo (human milk content)</td>
<td>210</td>
</tr>
<tr>
<td>7–12 mo (human milk + solid food)</td>
<td>270</td>
</tr>
<tr>
<td>1–3 years</td>
<td>500</td>
</tr>
<tr>
<td>4–8 years</td>
<td>800</td>
</tr>
<tr>
<td>9–18 years</td>
<td>1,300</td>
</tr>
<tr>
<td>19–50 years</td>
<td>1,000</td>
</tr>
<tr>
<td>&gt;50 years</td>
<td>1,200</td>
</tr>
<tr>
<td>Pregnancy or lactation</td>
<td>1,300</td>
</tr>
<tr>
<td>≤18 years</td>
<td>1,300</td>
</tr>
<tr>
<td>19–50 years</td>
<td>1,000</td>
</tr>
</tbody>
</table>

Vitamin D

Vitamin D facilitates absorption of calcium and mineralization of bone. It is found in liver, fatty fish, egg yolks and as an additive in foods such as milk, orange juice, and cereals. It can be taken as a supplement and is also synthesized in the skin through sunlight exposure. As many experts believe that the recommended dietary intake should be higher for children and younger adults, the Institute of Medicine is likely to raise the Recommended Daily Allowance for this vitamin, with final recommendations expected in 2011. Current recommendation for daily intake among experts in the field is 800 to 1,000 IU/day of vitamin D3 for men and women age 50 or older (95). The active form of vitamin D (1,25-di hydroxy vitamin D or calcitriol) is also beneficial in osteoporosis and is commonly prescribed as a supplement for patients who lack the 1-α-hydroxylase enzyme because of severe renal impairment. Cholecalciferol (vitamin D₃) is probably the preferred form of vitamin D supplement, but ergocalciferol (vitamin D₂) can also be used as an oral dose of 50,000 IU, weekly, for 4 to 12 weeks as replacement therapy when low vitamin D levels are found. Severe deficiency can lead to secondary hyperparathyroidism, osteomalacia or rickets, and is associated with increased risk of osteoporotic fracture (96,97).

As these supplements increase calcium absorption at the gut level, their use can result in an increased risk for kidney stones or for hypercalcuiaria, nephrolithiasis, or even nephrocalcinosis in patients at risk. Increased urinary calcium is the first sign of such toxicity and is easily evaluated with a 24-hour urine sample. Many institutionalized and house-bound elderly patients are deficient in vitamin D and can benefit from its administration (91,92,98,99). In general, total daily administration of vitamin D should not exceed 2,000 IU (100). As most studies have focused on postmenopausal women and the elderly, more research is needed in vitamin D effects on children, premenopausal females, men, and diverse ethnic and racial groups (101). Although there is evidence of benefit of vitamin D supplementation in rickets, hyperthyroidism, fall reduction, and BMD, standardization of 25(OH) vitamin D assays has not been achieved. The effect of sunscreen on activation of vitamin D, and determination of optimal sunlight exposure without increased risk of skin cancer, requires further investigation (102). As calcium and vitamin D are critical to the action of bisphosphonates, normal levels must be verified prior to start of bisphosphonate therapy.

### TABLE 39.10 Selected Food Sources of Calcium

<table>
<thead>
<tr>
<th>Food, Standard Amount</th>
<th>Calcium-mg</th>
<th>Calories</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fortified ready-to-eat cereals (various), 1 oz</td>
<td>236–1043</td>
<td>88–106</td>
</tr>
<tr>
<td>Plain yogurt, nonfat (13 g protein/8 oz), 8-oz container (low-fat&lt;sup&gt;a&lt;/sup&gt;)</td>
<td>452(415)</td>
<td>127(143)</td>
</tr>
<tr>
<td>Soy beverage, calcium fortified, 1 cup</td>
<td>368</td>
<td>98</td>
</tr>
<tr>
<td>Fruit yogurt, low fat (10 g protein/8 oz), 8 oz</td>
<td>345</td>
<td>232</td>
</tr>
<tr>
<td>Orange juice, fortified, 1 cup</td>
<td>308–344</td>
<td>85</td>
</tr>
<tr>
<td>Swiss cheese, 1.5 oz</td>
<td>336</td>
<td>162</td>
</tr>
<tr>
<td>Sardines, Atlantic, in oil, drained, 3 oz</td>
<td>325</td>
<td>177</td>
</tr>
<tr>
<td>Fat-free (skim) milk, 1 cup&lt;sup&gt;a&lt;/sup&gt;</td>
<td>306</td>
<td>83</td>
</tr>
<tr>
<td>1% low-fat milk, 1 cup (whole milk&lt;sup&gt;a&lt;/sup&gt;)</td>
<td>290(276)</td>
<td>102(146)</td>
</tr>
<tr>
<td>Plain yogurt, whole milk (8 g protein/8 oz), 8-oz container&lt;sup&gt;a&lt;/sup&gt;</td>
<td>275</td>
<td>138</td>
</tr>
<tr>
<td>Tofu, firm, prepared with nigari&lt;sup&gt;b&lt;/sup&gt;, ½ cup</td>
<td>253</td>
<td>88</td>
</tr>
<tr>
<td>Mozzarella cheese, whole milk, 1.5 oz</td>
<td>215</td>
<td>128</td>
</tr>
<tr>
<td>Pink salmon, canned, with bone, 3 oz</td>
<td>181</td>
<td>118</td>
</tr>
<tr>
<td>Collards, cooked from frozen, ½ cup</td>
<td>178</td>
<td>31</td>
</tr>
<tr>
<td>Molasses, blackstrap, 1 tbsp</td>
<td>172</td>
<td>47</td>
</tr>
<tr>
<td>Soybeans, cooked, ½ cup, green (mature)</td>
<td>130(88)</td>
<td>127(149)</td>
</tr>
<tr>
<td>Ocean perch, Atlantic, cooked, 3 oz</td>
<td>116</td>
<td>103</td>
</tr>
<tr>
<td>Oatmeal, plain or flavored, instant, fortified, 1 prepared packet</td>
<td>99–110</td>
<td>97–157</td>
</tr>
<tr>
<td>Pizza, cheese</td>
<td>100</td>
<td>255</td>
</tr>
<tr>
<td>White beans, canned, ½ cup</td>
<td>96</td>
<td>153</td>
</tr>
<tr>
<td>Broccoli (raw), 1 cup</td>
<td>90</td>
<td>25</td>
</tr>
<tr>
<td>Okra, cooked from frozen, ½ cup</td>
<td>88</td>
<td>26</td>
</tr>
<tr>
<td>Ice cream, vanilla, ½ cup</td>
<td>85</td>
<td>135</td>
</tr>
</tbody>
</table>

Source: Nutrient values from Agricultural Research Service (ARS) Nutrient Database for Standard Reference, Release 17. Adapted from 2002 revision of USDA Home and Garden Bulletin No. 72, Nutritive Value of Foods. Food sources of calcium ranked by mg of calcium and calories per standard amount. Bioavailability may vary. (All dairy are ≥20% of AI for adults 19–50, which is 1,000 mg/d.)

<sup>a</sup>Calcium content varies slightly by fat content; the more fat, the less calcium the food contains.

<sup>b</sup>Calcium content is for tofu processed with a calcium salt; other salts do not provide significant calcium. See http://www.nal.usda.gov/fnic/foodcomp/Data/SR20/nutrlist/sr20a301.pdf for a more comprehensive list of foods containing calcium.
Protein
Along with calcium and vitamin D supplementation, protein supplementation has been shown to improve healing and decrease mortality in persons who have sustained hip fractures. In a study by Schurch et al. (103), elderly patients received dietary protein supplements of 20 g/day for 6 months following hip fracture. Much of the rapid bone loss that usually occurs in the contralateral hip during the year after a fracture was avoided in these patients. The RDA for protein is 46 g/day for women and 56 g/day for men. This can be achieved with two to three servings of meat or beans, and two to three servings of milk and cheese per day (104). Too much protein, however, may be harmful, as it can incur an acid load; calcium from bone is a major source of serum alkaline buffering in this instance.

Exercise
A lifelong dedication to physical activity (60 minutes daily for all children aged 8 and older, and 30 minutes for adults) (105) and exercise is recommended. Therapeutic exercise is an essential element of the rehabilitation program for patients with osteoporosis, and should be tailored to the patient’s level of fitness and anticipated propensity to fracture. Exercise increases muscle and bone strength, joint flexibility, and balance, and prevents falls. Although genetic factors can determine a large proportion of bone mass and structure, up to 50% to 90%, controllable lifestyle factors contribute as well (106). A thorough history and physical examination and, if appropriate, bone density testing, can address deficits and define exercise precautions to avoid injury. An exercise program should incorporate both short- and long-term goals, which must be reviewed with the patient. Patient education concerning proper posture, body mechanics, and increasing strength and aerobic capacity is an essential component of both short-term and long-term interventions. Osteoporosis is a disease that is progressive and, if unchecked, can cause severe disability. In patients with physical disabilities, inactivity can be especially harmful because of a propensity for further deterioration with aging, and concomitant health problems. The PASIPD questionnaire (Physical Activity Scale for Individuals with Physical Disabilities) can be useful in assessing levels of activity in the persons with disability, such as home maintenance function, moderate and vigorous sport and recreation activities, occupation, and transportation (107).

For optimal bone health, an exercise program should include weight-bearing activities for 45 minutes 3 or 4 times per week, or weight lifting for 20- to 30 minutes 2 or 3 times per week (49). Weight-bearing, low-impact activities include walking or treadmill use; jogging, tennis, and soccer are high-impact activities. Muscle-strengthening activities such as jumping, weight lifting, and use of resistive equipment would be classified as moderate to vigorous activity. Swimming is beneficial from a cardiovascular standpoint, but is not a weight-bearing sport, and thus does not contribute to bone mass (108). Swimming does maintain muscle mass, however, which contributes to fall reduction. Balance training should be included to prevent falls for patients with fall risks (see Table 39-7). Regular exercise has been shown to improve bone mass (0.5% to 3.0%), with 20% to 45% reduction in hip fractures in older populations with moderate-to-vigorous physical activity (49,109–111). Further investigation is needed to ascertain the complex relationships between physical activity, bone mass, and osteoporotic fractures.

The following general principles should be considered when recommending therapeutic exercise.

Life Span Bone Phases
Bone health is determined by many factors (112). Increased physical activity, especially weight-bearing and resistive exercise, stimulates bone building via the piezoelectric effect of contracted muscle on bone. In defining general life phases in which bone changes are likely to occur, the clinician can target screening and exercise recommendations more easily. During childhood and adolescence, bone undergoes a growth phase, especially during puberty. This is followed by a maintenance phase during young and middle adulthood. A mid-life phase, ages 50 to 70, can be characterized as a bone loss phase; after age 70, a frailty phase is more likely. CDC surveillance data report about 16.7% of adults aged 45 to 64, and 23.1% of adults aged 65 to 74 are inactive. As Americans age, their participation in physical activity declines, with women persistently lagging behind men: 61% of adults reporting never engaging in vigorous physical activity, with 66% of women compared to 56% of men in this category of inactivity (113). Lower levels of physical activity (lack of vigorous exercise) correlated with lower levels of education (80% of those without high school diploma vs. 43% of those with bachelor’s degree or higher) and lower income (72% of poor families compared with 52% of families classified as not poor) (114).

The most critical period of bone growth is thought to be during puberty and adolescence, with the most gains, on average 25% to 30% of adult bone mass, between age 12 to 14 in girls and 13 to 15 in boys (115). Most of these gains reflect increased bone length and size, not bone density (116). Bone mass gained during this period was found to typically equal to the amount lost later in adult life (117). As bone mineralization lags behind growth in length, fracture rates increase during periods of rapid growth (118). In a school-based program that emphasized jumping activities over 7 months, only early pubertal girls improved in bone mass (1.5% to 3.1% at the femoral neck and lumbar spine when compared with controls), with no benefit noted in the late prepubertal experimental group of girls (119).

Exercise need not be strenuous to confer benefit (120–122). The gentle, slow movements involved in the practice of t’ai chi have been gaining in popularity with the elderly. Benefits have been reported in balance, strength, cardiovascular fitness, respiratory function and flexibility, and decreased injury. Wolf et al. reported a 47.5% decrease in multiple falls in assisted living residents aged 70 to 97 during a 1-year program for elderly patients (123). Li et al. demonstrated a 55% reduction in multiple falls, and reduced injury from fall (7% vs. 18% in controls) with a three times weekly t’ai chi program during a 6-month period (124). Later studies demonstrated delayed bone density...
loss in postmenopausal female t’ai chi practitioners (125–127). Early research in whole body vibration also shows preliminary benefit in balance in residents of long-term facilities (128), muscle strength (129,130), and bone mass (130).

**Exercise Principles**
The following general principles should be considered when recommending therapeutic exercise (131).

**Principle of Specificity**
Exercise should stress the specific physiologic system being trained. In patients with normal bone mass, exercise activities should stress sites most at risk for fracture (i.e., the hip, spine, and wrist). Exercise for osteogenesis may be prescribed for the purposes of prevention of osteoporosis even for those with minimal reductions in bone mass. High-impact exercise such as jumping and strength training is more beneficial to improving bone mass than low-to-moderate intensity exercise such as brisk walking, and has positive bone mass effect at spine and hip, in as little as 5 to 10 min/day (132). However, these exercises may not be appropriate for the frailty-phase patient population, because patients with very low bone mass and multiple fractures need skeletal protection while building strength, and increasing balance and flexibility. Isometric, core-strengthening exercises with the spine in neutral position should be emphasized in all patients; spinal flexion should be avoided in those patients with abnormal mineral density, particularly the frail (133).

**Principle of Reversibility**
The positive effect of exercise will slowly be lost if the program is discontinued. Hence, a lifelong dedication to exercise and physical activity is necessary for optimal bone health, as changes associated with aging can have profound physical and clinical consequences. On average, adults lose 5% of muscle mass per decade after age 30, with more rapid decline possible after age 65 (134). However, many patients with sarcopenia (age-related loss of skeletal muscle) and impaired balance remain highly functional. This muscle loss may be reversed by exercise—as shown by Fiatarone et al. (135). A comprehensive exercise program must increase bone mass and muscle strength (136) with continued variety to prevent abandonment of exercise goals. It is important to emphasize strengthening the bone while it still has adaptive ability; this is particularly important in persons with disability and the elderly. While passive standing is not recommended for active adults for exercise, in a person with severe disability such as spinal cord injury, through the use of a standing frame or standing wheelchair, it is a rehabilitative strategy shown to prevent bone loss (137).

**Principle of Progression**
To increase bone mass, the stimulus must exceed previous bone-loading activity. Thus, there must be a progressive increase in the intensity of the exercise for continued improvement. Kerr’s study of resistive exercises at the hip demonstrated that exercise benefits are site specific and result from progressive resistance exercise with maximum loading as opposed to an endurance regimen (138). To avoid injury, however, applied loads must be within the capacity of the bone to sustain mechanical stress. Progressive resistance is important for both bone health and improved functional capacity (139). Slowly increasing time or intensity of exercise by 10% weekly decreases risk of injury (140).

**Principle of Initial Values**
Those who initially have low capacity will have the greatest functional improvement from a given program (141). Nevertheless, it is important for inactive participants to begin with short exercise sessions of low intensity and progress systematically to limit risk of injury.

**Principle of Diminishing Returns**
There is a biologic ceiling to exercise-induced improvements in function. As this ceiling is approached, greater effort is needed to achieve minimal gain.

Optimal calcium intake is thought to be synergistic with physical activity throughout life in improving bone mass (142–144) and should be encouraged in all patient populations where medical contraindications do not coexist. In meta-analysis of randomized clinical trials, Shea et al. concluded that calcium supplements in postmenopausal women reduced bone loss by 2%, with an approximate 23% reduction in spine fractures (145). Vitamin D supplements may reduce spine fractures by as much as 7% (146). Exercise must also be supported with adequate caloric intake for energy expended for all patients, as indicated by adequate glycemic index. Weight loss, which sacrifices adipose tissue and depletes estrogen production stores sufficiently to cause amenorrhea in females, increases the risk of osteoporosis and stress fractures. More than 3 months without menstruation should prompt clinical evaluation in these patients (see Female Athlete Triad [FAT], below).

The prevention of falls and fractures through an ongoing exercise program that maintains proper nutrition, strength, and aerobic capacity should be coupled with adjunctive measures such as the provision of adequate support for the spine, pain management, and psychological support when developing objectives for long-term goals.

**Fall Reduction Strategies**
Approximately one-fourth of community-dwelling persons 65 years of age will fall (147); this number doubles for residents in nursing homes (148). Although less than 1% of falls result in fracture, according to the Northeast Hip Fracture Study Group, 90% of hip fractures are due to a fall (149). The propensity to fall may be as important as bone mass, given the frequency and severity of falls in this age group. Falls have many etiologies, including decreased neuromuscular coordination of the elderly (resulting in an inability to prevent loss of balance or break a fall’s impact), mental status issues such as confusion and dizziness with medication, and environmental factors such as poor lighting and loose rugs. New research has examined differences in frequency, mechanism, and risk factors for falling between men and women ("http://www.cdc.gov/ncipc/factsheets/nursing.htm") (150). The loss of muscle mass can be as much as 3% to 5% per decade after
age 30, and is also commonly associated with falls and fracture in the elderly. Fiatarone et al. found 71% of elderly hip fracture patients were sarcopenic (151).

Degree of demineralization of bone and the trauma of falling are well established as risk factors for fracture. There are rare reports of spontaneous hip fracture, although one historical study in 1981 demonstrated that 11% of fractures were spontaneous, with 25% of them associated with standing or sitting transfers and 60% during simple ambulation (152). More research is needed in this area. These studies support the notion that fractures of the proximal femur can be the result of muscle forces acting on the hip exceeding the mechanical ability of the femur to withstand stress (153). Because hip fractures have multifactorial causes (154), interventions must aim not only at increasing BMD at the hips, but also at increasing muscular strength, balance, and flexibility and reducing the forces of impact when a fall occurs.

A successful program aimed at preventing falls includes education on how to eliminate identifiable fall risks, and exercises that prevent falls through improved balance, such as gait training, coordination and functional exercises, and muscle strengthening (155–157). Preservation of autonomy after a fracture can be aided by modifications made to the home environment before the patient is released from the hospital. Home visits from an occupational therapist have also been shown to be helpful in this regard (158).

Falls in long-term care facilities present very specific challenges. According to the CDC, approximately 5% of adults age 65 and older live in nursing homes in the United States, and about three out of four of these residents fall each year, which is twice the rate of older community dwellers. About 2% to 6% of these falls result in death. The most common causes of these falls are similar to those of community dwellers listed above. The most effective strategies to prevent falls include post-fall assessment and follow-through on findings; staff education; medication review; environmental changes such as raised toilet seats, lower bed height and hallway handrails; hip pads; bed restraints do not lower fall risk, and in fact, may increase the risk of injury and death with falls (150).

Mechanisms of Falls

Although much attention has been given to methods for increasing bone density, the scientific literature on mechanisms behind falls is less studied; Hayes and Cumming have contributed much in this area (152,159–162). The severity of the fall is an independent risk factor for hip fracture (163) and is related to many factors, including the direction of the fall and the specific anatomic location of major impact (161). Whereas young adults tend to fall to the side or backward, the elderly tend to fall sideways or drop in place, especially those with unsteady gaits (159). It is falls to the side that result in major impact forces that greatly exceed the mechanical strength of the proximal femur and therefore result in fracture.

The risk of falling appears to increase linearly with the number of risk factors present. The most consistent predictors for falling include balance and gait abnormalities, impaired vision, decreased ADL function, polypharmacy, and cognitive impairment (147). Sedative use, particularly polypharmacy, is associated with falling independent of other risk factors. Benzodiazepines, phenothiazines, and antidepressants are used frequently in the elderly for dementia and depression, and use of these medications is associated with increased falling—particularly when taken in their longer-acting forms. Diuretics and antihypertensive agents are also associated with an increased risk of falling, as they can induce postural hypotension. An annual review of current medications, especially antihypertensives, psychotropics, sedatives, analgesics, antihistamines, and diuretics, should be included in fall risk assessment screening. Table 39-7 lists other major risk factors, including environmental hazards.

Fear of Falling

The elderly have a generalized fear of falling and are particularly fearful of not being able to get up after a fall. In comparison to fallers who were able to get up, Tinetti found that those unable to get up were more likely to suffer permanent decline in ADLs, be hospitalized more frequently, and have a higher mortality rate. The highest risk factors in this group were age greater than 80, depression, and problems with balance and gait (164). It is reasonable to expect fear of falling in the elderly, to see it reduce physical activity and conditioning, and thus become an independent risk factor for falling (165). Velozo modified the University of Chicago Fear of Falling Measure (UIC FFM) using Rasch analysis to develop a successful unidimensional construct that links a community-dwelling older person’s level of fear with functional activities, such as “stepping off a curb” or “standing on a moving bus” (166). This assessment can be useful in defining rehabilitative interventions in different population groups. Table 39-11, Fear of Falling Measure, lists the activities with clinical relevance.

For a fall to result in hip fracture, there must be impact near the hip that is not reduced by body mechanisms or absorbed by soft-tissue structures. The fall itself may be divided

### Table 39.11: Fear of Falling Measure

<table>
<thead>
<tr>
<th>1. Walk when icy</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. Carry bundles up poorly lit stairs</td>
</tr>
<tr>
<td>3. Use a step stool to reach in kitchen cabinet</td>
</tr>
<tr>
<td>4. Climb poorly lit stairs</td>
</tr>
<tr>
<td>5. Stand on a moving bus</td>
</tr>
<tr>
<td>6. Climb up bus stairs</td>
</tr>
<tr>
<td>7. Get in/out of bathtub</td>
</tr>
<tr>
<td>8. Carry bundles up well-lit stairs</td>
</tr>
<tr>
<td>9. Step off a curb onto the street</td>
</tr>
<tr>
<td>10. Walk on a crowded sidewalk</td>
</tr>
<tr>
<td>11. Climb up well-lit stairs</td>
</tr>
<tr>
<td>12. Use an escalator</td>
</tr>
<tr>
<td>13. Take a walk</td>
</tr>
<tr>
<td>14. Get in/out of a car</td>
</tr>
<tr>
<td>15. Carry a full plate to the dinner table</td>
</tr>
<tr>
<td>16. Pick something lightweight from the floor</td>
</tr>
</tbody>
</table>
into four phases (162,166): instability phase, where balance is lost; descent phase; impact phase; and post-impact phase.

A fall prevention program should address these four phases of a fall and present interventions at each level (167). General conditioning exercises, appropriate provision of assistive devices, adequate footwear, modification of medications, and attention to other risk factors for falling are targeted at preventing the instability that results in a loss of balance (168). There is ample evidence of the benefit of exercise increasing bone mass and preventing falls. These programs must be maintained, however, for continued benefit to accrue to bone mass and muscle strength (169). Several important determinants govern the forces applied to the femur as a result of falling. These forces include the person’s weight, thickness of subcutaneous tissue, height of the fall, configuration of the body during the fall, velocity at which the hip strikes the impact surface, and nature of the impact surface. Falling to the side raises the risk of hip fracture approximately sixfold, compared with the threefold increase in relative risk associated with a decrease of 1 SD in hip BMD (152,159–161). Contraction of the quadriceps and other muscles of the lower limb is likely to reduce velocity at impact and reduce impact forces at the proximal femur in falls to the side. Exercise programs aimed at increasing lower-extremity strength may therefore prevent hip fracture by reducing fall severity (138,164,169–172).

Meta-analysis by Moayyeri et al. demonstrated that moderate to vigorous physical activity is associated with hip fracture risk reduction of 38% in women and 45% in men (173). Low body weight and low BMI are associated with an increased risk of hip fracture in elderly men and women (154,163,170,171). Although the impact force most closely correlates with the individual’s weight, velocity on impact is most associated with height. Thus, although an individual may be heavier and the resultant force of impact more, he or she may have more padding over the trochanter in addition to stronger bone, thereby preventing fracture.

While the benefits of physical activity and exercise are well described, it is less clear which exercises are most efficacious in preventing falls and decreasing risk of osteoporotic fractures (174,175). Since weight-bearing exercises lead to site-specific improvement in BMD (176), it is reasonable to hypothesize that upper extremity strengthening and core strengthening would play a significant role in fall and fracture prevention, and improve overall function and general fitness, as lower extremity exercises have been shown to do. Evidence from the Tromso study suggests, however, that among middle-aged women, a high level of physical activity is related to increased risk of fracture at non–weight-bearing sites, including the wrist, proximal humerus, hand, and fingers (177). Carter et al. addressed the effects of improved resistance and agility training, but found no benefit in reducing fall risk in osteoporotic women aged 65 to 75. However, although trunk stabilization exercises were included in this “Osteofit” exercise program, only knee extension strength was evaluated as a measure of improvement in fall risk (178). Liu-Ambrose et al. however found decreased fall risk (37.4% to 43.3%) with regular exercise in women with low bone mass aged 75 to 85, but less increase in physical activity from resistive exercise (3.8%) than balance (29.2%) or stretching exercise (37.7%) (179). Thus, there is a clear need for further research in clarifying the relationship between specific exercise protocols and their contribution to fall and fracture reduction.

Recognition that increases in soft-tissue thickness around the hip substantially reduce peak force to the trochanter at fall impact has led to the development of hip protectors. Fractures have been prevented in women wearing hip protectors, and Lauritzen concludes that the use of hip protectors could reduce fracture by 53% (172). The Cochrane Database System Reviews of 2003 and 2007 indicated many trials have identified strategies to prevent falls (180,181), but standardization is lacking for outcome measures. The 2005 Cochrane review by Parker and Gillespie however showed no clear benefit from hip protectors in preventing fractures, citing poor compliance due to discomfort and practicality as important factors (182). Pads are designed in two primary configurations: a simple pad that covers the trochanter and reduces impact force by absorbing energy in the pad material, and one based on shunting energy from the fall away from the trochanter (183). The latter pad is designed in an inverted U shape and is filled with a colloidal-like substance that hardens on impact, permitting 68% reduction in peak force at the hip with impact. Softer pads tend to promote more compliance with wear (184).

Architectural modification of the flooring surface offers an approach to impact reduction that seems worth exploring, particularly in residential health care facilities and housing projects for the elderly. Research has shown that the least number of falls occurs with vinyl flooring, and wooden subfloors decrease the risk of fracture with falls compared with concrete subfloors (185). When compared with linoleum flooring, a thick pile carpet with underpad can reduce impact from a fall by 23%, but this solution seems impractical in light of the difficulties in ambulation presented by rugs of this design (186). Other flooring systems that might provide impact reduction without presenting an obstacle include low-pile carpeting and rubber-backed materials. Caution would have to be taken to air these products thoroughly so as not to introduce a confounding health risk in the form of toxic chemical exposure. There is continued research into shock-absorbent flooring, but it has proven to be expensive and impractical in homes and hospitals already built (185).

Psychosocial Considerations
Psychological issues have been noted to contribute significantly to disability in the osteoporotic patient after fracture. Depression is the most common psychological problem identified in these patients. A study of community-dwelling postmenopausal women found that those with osteoporosis had significantly higher depression scores than did those with normal bone density (187). Anxiety, fear, and other emotional reactions also affect post-fracture outcome. In a study of 200 women recovering from hip fractures, those with high depression scores following surgery were more likely to experience poorer recovery of function (188). In a 1993 study of 100 women with
osteoporosis-induced vertebral fracture, the women noted emotions as having greater importance than physical functioning, leisure and social activities, and ADLs (189). In this group, most reported fear of falling, fear of new fractures, frustration, anger, and feeling overwhelmed. Vertebral fracture patients also suffered loss of self-esteem, isolation, vulnerability, and embarrassment related to physical appearance (139).

There is increasing evidence that women with depression are more susceptible to low bone mass, with both antidepressant medications and imbalance of immune system inflammatory proteins implicated (190). Inflammatory proteins are induced by adrenalin, which is frequently elevated in depressed. Of these proteins, IL-6 is known to promote bone loss (191). Dudgeon et al. investigated the relationship of physical disability and chronic pain to changes in lifestyle. They found patients learned to suppress pain complaints to some degree because of perceived negative social consequences in the form of isolation, rejection, and consequent depression. Patients’ methods for coping with pain included distraction techniques (e.g., listening to music), which conferred some sense of control in a health care system in which the nature of their pain was not well understood, not addressed adequately, and peripheralized by health care providers (192).

**PHARMACOLOGIC AGENTS**

The ability of Food and Drug Administration (FDA)-approved pharmacologic agents to prevent or treat osteoporosis has been studied best in postmenopausal women; fracture data are more limited for men and for corticosteroid-induced osteoporosis. Thus, prescription treatments at present demonstrate fracture risk reduction best in those with osteoporosis by DXA and/or fragility fracture history, compared with osteopenic patients without fractures. Benefit and risks of therapy must be individualized to each patient. The therapeutic agents available for the treatment and prevention of osteoporosis are classified as decreasing bone resorption (antiresorptive agents) or stimulating bone formation (anabolic agents). The end result of each is the same: to preserve or improve bone mass, and thereby prevent fractures. As noted in Table 39-12, Bisphosphonate Comparison, and Table 39-13, Nutritional and Pharmacologic Treatments for Osteoporosis, most U.S. FDA-approved therapeutic agents decrease bone resorption. Reduction of hip fracture may lead to a reduction in mortality. Lyles et al. showed a 28% reduction in deaths from annual infusion of zoledronic acid within 90 days after repair of hip fragility fracture (193). A recent study by Gilchrist et al. with alendronate 70 mg weekly dosing in acute spinal cord injury patients demonstrated significant differences over placebo controls in total body BMD (+5.3%), and total hip BMD (+17.6%) at 1 year (194).

### Antiresorptive Agents

#### Bisphosphonates

The FDA-approved bisphosphonates, alendronate (Fosamax), risedronate (Actonel), ibandronate (Boniva), and zoledronic acid (Reclast) have the ability to preserve and increase bone mass at the spine, hip, and other sites to varying degrees, and in most cases prevent fractures at these sites. The FDA has approved their use for the prevention and treatment of osteoporosis for postmenopausal women, and men in most cases. The use of these medications in pediatric populations is promising, particularly in juvenile idiopathic arthritis, but more studies are needed (195,196). Their mechanism of action is well studied: they bind to the surface of bone and are taken up by osteoclasts during the bone resorption process. They subsequently precipitate early cell death by blocking essential lipid compounds within the osteoclast, slowing the resorption process. (See Tables 39-12 and 39-13 for summary of dosing, safety and efficacy of medications in this class and other treatments.)

Long half-life allows some bisphosphonates to accumulate and persist in bone, which promotes maintenance of bone density gains even after treatment is stopped. This permits treatment “holidays,” provided DXA and bone turnover markers remain

<table>
<thead>
<tr>
<th>Name</th>
<th>Treatment Dose (per d/wk/mo)</th>
<th>Increased BMD Sites*,b</th>
<th>Decreased Fracture Sites*,b</th>
<th>FDA Indications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ibandronate</td>
<td>2.5/na*/150 (3 mg IV/3 mo)</td>
<td>Spine, hip</td>
<td>Spine</td>
<td>RT</td>
</tr>
<tr>
<td>Zoledronic Acid</td>
<td>5 mg IV yearly</td>
<td>Hip</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note: Definitions are based on WHO assessment of bone mass measurement at any skeletal site in white women (176); data cited for increased BMD and Decreased Fracture Sites are not from comparable studies.

---

**TABLE 39.12 Bisphosphonate Comparison**

---

[Data adapted from NOF Clinician’s Guide to Prevention and Treatment of Osteoporosis.](#)
TABLE 39.13  Nutritional and Pharmacologic Treatments for Osteoporosis

<table>
<thead>
<tr>
<th>Medication</th>
<th>Usual Dosage</th>
<th>Mode of Action</th>
<th>Side Effects &gt; placebo (≥5%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Calcium</strong></td>
<td>1,200 mg/d</td>
<td>Decreased bone resorption</td>
<td>Increased urinary calcium</td>
</tr>
<tr>
<td><strong>Vitamin D₂ or D₃</strong></td>
<td>800–1,200 IU/d</td>
<td>Increased calcium absorption in GI tract</td>
<td>Low risk hypercalcemia</td>
</tr>
<tr>
<td><strong>Bisphosphonates</strong></td>
<td>–</td>
<td>Decreased bone resorption</td>
<td>Esophageal irritation (po) Osteonecrosis jaw (IV predominantly, in cancer pts.) atypical femoral fractures</td>
</tr>
<tr>
<td>Alendronate (Fosamax)</td>
<td>10 mg daily or 70 mg weekly (space po (mg daily or 35 mg weekly for osteopenia)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Risedronate (Actonel)</td>
<td>5 mg daily, 35 mg weekly or 150 mg monthly</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ibandronate (Boniva)</td>
<td>150 mg po monthly or 3 mg IV every 3 months</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zoledronic acid (Reclast)²</td>
<td>5 mg IV over 15 min yearly</td>
<td>Improved bone mineralization</td>
<td></td>
</tr>
<tr>
<td>Teriparatide (Forteo)</td>
<td>20 μg daily sq (max. 18-24 months Rx)</td>
<td></td>
<td>Leg cramps, dizziness</td>
</tr>
<tr>
<td>Calcitonin (Miacalcin)</td>
<td>200 MR IU/d (nasal spray)</td>
<td>Decreased bone resorption</td>
<td>Nasal irritation (rare) Possible increased risk of cancer, high blood pressure, deep vein thrombosis, stroke, heart disease, thromboembolic disease</td>
</tr>
<tr>
<td>Estrogen with or without progesterone</td>
<td>0.625 mg/d for 21–30 d (cycled 21 of 30 d)</td>
<td>Decreased bone resorption</td>
<td></td>
</tr>
<tr>
<td>Estrogen agonist/antagonist</td>
<td>60 mg/daily</td>
<td>Decreased bone resorption</td>
<td>Hot flashes, leg cramps, deep vein thrombosis</td>
</tr>
<tr>
<td>Raloxifene (Evista)</td>
<td>–</td>
<td>Decreased bone resorption via RANKL inhibition</td>
<td>Musc-skel pain, elevated cholesterol, cystitis pancreatitis</td>
</tr>
<tr>
<td>RANKL Inhibition</td>
<td>Denosumab (Prolia)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Stable. When ingested orally, these drugs must be taken at least one-half hour before any food, drink, or other medication in the morning, and the patient must maintain an upright position during that time to avoid the risk of esophageal irritation and, rarely, ulceration. Other side effects include visual disturbances and difficulty swallowing. Ninety-four percent of the case reports of osteonecrosis of the jaw with this class of drug are associated with cancer patients receiving multiple doses of IV bisphosphonates (83). While in one British study there was a higher risk of atrial fibrillation with zoledronic acid compared to placebo (1.3% vs. 0.4%), this risk is still under investigation with other bisphosphonates and has not led to any FDA-induced change in prescription recommendations (197). Bisphosphonates should be prescribed with caution in patients with severe renal dysfunction and are not considered safe for pregnant women (198). Compliance with medication dosing must be monitored. Siris et al. found only 43% of women over age 45 were compliant with refills of bisphosphonate prescriptions, and only 20% persistent without gaps in refills at 2 years; only compliant patients had significant reductions (20% to 45%) in fractures (199). It must also be stressed to the patient the need for concomitant calcium and vitamin D supplementation to maximize efficacy of the bisphosphonates, as noncompliance with supplements is also common and reported as low as 40% (200).

**RANKL Inhibition**

Research on the remodeling mechanisms involving osteoblasts and osteoclasts has resulted in the development of a new category of anti-resorptive medication, an inhibitor of RANKL (nuclear factor-κB ligand). Denosumab (Prolia) decreases bone resorption by inhibiting the earliest stages of osteoclast maturation. The antibody prevents RANKL from interacting with the receptor RANK, truncating the earliest stages of the osteoclast maturation cascade (See Figure 39-2). Denosumab is administered via subcutaneous injection (60mg) twice yearly. Side effects include musculo-skeletal pain, elevated cholesterol, cystitis, pancreatitis and infection. It is approved for the treatment of postmenopausal osteoporosis.

**Hormone Therapy**

**Estrogen**

Estrogen is important to bone development throughout life for both men and women. Because it acts on both reproductive and nonreproductive tissues in the body, consideration for use...
of its exogenous forms, alone or in combination with proges-
terone, must balance its benefits with the individual patient’s medical and family history. It was first approved by the FDA for postmenopausal osteoporosis in 1972, after approval in 1942 for relief of menopausal symptoms (201). A recent trial by the Women’s Health Initiative (202) reported the first definiti
data supporting a benefit of postmenopausal hormones in the prevention of fractures at the hip and spine by at least one third compared to placebo. One arm of this placebo-controlled trial involved the administration of an estrogen/proges-
tin combination tablet (0.625 mg estrogen, 2.5 mg progestin [Prempro, Wyeth-Ayerst, Philadelphia, PA]) versus placebo to approximately 17,000 postmenopausal women with an intact uterus. A second arm of the study evaluated estrogen alone in women who had hysterectomies (203). Both studies were halted early because evidence of harm was found in the form of increased incidences of stroke, deep vein thrombosis, and cognitive impairment. Increased risk of coronary heart disease, invasive breast cancer, and pulmonary embolism were found in the combined estrogen/progestin study. The development of these conditions was determined to outweigh the benefits that were achieved in counteracting osteoporosis and prevent-
ing colorectal cancer. Despite these concerns, estrogen supple-
mentation continues to be an important part of the overall treatment of osteoporosis in women. Hormone replacement therapy in women is FDA-approved for prevention of osteo-
porosis only, and is to be taken in conjunction with calcium and vitamin D. Several studies have indicated that withdrawal of hormone therapy can result in decline in bone density and higher rate of hip fracture (204,205). Research with lower-dose hormone therapy and newer “designer estrogens” offers promise of beneficial effects without the detrimental side effects.

Raloxifene

Selective estrogen receptor modulators (SERMs) are now called estrogen agonist/antagonists and have been developed to provide beneficial effects similar to those obtained with estrogen, but without the adverse effects. They have an agonistic effect on bone and lipoprotein production, while being antagonistic toward breast tissue, without effect on uterine mucosa. Only Raloxifene (Evista) is FDA-approved in this class, for the treat-
ment of postmenopausal osteoporosis, and prevention of bone loss in recently postmenopausal women; it is not recommended for relief of menopausal symptoms (74). The use of raloxifene provides modest increases in bone mass, but reduction in the risk of vertebral fracture is 40% to 50%, and there is no reduc-
tion in nonvertebral fractures (45). It also appears to reduce the risk of estrogen-dependent breast cancer, but increases the risk of deep vein thrombosis to a degree similar to that of estrogen. Newer research in this class focuses on increased benefit to bones, heart and breast tissue (206).

Combination Antiresorptive Therapies

The efficacy and safety of combined bisphosphonates with either hormone therapy or SERMs is being studied; as they have different mechanisms of action, there may be potential for an additive effect when used together. Bone et al. showed 8% increase in spine BMD in combination, compared to 6% with either alone (207). However, as there is no proven addi-
tive effect on fracture reduction, combination therapies are considered experimental (208).

Calcitonin

Calcitonin is a natural hormone secreted by parafollicular cells within the thyroid gland and has an inhibitive effect on osteo-
clasts. Although the PROOF trial noted a decrease in spine fractures by 33% with 200 IU daily dose, (209), later stud-
ies showed no significant difference in nonspine fracture rates after 5 years. Its use subsequently declined. Current research is focused on oral and higher-dose nasal spray forms. It may have limited potential to decrease pain after acute compression frac-
ture in the spine, and has few side effects of significance: nasal congestion, nosebleeds, and nausea.

Anabolic Agents

Teriparatide

The FDA approved the use of teriparatide (Forteo), a recombi-
nant human PTH fragment (PTH 1 to 34) for the treatment of osteoporosis in 2002 after initial studies showed increased absorption of calcium and phosphorus, and a remarkable increase in bone turnover, with bone formation outweighing bone resorption (210). It is administered by subcutaneous injection daily (20 µg, in a preassembled multiple-dose pen device). It has an anabolic effect on bone through thicken-
ing of the bone cortex, and increasing connections within the bone matrix. It is approved for treatment of postmenopausal osteoporosis and for men with idiopathic or hypogonadal osteoporosis who are at high risk for fracture, and for those who have failed or are intolerant of previous treatment (74). In postmenopausal women with osteoporosis, BMD increased by 9.7% in the spine and 2.6% at the hip, with fracture reduction of 65% at the spine and 53% at nonspine sites (211). Data for fracture risk improvement in men have not been established, but a 1998 study showed BMD increases of 5.9% in spine and 1.2% at the hip in this population (212). Common side effects include dizziness and leg cramps. Although the approved dose regimen has not been found to increase risk of osteosarcoma in humans as has been observed in laboratory animal studies, this medicine is presently not prescribed for patients with a his-
tory of bone carcinoma or bone metastases. Nor is it thought safe in pediatric patients, or adult patients with hypercalcemia, Paget’s disease, or kidney disease (213). Neither efficacy nor safety of treatment beyond 2 years is established. Because of the biphasic circadian rhythm of endogenous PTH secretion, and its sensitivity to nutritional intake and growth hormone, there may be novel pharmaceutical treatments in the future in this area (214).

Testosterone

Testosterone is of value only in the treatment of secondary osteoporosis in hypogonadal men. Prostate-specific antigen and serum lipid status should be monitored during treatment.
Anabolic steroids may actually have a beneficial effect on bone mass; however, their side effects include liver toxicity, masculinization, and increased cholesterol levels, which prohibits their use in osteoporosis.

Cytokines
In theory, a number of cytokines may function as growth factors (transforming growth factor β, IGF-I, etc.) with potential benefit in osteoporosis. Their benefits as established by clinical trials, however, are not presently common in clinical practice, although IGF-I and DHEAS may be used to assess bone health in pediatric populations in the future (i.e., anorexia nervosa and cerebral palsy) (215).

Non–FDA-Approved Drugs
Current data suggest that high-dose sodium fluoride, which is a positive bone former, may actually worsen osteoporosis by increasing the risk of nonspinal fracture. Whether sodium fluoride in newer formulations, including a low-dose sustained-release preparation, will prove of benefit is unclear. Sodium fluoride must be viewed as an experimental therapy with some concerns regarding its overall benefit in osteoporosis (216–218).

Other Bisphosphonates
Etidronate, tiludronate, pamidronate, ibandronate, and zoledronate are currently FDA approved for the treatment of such conditions as Paget’s disease, hypercalcemia of malignancy, heterotopic ossification, and myositis ossificans. Etidronate is not as potent as the newer bisphosphonates in the treatment of osteoporosis and can also impair the laying down of mineral during new bone formation if given in high doses (219).

REHABILITATION MANAGEMENT
POST-FRACTURE
Osteoporosis is a silent disease that can progress from minimal impairment of the skeleton’s capacity to bear stress to a disease characterized by frailty, fracture, deformity, chronic pain, handicap, and loss of independence. Patients present with an individualized spectrum of complaints warranting different degrees of investigation and intervention (see Fig. 39-3). Rehabilitation management depends on accurate determination of the degree of bone loss, the risk factors for osteoporosis, the degree of frailty and propensity to fall, the capacity for participation in ADLs and safe exercise, and pain impairment level. All patients with chronic disability should be investigated for secondary causes of osteoporosis and comprehensive treatment initiated where warranted.

Radiologic Imaging
Imaging of symptomatic skeletal sites to evaluate for the presence of fractures and ascertain degree of associated deformity may be required. Plain radiographs often suffice to determine site of fracture, but if negative, MRI may be needed to detect first signs of inflammation associated with vertebral microfractures of individual trabeculae. This can determine acuity of compression fracture or occult hip fracture. MRI is also useful if neurologic involvement is suspected with new compression fracture. For example, severe vertebral collapse can precipitate foraminal narrowing and nerve impingement, or retropulsed bone fragments that may compromise spinal cord function. MRI can also identify avascular necrosis, disc herniations, and facet pathology post-fracture. When MRI is contraindicated or poorly tolerated, CT scan can be performed; it is particularly helpful in identifying metastatic disease, and fracture lines that are potential routes for cement extravasation with vertebroplasty. Skeletal scintigraphy can be also helpful in differentiating acute from chronic compression fracture sites, fracture pain from arthritic complaints, and bone malignancy.

Bone mass measurement with DXA is of value in determining the goals and intensity of the therapeutic exercise program post-fracture. However, if not available, the diagnosis of fragility fracture at the spine or hip also establishes the diagnosis of osteoporosis, to guide medical precautions during the post-fracture rehabilitation phase. When the fracture is healed, DXA can be obtained in most patients and can provide a baseline to monitor response to treatment over time. (See Fig. 39-3, Tables 39-5 and 39-8, and Clinical Evaluation, above, for appropriate screening and treatment guidelines.) Rehabilitation goals include pain reduction, improvement or maintenance of bone mass and muscle strength and flexibility, and establishment of medical precautions for safe exercise, fall reduction (see Table 39-7), and maximum recovery of independence. Strategies to achieve these goals should also address secondary causes for osteoporosis and any deformity, pain, or contractures, while improving core and peripheral muscle strength, balance, and gait.

The plan of care should include appropriate diagnostic tests and pain management strategies, prescriptions where appropriate for osteoporosis medications or supplements (see Tables 39-12 and 39-13), rehabilitation therapy, assistive devices and home modification to prepare the patient for maximum independence in exercise, fall reduction, and safe ADLs. Clearly define all medical precautions for patient and therapists, to prevent injury or medical complications during exercise and ADLs; recommend diet and lifestyle behaviors to optimize bone health (see Fig. 39-3).

Commonly, older patients present with many significant risk factors for low bone mass, and history of previous fractures, with acute or chronic disability. As discussed above, fractures can occur in virtually any bone, but often occur in one of three common sites: distal radius, proximal femur, and/or vertebra due to the biomechanical strategies of falling. The pain associated with these fractures is usually severe but self-limiting. The loss of function resulting from these fractures can be severe, affecting mobility and ADLs, and may lead to loss of independence and to subsequent institutionalization.

As stressed above, screening for calcium intake and vitamin D deficiency is very important in the elderly and in persons with disability, because of the prevalence of vitamin D deficiency in these populations, and the synergistic benefit of calcium and vitamin D in exercise and fall reduction programs. Weight-bearing and progressive resistance exercises targeted to
areas most commonly fractured can increase bone formation at those sites, with consideration of the general fragility of the skeleton individualized for each patient by therapists.

**Vertebral Fracture Rehabilitation Course**

Up to 33% of vertebral fractures are silent (see Table 39-2). Although fractures of the proximal femur and distal forearm are associated with significant pain, fractures of the vertebrae can be associated with minimal trauma, such as coughing or straining on the commode, and can be asymptomatic (139). Shen et al. in 2007 estimated that only 20% to 25% of vertebral compression fracture patients seek medical attention (220). Diagnostic procedures, most frequently the radionuclide bone scan or MRI, with STIR (short tau inversion recovery) imaging, may be used in evaluating the acuity of vertebral fractures noted on plain radiographs, particularly with active pain complaints at the site.

Vertebral fractures in the osteoporotic patient typically involve the anterior portion of the vertebral body, and occur most frequently at the thoracolumbar junction, T8 through L2 (221). This portion of the spine is made of predominantly cancellous bone (65% to 75% trabecular and 25% to 35% cortical). The midradius, by contrast, is about 95% cortical bone in content. In primary osteoporosis, one sees decreases in the density of trabecular bone by about 40% by 75 years of age (222).

Most patients with compression fractures present with acute or chronic back pain, and complain of sharp pain that increases with movement, particularly bed mobility and transfers, and is alleviated with rest. After acute fracture, patients can complain of pain with even basic ADLs like walking, combing hair or donning clothing. Severe and frequently disabling pain may persist for 2 to 3 weeks, but usually subsides by 6 to 8 weeks from time of fracture. It is unlikely that osteoporosis produces acute severe pain in the absence of a fracture. Spine pain can also be secondary to mechanical derangement of the spine such as kyphosis, especially when severe. Paraspinal muscle spasm, arthritis, nerve impingement or costal–iliac impingement syndrome, a painful rubbing of the rib cage against the iliac crest, can also be sources of pain (223). Signs and symptoms of vertebral fracture may be mimicked by neoplasm, herpes zoster, polymyalgia rheumatica, pancreatic disorders, and abdominal aortic aneurysm. It is important to investigate and treat the pain complaint promptly. If the patient experiences pain over prolonged periods, he or she may suffer consequences such as depression, sleep disturbance, and functional decline. The persistence of pain beyond 6 months at the site of a previous vertebral fracture may suggest causes other than original osteoporotic fracture; other etiologies of back pain in this group should be reconsidered, including progressed loss of height at original fracture site.

Plain film x-rays have poor sensitivity in diagnosis of osteoporotic sacral fractures (called sacral insufficiency fractures, or SIF) as well. There may be other associated fractures including rib fractures in addition to the sacral fracture in many cases. Therefore, a technetium-99 bone scan may be helpful in identifying such fractures, as well as their acuity. The classic “H” or “Honda” sign seen on bone scan, representing combined bilateral vertical and horizontal sacral fractures, is present inconsistently and may vary from 15% to 68% (224). MRI demonstrating bone marrow edema as low-signal intensity on T1 weighted, high-signal intensity on T2 weighted images and the T2 weighted STIR images are particularly sensitive (225). However, CT is regarded as the gold standard in diagnosing occult fractures.

The osteoporotic patient complaining of acute pain resulting from vertebral fracture should be managed initially with rest, immobilization of fracture site, and analgesic agents. Because vertebral fractures generally heal well, management is directed at pain control and providing adequate rest and immobilization of the fracture site (Table 39-14). The unwanted side effects associated with analgesic agents may complicate treatment. Common pharmacologic interventions for acute back pain are summarized in Table 39.14.

| TABLE 39.14 Rehabilitation Management of Back Pain in Patients with Vertebral Fracture |
| Acute back pain |
| Limit bed rest during the day; encourage good sleep and nutrition habits. |
| Recommend analgesics to facilitate optimal function; utilize opiates with caution. |
| Consider transdermal medicines to limit sedation (i.e., lidocaine or NSAID patch). |
| Prescribe medications for constipation if needed, if natural remedies fail. |
| Consider back brace (i.e., CASH brace, or rigid TLSO if risk for cord compression). |
| Monitor for signs of radiculopathy and spinal cord compression at site of fracture. |
| Teach proper bed positioning and mobility techniques, and emphasize spine-neutral principles during transfers, ADLs, and exercise. |
| Train caregivers to assist patients safely with minimal spine loading. |
| Prescribe appropriate ambulation assistive device where needed. |
| Coordinate physical therapy and occupational therapy services as appropriate. |

| Chronic back pain |
| Improve posture, transfers, and gait pattern to limit vertebral compression forces. |
| Consider postural support orthotic to decrease ligament stretch. Adjust analgesics, as pain warrants; establish opiate management contract with patient. |
| Prescribe a sound, on-going therapeutic exercise program. Consider vertebroplasty if conservative measures fail to improve pain profile. Evaluate and treat psychological and social consequences: consider relaxation techniques, biofeedback, support groups and self-management skills training. |

fractures include narcotics such as codeine, but in some cases transdermal applications (diclofenac, lidocaine), or oral tramadol may suffice. Within 3 to 4 weeks, weaning trials with other analgesic agents such as acetaminophen or nonsteroidal anti-inflammatory drugs (NSAIDs) should be attempted, in conjunction with other pain therapies. NSAIDs, however, must be used sparingly and with caution in the elderly. Primary pain management should incorporate rest, orthoses, and physical agents, with pharmacologic agents serving as adjunctive therapy. A program of progressive activity is indicated after brief initial bed rest. Because of mechanical forces translated into the spine with bed mobility and transfers, these movements are typically more painful than ambulation at the onset of rehabilitative phase post-fracture. The use of a sheepskin, egg crate, or gel flotation pad on the mattress frequently enhances patient comfort. A stool softener and laxative will help prevent straining with bowel movements. Use of a bedside commode may prove easier than a bed pan and requires less energy expenditure. Progressive transfer and ambulation training should be provided by the therapist, followed by a gentle progressive resistive exercise program for the limbs. Resistive exercises are unlikely to cause fracture in osteoporotic bones of the limbs, but can cause a new fracture or progression of acute compression fracture if resistive forces are translated into the spine. Slow introduction of isometric exercises for abdomen and back muscles is considered safe. Strict adherence to neutral spine positioning during exercise is recommended.

A fracture of the spine can result in 3-5 times increased risk of further spine fracture, and a 1.2 to 1.9 increased death rate (226). Clusters of vertebral fractures may occur in patients age 50 and older, in rapid progression. The cause of this most aggressive form of osteoporosis is unclear, but may be associated with an accelerated trabecular bone loss soon after menopause, malnutrition, immobility, and in some cases in adjacent vertebrae to previous sites of vertebroplasty (see below). After a number of fracture events in the spine, the collapsed and/or anteriorly wedged vertebrae may lead to deformity of the back, with subsequent kyphosis, loss of height, and chronic pain in the area of the thoracolumbar junction, secondary to mechanical deformity and paraspinous muscle spasm. This chronic back pain is typically of lesser intensity than the pain associated with the acute fracture event; it radiates laterally, is associated with exertion, and is relieved to a certain extent with rest. In addition, with progressive spinal deformity and height loss, an abdominal protruberance and resultant gastrointestinal discomfort (bloating and constipation) may occur, as well as some degree of pulmonary insufficiency secondary to thoracic cage deformity. In patients with multiple fractures, and severe spinal kyphosis, costal–iliac impingement syndrome may result.

Although chronic back pain is a common complaint of the elderly, the extent to which osteoporosis contributes to this pain is questionable. In a study involving 242 women 55 years of age, 30% had complaints of back pain, but there was no relationship between this pain and spinal curvature. In a group of older women (60 to 79 years of age), back pain affected a similar proportion (30%) but was twice as likely to occur in women with kyphosis or a loss in height exceeding 4 cm (227). There is not an absolute relationship, therefore, between osteoporosis and kyphosis. It is recognized that kyphosis may be secondary to chronic poor posture and age-related changes in muscles, ligaments, and intervertebral discs. Seventy percent of women over 60 years of age may demonstrate kyphosis without evidence of vertebral deformity. With aging, there is a progression of kyphosis (228), but back pain does not appear to be associated with the kyphosis unless vertebral deformity is such that there is a reduction in the height of the vertebral body greater than 4 SDs from normal (227).

Investigations into causes of back pain using bone scintigraphy showed a high incidence of facet joint disease in osteoporotic women with previous vertebral fracture, most prominently at the level of the vertebral collapse. Smaller lesions were commonly found in the facets above and below this level. Back pain with neurologic symptoms in the lower extremities may occur when vertebral fracture results in retropulsed fragments, or with associated foraminal narrowing, causing nerve impingement and radicular symptoms (229).

The pain of costal-iliac impingement syndrome is particularly difficult to treat. It is typically localized to the site of actual mechanical irritation, but can refer to the lower back and into the leg. Lateral bending and rotation can elicit this pain. Diagnosis is made by provocation of pain when palpating the lower ribs and the iliac crest contact site, and with lateral bending and rotation of spine. Injection of lidocaine into the margin of the iliac crest and lower ribs can be both diagnostic and therapeutic. Postural training, strengthening of abdominal and lumbar musculature, and the use of a wide, soft belt or CASH (cruciate anterior sternal hyperextension) brace have been reported as beneficial in relieving symptoms by lifting the ribs to avoid contact with the iliac crest. Concomitant trial with transdermal pain medications such as diclofenac or lidocaine should be considered, but narcotic medications may be needed. In severe cases, resection of the lower ribs has been beneficial (230).

In summary, the management of chronic back pain associated with osteoporotic vertebral fracture should include a program of strengthening paravertebral, abdominal, and gluteal muscles and safe programs to improve balance, flexibility, and posture. Relief of stress on the spine through use of proper body mechanics is encouraged. In severe cases, an orthosis can be of benefit. An assessment of ADLs may lead to the use of other techniques and devices that can help the patient avoid situations that aggravate pain. The strategic placement of a pillow or towel roll behind the back frequently increases sitting tolerance in patients with kyphosis. Physical agents such as heat, ice, transcutaneous electrical nerve stimulation, and acupuncture can be of benefit. Hypnosis, behavioral modification, biofeedback, and counseling have also been of benefit in the treatment of chronic pain.

**Back Supports and Bracing**

The degree and types of skeletal pain and disability among patients with osteoporosis present a complicated challenge.
to provide adequate mechanical support for the spine (231). When many solutions are available to answer a given problem, it usually indicates that there is no single good solution. Such is the case with mechanical supports for the osteoporotic spine. These orthoses may be used for pain relief and stabilization of the spine in both acute fracture and long-term care, and to promote healing and improved function. In the long-term treatment of the osteoporotic spine, orthoses may prevent further fracture.

When prescribing a brace or corset, one must understand the biomechanics of the spine, the types and causes of vertebral fractures, and the principles of bracing, including indications and hazards of individual orthoses (232). It is important to understand the functions of the lower thoracic and upper lumbar spine, where most compression fractures occur. The articulation of the thoracic vertebrae with the ribs, as well as the overlapping of the spinous processes, significantly limits its mobility in flexion and extension; rotation is relatively free, however. The lumbar spine has limited lateral flexion and axial rotation secondary to the relatively vertical orientation of the facet joints, so flexion and extension account for the majority of its movement. One must also understand the kinematic function of the lumbar spine, including compensation for increased functional mobility (233). While considering these functions of the spine, it is important to remember that movements that cause loading of the vertebral bodies increase the risk of fracture if bone density cannot support the resulting increase in applied force. Hence, bracing to help prevent this additional loading of the vertebral bodies must restrict flexion, which loads the anterior column of vertebrae. Restriction of flexion is therefore one goal of bracing, in conjunction with decreased pain, increased function, and prevention of soft-tissue shortening which may contribute to deformity.

There are several commonly used orthoses to stabilize osteoporotic vertebral fractures: postural training supports (PTS)-a weighted kypho-orthosis, thoracolumbar support such as the CASH brace, lumbosacral corset, and thoracolumbosacral orthosis (TLSO). All orthoses work on the principle of a three-point force system. Generally, the more rigid orthoses are used for acute thoracolumbar fractures, whereas the non-rigid orthoses such as the PTS and lumbosacral corset are used more commonly in the management of stable fractures and painful conditions. All orthoses described may not adequately prevent gravity-related axial compression, which may ultimately result in new fracture. Chronic use of spinal orthotics is generally discouraged because of the increased likelihood of weakening or atrophy of the trunk muscles, and reduced spinal mobility. Weakness of the supporting musculature may in time predispose to increased risk of vertebral fracture.

The PTS has been described as an inexpensive, unobtrusive device that promotes improvement in posture and decreases back pain by producing a force posteriorly below the inferior angles of the scapulae or by acting as a proprioceptive reinforcement (234). A TLSO such as the rigid clamshell brace is a long spinal orthosis that provides virtual fixation from the pelvis through the shoulders, and is commonly prescribed when vertebral fracture results in retropulsed fragments or severe spinal stenosis that could compromise the spinal cord. Also, the lumbar region of the spine is one of the most difficult areas of the body to immobilize, requiring more than a simple lumbar orthosis after retropulsion has occurred. Although the TLSO affords the greatest immobility, it is cumbersome and hot, and noncompliance with wearing schedule is very common. When neurologic compromise is not an imminent risk, a semi-rigid TLSO such as SpinoMed or CASH brace is often used after vertebral fracture. Custom-fitted rigid TLSOs are also more expensive than the PTS and CASH brace.

When prescribing a brace or corset, one must understand the biomechanics of the spine, the types and causes of vertebral fractures, and the principles of bracing, including indications and hazards of individual orthoses (232). It is important to understand the functions of the lower thoracic and upper lumbar spine, where most compression fractures occur. The articulation of the thoracic vertebrae with the ribs, as well as the overlapping of the spinous processes, significantly limits its mobility in flexion and extension; rotation is relatively free, however. The lumbar spine has limited lateral flexion and axial rotation secondary to the relatively vertical orientation of the facet joints, so flexion and extension account for the majority of its movement. One must also understand the kinematic function of the lumbar spine, including compensation for increased functional mobility (233). While considering these functions of the spine, it is important to remember that movements that cause loading of the vertebral bodies increase the risk of fracture if bone density cannot support the resulting increase in applied force. Hence, bracing to help prevent this additional loading of the vertebral bodies must restrict flexion, which loads the anterior column of vertebrae. Restriction of flexion is therefore one goal of bracing, in conjunction with decreased pain, increased function, and prevention of soft-tissue shortening which may contribute to deformity.

There are several commonly used orthoses to stabilize osteoporotic vertebral fractures: postural training supports (PTS)-a weighted kypho-orthosis, thoracolumbar support such as the CASH brace, lumbosacral corset, and thoracolumbosacral orthosis (TLSO). All orthoses work on the principle of a three-point force system. Generally, the more rigid orthoses are used for acute thoracolumbar fractures, whereas the non-rigid orthoses such as the PTS and lumbosacral corset are used more commonly in the management of stable fractures and painful conditions. All orthoses described may not adequately prevent gravity-related axial compression, which may ultimately result in new fracture. Chronic use of spinal orthotics is generally discouraged because of the increased likelihood of weakening or atrophy of the trunk muscles, and reduced spinal mobility. Weakness of the supporting musculature may in time predispose to increased risk of vertebral fracture.

The PTS has been described as an inexpensive, unobtrusive device that promotes improvement in posture and decreases back pain by producing a force posteriorly below the inferior angles of the scapulae or by acting as a proprioceptive reinforcement (234). A TLSO such as the rigid clamshell brace is a long spinal orthosis that provides virtual fixation from the pelvis through the shoulders, and is commonly prescribed when vertebral fracture results in retropulsed fragments or severe spinal stenosis that could compromise the spinal cord. Also, the lumbar region of the spine is one of the most difficult areas of the body to immobilize, requiring more than a simple lumbar orthosis after retropulsion has occurred. Although the TLSO affords the greatest immobility, it is cumbersome and hot, and noncompliance with wearing schedule is very common. When neurologic compromise is not an imminent risk, a semi-rigid TLSO such as SpinoMed or CASH brace is often used after vertebral fracture. Custom-fitted rigid TLSOs are also more expensive than the PTS and CASH brace.

The inexpensive abdominal corset has also been used to decrease pain, and increase function after an acutely painful vertebral fracture; it restricts movement via both mechanical and sensory feedback. This orthosis may also generate heat, pressure, or a massage-like effect that may be soothing for muscles in spasm. The use of corsets may also relieve pain by increasing hydrostatic support of the spine through increased intra-abdominal pressure, thereby placing an anteriorly directed force on the vertebral bodies. Again, this device can be hot, but it is not as bulky as a TLSO or CASH brace, can be worn under clothing, and has a higher compliance rate in patients with uncomplicated lumbar fracture.

Kaplan et al. conducted a pilot study to compare the effects of back supports on back strength in women age 40 and older with diagnosis of osteopenia or osteoporosis, who were randomly assigned to one of three groups: postural exercises (PE) alone, PE and conventional thoracolumbar support (CTLS), or PE and PTS. Compliance was poor among those who wore the TLS. The PTS and PE groups increased back strength significantly, implying that the more rigid CTLS inhibited strengthening of this area—a known complication of rigid bracing across any joint (235). Lynn and Sinaki found that kyphotic, osteoporotic patients had more postural sway and greater reliance on their hips to provide balance than did those with normal posture, who based balancing strategies in their ankles (236). Subsequently, they demonstrated that use of the PTS in these patients improved balance and decreased pain with only 1 month of PTS even in the absence of strong back muscles (237). The weighted PTS is thought to reduce pain by encouraging correct gravitational alignment and reduce strain to compensating muscles. In combination with back extensor muscle-strengthening exercises, the weighted PTS can contribute to the management of kyphotic pain.

In summary, rigid TLSO bracing is typically prescribed after vertebral compression fracture with retropulsed fragments and associated risk of neurologic impairment; noncompliance with wear is common, and associated risks must be discussed with patients. Other supports such as a non-rigid TLSO brace simple or a lumbosacral corset are better tolerated, and can be valuable after uncomplicated osteoporotic fracture. They reduce pain and facilitate early recovery of functional mobility and ADL function. The PTS can be useful in improving chronic pain and balance deficits in patients with kyphosis, but is not typically used after vertebral fracture. All spinal orthotics are thought to diminish vertebral fracture. All spinal orthotics are thought to diminish vertebral fracture.
Vertebroplasty and Kyphoplasty

Vertebroplasty and its derivative kyphoplasty are spinal procedures first introduced in 1984, and 1998, respectively, in which a radio-opaque bone cement, polymethylmethacrylate, is injected into vertebral fractures with an 8-13 gauge bone needle under fluoroscopy or CT guidance. Figure 39-5 shows the outcome of injection, in a L4 compression fracture after vertebroplasty. Kyphoplasty requires slightly more time, as space for cement is created first by inserting an inflatable balloon tamp prior to injection. The procedure can be in-patient or out-patient, in part dependent on medical status of the patient at time of procedure. These procedures are thought to stabilize the anterior column and any endplate fractures of the compressed vertebrae, in order to restore vertebral height, spinal alignment, and function to the greatest degree possible (238). There is growing evidence for efficacy in relieving unremitting pain from osteoporotic or pathologic vertebral compression fractures—up to 80% to 90% (239–243). However, well-controlled studies comparing these two procedures, or comparing either procedure with populations not receiving these interventions, are not yet available.

Although 700,000 vertebral compression fractures occur in the United States annually, one must clearly establish...
by thorough history, physical examination, and correlative radiologic study that the compression fracture site in question is indeed the locus of debilitating pain. This is especially important because the typical age group of patients that develop osteoporotic vertebral fractures is also prone to degenerative joint disease of the spine, and consequent back pain. Although the thoracolumbar vertebrae are most likely sites for intervention, vertebral augmentation can also be performed at the cervical spine. An evolving intervention, sacroplasty, attempts to replicate the success of vertebral augmentation procedures for SIF (244).

Absolute contraindications for these procedures include uncorrected coagulopathy, preexisting systemic or spine infection, and on-going neurologic deficits secondary to the fracture site. Reported complications include extravasation of cement, bleeding, infection and neural injury, fractures, cement embolism, and predisposition to subsequent vertebral fractures at fragile adjacent vertebrae (12% to 52%) (245). There have been no comprehensive study reviews to date that have clarified this latter issue. Similarly, research is needed to evaluate whether the addition of back ergonomics and other post-compression fracture rehabilitation protocols such as postural and mobility training and back ergonomics 8 would improve functional outcomes with these procedures. Treating traumatic fractures in young persons with normal bone density is not considered appropriate with either technique (246).

A multidisciplinary team approach is beneficial to ensuring maintenance of function in this population. Nonpharmacologic interventions should be preferentially used to manage chronic back pain. This program should be supplemented by encouraging adjustments in lifestyle, medication use, physical agents, orthoses, and other therapies considered useful for chronic pain. These interventions are used only after ruling out other causes of back pain in the elderly, and after assessing the degree to which depression is contributing to the symptoms.

**Hip Fracture**

Hip fractures may be divided into three categories according to the anatomic area in which they occur. Intracapsular fractures (femoral neck fractures) are located distal to the femoral head but proximal to the greater and lesser trochanters. These fractures frequently disrupt the blood supply to the femoral head and are therefore associated with nonunion and osteonecrosis of the femoral head, and typically require hemiarthroplasty for stabilization of the joint (247). Fractures occurring between the greater and the lesser trochanter are not associated with the complications seen in the intracapsular region. However, they are associated with malunion and shortening of the leg as a result of osteoporotic bone and the deforming forces exerted on this area of the proximal femur. Intertrochanteric and femoral neck fractures occur with equal frequency and together account for 90% of hip fractures. Subtrochanteric fractures occur just below the lesser trochanter and are responsible for only 5% to 10% of all hip fractures. Because 90% of peripheral osteoporotic fractures result from a fall, radiographs should also exclude fracture of the pubic ramus, acetabulum, and greater trochanter. Pain in this area after fall can also be associated with trochanteric bursitis and sacroiliac dysfunction. In all patients, the goal of treatment is to return the patient, pain-free, to maximum level of mobility as soon as possible, address fall risks, and implement appropriate screening and treatment for underlying bone disease.

Rehabilitation begins on the first day after surgery with a progressive ambulation program. Special precautions to prevent deep venous thrombosis with anticoagulant medication such as enoxaparin or coumadin, and early mobilization are routine. Most patients begin walking on the first or second post-operative day, with weight-bearing status established by the orthopedic surgeon. When long-term stability of the surgical fixation is in question, weight-bearing status is decreased to minimize the possibility of hardware failure. Duke and Keating showed that mobility on day 2 following surgery was a significant predictor of independence at 2 weeks; independent mobility was defined as the ability to walk at least 15 meters with a walker and transfer to and from bed independently (248).

In a study involving a large patient population aged 65 years or older, Cree et al. found that almost all patients with cognitive impairment who had sustained a hip fracture were functionally dependent 3 months after the fracture and had acquired a new disability in transfers between chair and bed (249). Most of those who functioned independently prior to fracture functioned independently afterward, with post-fracture assistance most often needed in bathing and dressing. In those patients who were considered to be of high mental status, post-fracture dependence correlated with advanced age, a greater number of comorbidities, hip pain, previous employment in a high prestige occupation, and poorer self-related health (250). Magaziner et al. found little change in the areas of greatest disability in the respective 12- and 24-month follow-ups to hip fracture in previously independent patients: climbing five stairs (90% and 91%), toilet transfers (66% and 63%), and tub and shower transfers (83% at both intervals) (251).

The proportion of patients discharged to home after hip fracture ranges from 40% to 90%, although many patients remain institutionalized (252,253). Factors associated with permanent institutionalization are assistance with ADLs, age greater than 80 years, lack of involvement of family members, incontinence, and insufficient physical therapy at a skilled nursing facility (252). In a hospital-based comprehensive rehabilitation unit, more than 90% of these patients are discharged home. Although this high percentage has clearly been influenced by the selection process, the scope and intensity of rehabilitation services and the ability to manage acute exacerbation of comorbidities in close geographic relationship with an acute medical-surgical hospital may promote higher function. Discharge home is positively associated with presence of another person in the home, ability to walk independently before the fracture, ability to perform ADLs (252,253), absence of preexisting dementia, younger age, and preexisting social network (254).

The risk of death after a hip fracture is increased by a factor of 2.8 within the first 3 months after fracture, and is more likely in those in poor health (255), with mortality rate after
Wrist Fracture

Fractures of the wrist are the most common type of fractures in women less than 75 years of age. A prior wrist fracture doubles the risk of any future osteoporotic fracture in postmenopausal women and triples the risk for a second wrist fracture (260). These fractures increase in number after menopause, and although they usually occur in relatively healthy active women, they may be the first sign of an underlying problem such as low bone mass. The primary goal of treatment is return of pain-free normal function of the hand and wrist. Initial casting does not usually extend above the elbow. During the period of immobilization, usually 6 to 8 weeks, strength and flexibility should be maintained in the upper extremities. Active and passive range-of-motion exercises should be provided for the fingers and shoulder on the affected side, with close monitoring for signs and symptoms of median nerve impingement (carpal tunnel syndrome) secondary to cast impingement with post-fracture edema. These exercises should be continued after cast removal, with wrist, forearm, and elbow range of motion and exercise added. At that time, a local wrist splint may be used to support and protect the wrist. As a result of the wrist fracture, particularly on the dominant side, the patient may require assistance with ADLs such as getting dressed, combing hair, and brushing teeth.

MacDermid et al, reported the rate of improvement to normal pain-free function after distal radius fractures to be highest in the first 6 months after the fracture, with gains generally leveling off during the next 6 months (261). Of note, even with well-healed fractures, pain may persist because of ligament or triangular fibrocartilage complex (TFCC) injuries, which are common and can often be missed in the initial work-up for a wrist fracture (262–264).

FEMALE ATHLETE TRIAD

With the advent of Title IX in the early 1970s, high school age girls’ participation in organized sports has increased 1000% (265). With increased energy expenditures as training becomes more intensive, caloric intake must be increased accordingly. Carbohydrates should compose approximately 60% of daily caloric intake by American Dietetic Association guidelines, with fats comprising 25% to 30%, and proteins 15%. Carbohydrates are therefore the primary source of muscle energy, and adequate total caloric intake prevents sacrifice of body fat and proteins. Although vitamins, minerals, and water also are essential for health, they do not provide energy.

With increased energy expenditures as training becomes more intensive, caloric intake must be increased accordingly. Not increasing caloric intake after increased exercise can precipitate weight loss, and sacrifice of adipose tissue for training needs. This can deplete estrogen stores sufficiently to cause amenorrhea in female athletes. This increases the risk of osteoporosis and stress fractures, especially in conjunction with eating disorders, such as is seen in FAT. Eating disorders have been reported as high as 62% in college female athletes, and are becoming more common in adolescents. Although secondary amenorrhea in the general population is approximately 5%, it can increase to 10% to 20% in college athletes and 50% in elite athletes (266). Delayed menarche or absent menses for more than 3 months, particularly during times of increased training, should prompt clinical evaluation, including history of intentional weight loss strategies (i.e., diuretics, bulimia or anorexia, laxatives, avoidance of “fear foods”), distorted body image, and refractory pain sites. Evaluation with nutritionist is recommended, with psychologist or psychiatrist if warranted, if eating disorder is suspected. In-patient treatment should be considered in patients refractory to treatment efforts. Monitoring for bulimia in light- and middle-weight category male wrestlers, especially prior to competition, is also needed. Although these athletes are not as likely to fracture, there is concern they will be at increased risk of fracture as they age, if bulimic behavior coexists with the bone formation phase of adolescence.

In addition to basic osteoporosis laboratory tests (see Table 39-8) hormonal evaluation (pregnancy test, Thyroid Stimulating Hormone (TSH), Follicular Stimulating Hormone (FSH), and prolactin) is required in these patients, including estradiol level checked between days 2 and 4 of expected menstrual cycle. Any deficiencies must be addressed clinically. During childhood and adolescence, it is normal to see ongoing bone growth at bone metaphyses by plain radiographs,
and corresponding elevation of bone turnover markers such as urine or serum NTX. This precludes their use to monitor treatment efficacy in this population. Although IGF-I and DHEAS are useful markers in FAT research, at present they cannot be used clinically in young patients at high risk, such as is seen in anorexia nervosa or cerebral palsy (267). Similarly, applicability of DXA T-scores for bone mass assessment is not recommended in pediatric populations; Z-score monitoring is under investigation and will likely prove useful clinically in the future.

With FAT, stress fractures usually occur at the tibia or hip. If plain radiographs are negative, MRI is needed, as it can differentiate best between fracture and soft-tissue injury, such as anterior tibialis muscle tear. Calcium and vitamin D supplementation, oral contraceptive therapy, and a minimum 6 weeks non-weight bearing in the fractured extremity is the standard of care in this population. Bed rest in hospitalized anorectic patients should be used with caution, so as to limit further bone loss. Resistive exercise regimens for the other limbs are recommended during the bone-healing phase, provided the patient follows strict nutritional guidelines for caloric intake for energy expenditures. Return to regular exercise and competitive sports requires close monitoring of caloric intake and weight, with nutritional and mental health counseling if needed. Although oral contraceptives can be used to regulate menses and improve bone mass in these patients, irreversible osteoporosis is common.

In summary, physical medicine and rehabilitation expertise can be of great value in the management of patients with osteoporosis and is currently underused. These strategies are used to reduce disability resulting from impairments in bone mass and structure, muscle strength, and coordination. With a timely, focused, coordinated plan of care, preventively or after fracture, disability can be minimized.

**ELECTRONIC MEDICAL RECORD SYSTEMS COORDINATION**

More comprehensive osteoporosis screening and treatment will require expanded utilization of existing medical and electronic resources to improve communication between physicians and their patients. Multiple studies in Europe and America have shown that electronic medical record utilization, in strict compliance with patient privacy guidelines, has had good-to-excellent results in improving management of low bone mass (268–271). The best results to date in the United States have been in private health maintenance organizations. Newman et al. reported in 2003 that EMR targeting within a small, rural Pennsylvania Health Maintenance Organization of at-risk women 55 and older for DXA screening and osteoporosis treatment resulted in reduced hip fractures over a 5-year period in women 65 and older, with $1.5 million savings per annum over baseline in direct costs of care (272). Later EMR targeting of those patients who required chronic glucocorticoid medication resulted in 96% compliance with calcium, vitamin D, and prescription medication where indicated, significant improvement in bone density at hip and spine, and improved exercise frequency and vitamin D levels (273). A comprehensive screening, education and treatment program in California for patients age 65 or older at risk for osteoporosis reduced hip fractures by 37%, saving a larger private health plan approximately $30 million over 5 years (274). Studies have also demonstrated the benefit of a nurse coordinator or manager to improve appropriate treatment and decrease hip fractures (275) and provide patient education (276).

**SUMMARY**

Osteoporosis is a disease that is defined as both an intermediate outcome (BMD <2.5 SD from young adult baseline) and as a metabolic disease that can facilitate a fragility fracture of the spine or hip from a fall from a standing height. In this light, low BMD can be considered a precipitating factor for the latter, just as uncontrolled hypertension can lead to stroke. Research has shown that the negative outcomes of osteoporosis, like other disease entities, do respond to traditional public health strategies to improve clinical, radiologic, and laboratory screening and the success of expanding treatment options. Recent evidence-based research has better defined risks and treatment strategies for female and male patients of all ages, and specific ethnic groups in the last decade. Further research is needed to expand our knowledge of causes of secondary osteoporosis in males, and optimal screening and treatment strategies (277). Similarly, further investigation into the genetic basis for osteoporosis, optimal bone-building strategies for pediatric and adult populations, including persons with disability, and the role of exercise in preventing falls and optimizing BMD is needed.

The decision to screen for low BMD should be based on an individual’s risk factors and on treatments being considered. It is important to ensure that adults receive up to 1,200 mg of calcium per day in diet and supplements, and 800 to 1,000 IU of vitamin D daily, provided there are no medical contraindications to this regimen. Protein intake is also critical, with supplementation, particularly after hip fracture, considered for those at risk. Education of patients regarding secondary causes of osteoporosis and lifestyle choices, such as smoking, and the need to engage responsibly with health care professionals in an active preventive and treatment program is critical to reducing the disabling consequences of this disease. Regular weight-bearing exercise, particularly those aimed at improved balance, can help prevent osteoporotic fractures through osteogenic effects and prevention of falls. Many medications, including bisphosphonates and hormone replacement therapy, have shown success in decreasing bone resorption, to reduce fracture risks for those with low BMD (osteopenia) and osteoporosis. Teriparatide, which induces bone formation, and denosumab, which slows resorption via inhibition of RANKL, also help to prevent progression of osteoporosis and fractures.

Physical medicine and rehabilitation strategies can be of value in the management of patients with osteoporosis, but are
currently underused. There is increasing evidence that expanding traditional rehabilitation interventions (i.e., pain management, posture and exercise programs, and fall reduction strategies such as proper footwear, and appropriate environmental adaptations) in combination with adjustment of medications contributing to polypharmacy—particularly sedatives, antidepressants, diuretics, and antihypertensive agents—can significantly reduce the risk of falling, which in turn reduces the incidence of all fractures. The benefits of assistive devices such as canes, walkers, and wheelchairs for those with disturbed balance or deficits in gait should also be emphasized as frailty develops with aging and disability. Interventions such as vertebroplasty are becoming more prevalent for stabilizing spine integrity after compression fracture, but require further controlled studies for long-term benefit and efficacy. Targeting persons with chronic disability, who are at increased risk of bone loss compared to their more mobile peers, is particularly needed.

By utilizing both medical and traditional rehabilitation strategies, physiatrists can play a unique role in the interdisciplinary model to optimize screening and treatment efforts within the medical community. The implementation and coordination of these efforts in all in-patient rehabilitation settings, particularly during the seminal post-fracture period, could prevent further fracture, and reduce disability resulting from impairments in bone mass and structure, muscle strength, and coordination. To advocate for our patients in order to increase the likelihood of independent community dwelling, and to improve the quality of life for those in extended care facilities who have double the risk of falls, is at the core of our mission as physiatrists.

REFERENCES


Arthritic diseases are having an increasing impact on public health (1). It is estimated that musculoskeletal disorders are the most frequent cause of disability in developed countries (2). Currently, 46 million people are affected by arthritis, and this number is predicted to rise to 67 million by 2030. These individuals often sustain impairments, functional losses, and disability. The annual cost in medical care and lost wages is $128 billion (3). Twenty-first century physicians who care for these patients must be aware of current pharmacologic and rehabilitation strategies to control disease, limit impairment, preserve or enhance function, and reduce disability (1).

New trends have been developing in the management of inflammatory arthritis: (a) early use of disease-modifying anti-rheumatic drugs (DMARDs) in combination with reduction in disability (4), (b) a transition from passive patient involvement in treatment to self-efficacy, (c) a change from prescribing rest, to one of advocating relative rest and activity, (d) an increased use of rehabilitation early in the disease course, (e) a greater use of complementary and alternative therapies by patients and health care providers, and (f) an increased interest in which rehabilitation measures are available and are efficacious in achieving preventative and restorative goals for arthritis patients.

The specialty of rheumatology is dedicated to the understanding and control of disease activity. Physical Medicine and Rehabilitation has its roots in maintenance and restoration of function as well as prevention of dysfunction. This is achieved by having a keen knowledge of the impairments associated with the various rheumatic diseases (RDs) and the formulation of individualized treatment plans to maximize patient function by utilizing education, physical modalities and techniques, exercise, assistive and adaptive devices, energy conservation, joint protection, and vocational planning. The care of arthritis patients takes a multidisciplinary team effort. It is currently accepted that early diagnosis, medical treatment, and rehabilitation are preferable. Orthopedic interventions are often necessary and need to be appropriately timed.

This chapter addresses inflammatory diseases and their impact on the person as a whole, on specific joints, and on involved organ systems. Appropriate rehabilitation strategies are presented.

Inflammatory arthritis may involve a single joint or multiple joints, periarticular structures, and other organ systems. It may be an acute process that completely resolves (e.g., septic joint), or may be a chronic process (e.g., rheumatoid arthritis [RA]). Inflammatory arthritis may involve all the joint structures: synovium, cartilage, tendons, capsule, bone, and surrounding muscle. It is often part of a systemic RD (i.e., connective-tissue disease) such as RA, juvenile idiopathic arthritis (JIA) (formerly known as juvenile rheumatoid arthritis [JRA]), systemic lupus erythematosus (SLE), dermatomyositis/polymyositis (DM-PM), progressive systemic sclerosis (PSS), or mixed connective-tissue disease (MCTD). These diseases are usually chronic, remitting, and relapsing; are variable in their course; and affect multiple organ systems in addition to joints. They most often require long-term treatments with pharmacological agents, which may have a significant impact on appearance, sleep, psychological function, and reproductive ability. The physiatrist directing the rehabilitation team must take these issues into account when evaluating and devising a treatment plan.

This chapter emphasizes the importance of a good initial evaluation, early treatment intervention, and periodic reevaluation of the arthritis patient. It describes the rational and stage-specific manner in which rehabilitative treatments are applied to the problems of patients with RDs and the current scientific evidence, or in the absence thereof, the seasoned clinical judgment for their use. Equally important is to stress the critical need of engaging the patient in his or her own treatment (5).

ARTHRITIC DISEASES

Classification

The important determinants in classifying arthritis are whether the disorder is inflammatory versus noninflammatory, symmetric versus asymmetric, or accompanied by systemic and extra-articular manifestations (Fig. 40-1). A detailed history and physical examination, and appropriate laboratory and x-ray studies will often allow a specific diagnosis to be made. Clinical features that suggest inflammatory rather than noninflammatory disease include: acute painful onset, fever, erythema...
of the skin over the joint or joints involved, warmth of the joint or joints, and tenderness that usually parallel the degree of inflammation. Laboratory and x-ray findings that suggest an inflammatory process include an increased peripheral white blood cell count with left shift, an elevated erythrocyte sedimentation rate (ESR), a group II joint fluid (Table 40-1), and x-ray demonstration of soft-tissue swelling, periostitis, bony erosions, or uniform cartilage loss (Table 40-2).

### TABLE 40.1 Synovial Fluid Analysis

<table>
<thead>
<tr>
<th>Fluid Group</th>
<th>Color</th>
<th>Clarity</th>
<th>Viscosity</th>
<th>Mucin Clot</th>
<th>Cells/mm³</th>
<th>Percentage of White Blood Cells That Are Polymorphonuclear Leukocytes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>Pale yellow</td>
<td>Transparent</td>
<td>High</td>
<td>Good</td>
<td>&lt;25%</td>
<td>&lt;10%</td>
</tr>
<tr>
<td>Group I (noninflammatory)</td>
<td>Yellow or straw</td>
<td>Transparent</td>
<td>High</td>
<td>Good</td>
<td>&lt;2,000</td>
<td>&lt;25%</td>
</tr>
<tr>
<td>Group II (moderately inflammatory)</td>
<td>Yellow or straw</td>
<td>Transparent to opaque, slightly cloudy</td>
<td>Variably decreased</td>
<td>Fair to poor</td>
<td>3,000–50,000</td>
<td>&gt;70%</td>
</tr>
<tr>
<td>Group III (highly inflammatory, septic)</td>
<td>Variable; yellow-gray, purulent</td>
<td>Opaque, cloudy</td>
<td>Low</td>
<td>Poor</td>
<td>50,000–100,000 (usually 100,000 or more)</td>
<td>&gt;75%, usually close to 100%</td>
</tr>
<tr>
<td>Group IV (hemorrhagic)</td>
<td>Red</td>
<td>Opaque</td>
<td>High</td>
<td>Good</td>
<td>Up to normal count in blood</td>
<td>May be the same as normal blood</td>
</tr>
</tbody>
</table>
Inflammatory arthritis falls into four different groups and may be monoarticular or polyarticular (6,7):

1. Inflammatory connective-tissue disease (e.g., RA, JIA, SLE, PSS, DM–PM, MCTD, psoriatic arthritis [PSA])
2. Inflammatory crystal-induced disease (e.g., gout, pseudogout, basic calcium phosphate)
3. Inflammation induced by infectious agents (e.g., bacterial, viral, spirochete, tuberculosis, and fungal arthritis)
4. Seronegative spondyloarthopathies (e.g., ankylosing spondylitis [AS], PSA, Reiter’s syndrome [RS], inflammatory bowel disease [IBD]).

Noninflammatory arthritis may be classified as:

1. Degenerative, posttraumatic, or overuse (e.g., osteoarthritis [OA], posttraumatic aseptic necrosis [AN])
2. Inherited or metabolic (e.g., lipid storage disease, hemochromatosis, ochronosis, hypogammaglobulinemia, hemoglobinopathies).

Criteria for classification may be found in the Primer of the Rheumatic Diseases, 13th edition (8). A number of these diseases have systemic manifestations (Table 40–3), many of which should be addressed in addition to treatment of the arthritis itself. The physiatrist must be aware of the impact of a chronic, unpredictable illness on various life stages. For example, systemic disease may influence a young mother quite differently from a postmenopausal woman. In some diseases (e.g., RA, SLE, JIA, gout, AS, PSS, DM–PM), a number of set criteria delineated by the American College of Rheumatology (ACR) must be fulfilled before a definite or probable diagnosis can be made.
TABLE 40.3 Systemic Manifestations of RDs

<table>
<thead>
<tr>
<th>System</th>
<th>Disease</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skín</td>
<td>Juvenile idiopathic arthritis</td>
</tr>
<tr>
<td></td>
<td>Psoriatic arthritis</td>
</tr>
<tr>
<td></td>
<td>Reiter’s syndrome</td>
</tr>
<tr>
<td></td>
<td>Colitc arthritis</td>
</tr>
<tr>
<td></td>
<td>Sarcoid arthritis</td>
</tr>
<tr>
<td></td>
<td>Septic arthritis (especially <em>Neisseria gonorrhoeae</em> and meningitis)</td>
</tr>
<tr>
<td></td>
<td>Hyperlipoproteinemia</td>
</tr>
<tr>
<td></td>
<td>Systemic lupus erythematosus</td>
</tr>
<tr>
<td></td>
<td>Amyloidosis</td>
</tr>
<tr>
<td></td>
<td>Dermatomyositis</td>
</tr>
<tr>
<td>Nasopharynx and ear</td>
<td>Reiter’s syndrome</td>
</tr>
<tr>
<td></td>
<td>Rheumatoid arthritis</td>
</tr>
<tr>
<td>Eye</td>
<td>Juvenile idiopathic arthritis</td>
</tr>
<tr>
<td></td>
<td>Reiter’s syndrome</td>
</tr>
<tr>
<td></td>
<td>Rheumatoid arthritis</td>
</tr>
<tr>
<td></td>
<td>Sarcoid arthritis</td>
</tr>
<tr>
<td>GI tract</td>
<td>Colitc arthritis</td>
</tr>
<tr>
<td></td>
<td>Scleroderma</td>
</tr>
<tr>
<td></td>
<td>Progressive systemic sclerosis</td>
</tr>
<tr>
<td>Heart and circulation</td>
<td>Amyloidosis</td>
</tr>
<tr>
<td></td>
<td>Polymyositis</td>
</tr>
<tr>
<td></td>
<td>Juvenile idiopathic arthritis</td>
</tr>
<tr>
<td></td>
<td>Reiter’s syndrome</td>
</tr>
<tr>
<td></td>
<td>Ankylosing spondylitis</td>
</tr>
<tr>
<td>Respiratory tract</td>
<td>Sarcoidosis</td>
</tr>
<tr>
<td></td>
<td>Polymyositis</td>
</tr>
<tr>
<td></td>
<td>Rheumatoid arthritis</td>
</tr>
<tr>
<td>Nervous system</td>
<td>Systemic lupus erythematosus</td>
</tr>
<tr>
<td></td>
<td>Rheumatoid arthritis</td>
</tr>
<tr>
<td>Renal system</td>
<td>Amyloidosis</td>
</tr>
<tr>
<td></td>
<td>Gout</td>
</tr>
<tr>
<td></td>
<td>Systemic lupus erythematosus</td>
</tr>
<tr>
<td></td>
<td>Progressive systemic sclerosis</td>
</tr>
<tr>
<td>Hematologic system</td>
<td>Rheumatoid arthritis</td>
</tr>
<tr>
<td></td>
<td>Systemic lupus erythematosus</td>
</tr>
</tbody>
</table>

Demographics

Many types of arthritis have a specific distribution in terms of age, gender, race, and geographic appearance. Severity of disease may vary with age and gender. Genetics and occupation may also be influencing factors. It is helpful to be familiar with those portions of the population that are more susceptible to certain diseases (Table 40-4) (9).

Etiology and Pathophysiology

Because the designation RD includes such a broad spectrum of processes and syndromes, a classification system (see Fig. 40-1) that groups arthritides to some extent by etiology may be useful.

Rheumatoid Arthritis

RA is the most common of the inflammatory arthropathies and is often difficult to diagnose in its early stages. The etiology of RA remains unknown, although much has been learned in the past two decades about the inflammatory process, its relationship to the immune system, and molecular genetic regulation (10). Of the two hypotheses in vogue, one suggests that RA is an autoimmune disorder; the other proposes that specific external agents initiate the response, which then is perpetuated or amplified by the immune host response. Data in support of the first hypothesis are derived from the fact that antibodies against autologous immunoglobulin G are present in many patients with RA, which may represent a primary abnormality in the regulation of cells that control immunoglobulin synthesis. Rheumatoid factor (RF) has been associated with more severe disease and almost exclusively with extra-articular disease. This primary defect may alter the control mechanisms, so that stimulation and control of these events are unbalanced and the response to endogenous immune products goes awry.

A more likely explanation for the etiology of RA is that specific external agents initiate an inflammatory response, and in the susceptible host, the inflammation leads to continual disease activity. Infectious agents can cause synovitis. Some replicate in the joint space (e.g., *Mycobacterium*, *Staphylococcus*), and some enter the joint space and cause synovitis initiating a local immune response (e.g., rubella, spirochete). Lyme disease is an example of a spirochete-initiated disease that can be associated with chronic arthritis. Another type of arthritis follows gastrointestinal (GI) disease (e.g., *Shigella*, *Salmonella*, *Yersinia*). No organism is recovered from the joint, although a reactive arthritis occurs, and the inflammatory process is initiated by a remote infection. All efforts to associate an infectious agent with RA have failed, despite sophisticated electron microscopy and molecular biology techniques. Exposure to cigarette smoke is an additional environmental factor believed to play a role in increasing the risk for development of RA (11,12). Several components need to be acknowledged in the understanding of this process: an inciting agent, most likely exogenous and possibly a wide range of antigens; a genetic susceptibility; and an abnormality in the host immune response.

The mechanism of tissue injury in RA has been demonstrated to include the following components of the immune system and its associated mediators of inflammation. In the affected host, a stimulus initiates an inflammatory response directed against self or nonself, which sets into motion complement, leukocyte phagocytosis, lysosomal enzyme release, and several small mediators that may initiate clotting and fibrolysis. In the joint, the local reactions are helper T-cell (in particular CD4 T cells) mediators that are attracted to macrophages and dendritic cells. Antibody synthesis is initiated, thus perpetuating the immunologic activities already begun. Cytokines play a key role in the perpetuation of synovial inflammation. Many have been found in the RA synovium (11,13). These cytokines are targets for clinical drug trial interventions. Some of the joint mononuclear cells can produce proteinases, prostaglandins (PGs), and other small mediators of inflammation.
Tissue inflammation causes the synovial membrane to become hyperplastic and neoangiogenesis occur, with the production of chemokines that increase the influx of more inflammatory cells. Synovial fluid enzymes directly affect the articular cartilage. Bone erosions develop when the synovial membrane has invaded the cartilage. Production of metalloproteinases, fibroblasts, and mononuclear phagocytes produced by the synovium is controlled by cytokines (IL1, TNF-α, and TGF-β). These cytokines influence chondrocytes to produce less collagen and proteoglycan, and increase collagenase synthesis, which degrades type II collagen (14). As the high-intensity inflammation subsides in the joint, repair takes place, often with the proliferation of fibroblasts and scar tissue. Although it is unclear what triggers this, once the process is in place it often continues for a longer period than would be expected to successfully clear the antigen. Hence, the host immunoregulatory system, which is genetically controlled, must be abnormal. In addition, people with RA often have systemic manifestations with other organ damage mediated by immunologic processes. An example of immune abnormality in RA is the inappropriate reduction in thymic function (15).

Systemic Lupus Erythematosus

SLE is a multisystemic disease that is associated with abnormalities of immune regulation and immune complex-mediated tissue injury. It has been called a classic autoimmune disease, as a result of an abundance of autoantibodies generated against cytoplasmic and nuclear cellular components. The hallmark of these is generation of IgG antibodies to double stranded DNA. The etiology of SLE is obscure, but viral inclusion bodies have been implicated because of electron microscopic observations made in lymphocytes and vessel walls. Virus has never been isolated in patients with SLE; even those diseases with documented infectious etiologies are often multifactorial. Family members with SLE are more likely to have immunologic abnormalities than are controls. Hormonal influences are important in the expression of SLE, and women in the child-bearing years appear to be at greater risk. Women taking progestation-based oral contraceptives are at higher risk than those taking estrogen-based oral contraceptives. The pathogenesis of SLE depends on abnormalities in humoral and cellular immunity. Lymphopenia is common and is inversely related to disease activity. B-lymphocytes are normal in number but are hyperactive. T-lymphocytes are often decreased, most markedly in the T-suppressor lymphocyte subpopulation. Natural killer cell activity is diminished, but there are an increased number of lymphocytotoxic antibodies (16).

Progressive Systemic Sclerosis

PSS is a progressive disorder in which microvascular oblitative lesions in multiple organs terminate in fibrosis and atrophy. The hallmark of this disease is induration of the skin. Patients with PSS have capillary abnormalities and small artery lesions that appear late with organ involvement. The pathogenesis of organ involvement is most likely due to injury to the endothelial cell lining of vessels. Disturbing the lining activates the clotting system, with the release of vasoactive peptides. These factors stimulate smooth muscle cells to migrate in, proliferate, and deposit connective tissue, which results in the proliferative vascular lesions of PSS. The etiologic agent is obscure, and no strong hypotheses exist as to its nature (17).

Idiopathic Inflammatory Myopathies

Polymyositis (PM), dermatomyositis (DM), and inclusion body myositis (IBM) are part of a heterogeneous group of diseases characterized by inflammation of muscle and skin, often associated with profound weakness of all striated muscle, including the heart, and elevated levels of skeletal muscle enzymes. There are six types, each of which may have a different etiology (18);

- **Group 1**: Primary idiopathic PM
- **Group 2**: Primary idiopathic DM
- **Group 3**: DM-PM associated with neoplasia
- **Group 4**: DM-PM associated with vasculitis (juvenile dermatomyositis [JDM])
- **Group 5**: DM-PM associated with collagen vascular disease
- **Group 6**: IBM

### TABLE 40.4 Epidemiology of Inflammatory Arthritis

<table>
<thead>
<tr>
<th>Disease</th>
<th>Incidence (per 100,000)</th>
<th>Prevalence (per 1,000)</th>
<th>Peak Visit (years)</th>
<th>Gender</th>
<th>Frequency by Race</th>
</tr>
</thead>
<tbody>
<tr>
<td>RA</td>
<td>32.7</td>
<td>10 (1%)</td>
<td>25–50</td>
<td>2.5:1 female</td>
<td>Higher in whites and Native Americans (Pima Indians)</td>
</tr>
<tr>
<td>JIA</td>
<td>3.5–13.9</td>
<td>1–2</td>
<td>1–3</td>
<td>2:1 female</td>
<td>Lower in African Americans, Asian and Japanese</td>
</tr>
<tr>
<td>SLE</td>
<td>2–8</td>
<td>0.5–1</td>
<td>15–40</td>
<td>9:1 female</td>
<td>Three times higher in blacks, Chinese; increased in Haida, North American Indian</td>
</tr>
<tr>
<td>AS</td>
<td>7.3</td>
<td>1.5</td>
<td>25–44</td>
<td>3:1 male</td>
<td>Higher in Central European</td>
</tr>
<tr>
<td>PM</td>
<td>1.0</td>
<td>0.1</td>
<td></td>
<td>2:5:1 female</td>
<td>Three times higher in black females</td>
</tr>
<tr>
<td>PSS</td>
<td>0.9–1.9</td>
<td>0.29</td>
<td>30–50</td>
<td>4:1 female</td>
<td>Increased in Southern U.S. higher among African Americans</td>
</tr>
<tr>
<td>Gout</td>
<td>&gt;120</td>
<td>27.5</td>
<td>45–65</td>
<td>10:1 male</td>
<td>African Americans, Asian-Pacific Islanders</td>
</tr>
</tbody>
</table>

RA, rheumatoid arthritis; AS, ankylosing spondylitis; JIA, juvenile idiopathic arthritis; PM, polymyositis; SLE, systemic lupus erythematosus; PSS, progressive systemic sclerosis.
Two leading hypotheses may explain the etiology of DM-PM: viral infection and abnormal recognition of self. Immunoglobulins have been demonstrated in vessel walls, especially in intramuscular blood vessels, suggesting that these deposits are immune complexes to muscle. These deposits are seen in a variety of muscle-wasting conditions and may be nonspecific. Cellular immunity is abnormal with DM-PM, as demonstrated by myotoxic activity of the lymphocytes in patients with DM-PM. Skeletal muscle antigens cause lymphocytes of patients with DM-PM to proliferate, suggesting that the lymphocytes are inappropriately responding to these antigens.

**Crystal-Induced Synovitis**

Crystal-induced synovitis can be caused by uric acid, calcium pyrophosphate, hydroxyapatite, and cholesterol crystals. Best understood is gout, a familial disorder in which there is a deficiency of hypoxanthine-guanine phosphorotransferase, resulting in an overproduction of uric acid. Hyperuricemia results, and as the concentration of urate in the blood increases, monosodium urate crystals precipitate in the tissue. It has been shown that injecting urate crystals subcutaneously will cause tophus formation, and when urate is injected into joints, gouty attacks will ensue. Other factors involved in the pathogenesis of the gouty attack include elevated temperature, which increases the joint urate concentration, lowered pH, trauma, and aging (19).

Pseudogout, or calcium pyrophosphate dihydrate deposition (CPPD), can be hereditary or sporadic. The etiologic agent is the calcium pyrophosphate crystal, which is formed secondary to a disorder of local pyrophosphate metabolism. The crystals adhere to leukocytes, and often immunoglobulin is absorbed, which stimulates phagocytosis and the perpetuation of inflammatory arthritis. These calcium-containing crystals get deposited in the pericellular matrix of cartilage and can present as chondrocalcinosis (20).

**Spondyloarthropathies**

Spondyloarthropathies are polyarticular disorders that primarily involve the sacroiliac joints, vertebral column and, to a lesser extent, larger, peripheral joints (shoulder and hip). There may also be an association with a variety of extraspinal lesions including the eyes, the GI track, cardiovascular system, lungs, kidneys, and skin. These arthritides share a number of additional features, including mucocutaneous lesions, sacroiliitis, heel pain, and the B27 antigen. Antecedent GI infection, caused by *Salmonella*, *Shigella*, and *Yersinia*, has stimulated interest that these diseases may be caused by a gram-negative organism (21). The most convincing data in support of this came from the well-documented *Shigella* epidemic, in which RS occurred in 344 of 150,000 infected persons in one study (22) and in nine of 602 in another study. No case occurred in anyone who was not infected. Arthritis with some additional features of RS has followed *Salmonella* and *Yersinia* infections. In PSA and AS, the data are less convincing, but evidence has linked the development of guttate psoriasis and streptococcal infection. The presence of B27 antigen appears to be the crucial link in the expression of the disease. Antecedent urethritis has also been associated with acute arthritis, and chlamydia is the organism most frequently identified.

The pathology occurs at the entheses (insertion of tendon to bone). The axial spine may fuse, but if the peripheral joints are involved, there is erosion and often bony reaction and periosteal new bone growth. Usually, there is no particular osteoporosis (OP). In summary, the etiology of spondyloarthropathies is most likely an infective agent, possibly Gram-negative bacteria, that interacts with a susceptible gene host: B27 in AS and RS, and perhaps B27, B38, and C6 in PSA (23).

**Infectious Arthritis**

A wide variety of infectious agents can cause arthritis secondarily to the infection itself or as a consequence of the host’s immunologic response. The organisms can be viral (e.g., hepatitis, rubella, mumps, herpes); bacterial (e.g., Gram-positive *Staphylococcus*, *Streptococcus*, and *Pneumococcus*; Gram-negative *Neisseria* and *Hemophilus influenzae*; *Pseudomonas*, mycobacterium tuberculosis), spirochete (Lyme disease), or fungal. Recently, interest has developed in the role of hepatitis C and the development of an RA-like arthritis (24).

**PATIENT EVALUATION**

A detailed history, physical examination, laboratory, and x-ray findings are essential to the proper diagnosis and management of RDs. Many schemes have been developed in an attempt to construct an organized approach to the classification of RDs, including algorithms that sort signs and symptoms around the presence or absence of inflammation, symmetry, and number of involved joints. However, these categories are not very helpful in sorting out the underlying pathophysiologic processes that need therapeutic intervention. A practical approach is suggested by James Fries, in which eight specific types of musculoskeletal pathology are distinguished (25). Patients can have more than one type of pathology; for example, enthesopathy and synovitis present in patients with PSA (Table 40-5). A more recent approach to patient evaluation, incorporating the World Health Organization (WHO) biopsychosocial model for patient evaluation, may be useful (26). Patient assessment from this approach will seek out not only the primary etiology and how the disease process affects the organ or system but also the associated effects on the patient and his or her interactions with the environment.

**History and Examination**

A detailed description of the symptom onset, prodromal symptoms, setting, pattern, and sequence will greatly aid in establishing a differential diagnosis. Exacerbating or remitting factors, functional impairments, and therapeutic effects should also be documented (27).
Pain is the most likely presenting symptom. Although often difficult to define, anatomic location and symmetry, character (e.g., burning, aching), and severity or intensity (graded by a numeric ten-point scale) should be included.

Stiffness may be seen as well. This term should be well-defined and its meaning to the patient should be understood. Timing, duration, and location of stiffness should be noted. For example, in the diagnostic criteria of RA, morning stiffness lasting at least 1 hour is included. This stiffness also known as a “gelling phenomenon” can be noted after sitting or maintaining a fixed position for some period of time which in some cases may be as short as 20 to 30 minutes but typically greater than 1 hour.

Range-of-motion (ROM) limitations may accompany complaints of stiffness. These symptoms in the rheumatologic population are usually not transient and should not be confused with gelling or stiffness. Noting symptom onset may help to differentiate an acute from a chronic process, such as nonreducible joint subluxation. Most important is to determine if the loss of ROM is fixed or likely to be successfully ranged, as a chronic ROM limitation may indicate ankylosis or autofusion of a joint. Passive and active ROM testing should be performed to rule out weakness as an etiology for ROM limitation.

Joint swelling should be documented by onset, persistence, location, and quantity (i.e., serial circumference measurements). This helps differentiate acute inflammatory from non-inflammatory involvement. The former is best not treated with heat while the latter can be treated easily with either a heat or cold modality. When pain is noted with swelling, synovitis or bursitis may be present.

It is important to differentiate weakness from fatigue. Careful documentation of patient reports on muscle groups involved and relationship to functional limitations is needed. Proximal muscle weakness may be indicative of inflammatory myopathy such as PM. Persistent rather than intermittent complaints of weakness may indicate other processes, such as neuromuscular disease (e.g., Guillain-Barré’s syndrome).

Fatigue may be one of the earliest symptoms. Some patients may have these complaints even though pain and swelling have been objectively controlled.

### Biomechanics

Evaluation of mobility can be performed with visual analysis or automated measures. The former has been standardized, and the latter have been significantly advanced with video-based high-speed systems. There is an increased ability to reliably measure motion in three dimensions: in real time, ground reaction forces, and pressures on the bottoms of the feet to calculate moments of force at various joints. These procedures are now more available and being performed fairly frequently. In addition, newer instrumentation has been developed to describe foot pressure profiles and describe forces and their influences on the foot (28).

Gait abnormalities for several RDs have been noted. The RA gait has been termed apropulsive because of the absence of pushoff from the ball of the foot. Similarly, studies of differences in gait before and after surgical procedures have been reported that describe which biomechanical changes occur as a result (29).

### TABLE 40.5 Evaluation of RDs

<table>
<thead>
<tr>
<th>Pathology</th>
<th>Examples</th>
<th>Laboratory Tests</th>
<th>Other Organs Involved</th>
</tr>
</thead>
<tbody>
<tr>
<td>Synovitis</td>
<td>Rheumatoid arthritis</td>
<td>Latex, x-rays</td>
<td>Lung, heart, skin nodules</td>
</tr>
<tr>
<td>Enthesopathy</td>
<td>Psoriatic arthritis/Reiter’s syndrome</td>
<td>x-rays</td>
<td>Skin</td>
</tr>
<tr>
<td>Crystal arthritis</td>
<td>Gout</td>
<td>HLA-B27, sacroiliac joint, X-rays, MRI</td>
<td>Heart</td>
</tr>
<tr>
<td>Joint infection</td>
<td>Psoriatic arthritis/Reiter’s syndrome</td>
<td>Serum uric acid, joint fluid</td>
<td>Skin, mucous membrane</td>
</tr>
<tr>
<td></td>
<td>Ankylosing spondylitis</td>
<td></td>
<td>Skin, kidney</td>
</tr>
<tr>
<td>Joint effusion</td>
<td>Fungal</td>
<td>Joint culture, joint fluid</td>
<td>Vaginal infection</td>
</tr>
<tr>
<td>Vasculitis</td>
<td>Trauma</td>
<td></td>
<td>Bacteremia</td>
</tr>
<tr>
<td>Tissue conditions</td>
<td>Reactive arthritis</td>
<td>Joint fluid</td>
<td>Hepatitis</td>
</tr>
<tr>
<td>Local</td>
<td>Pseudogout</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Generalized</td>
<td>DM-PM</td>
<td>Muscle biopsy</td>
<td>Thyroid livers</td>
</tr>
<tr>
<td></td>
<td>Scleroderma</td>
<td>EMG</td>
<td>Any organ</td>
</tr>
<tr>
<td></td>
<td>SLE</td>
<td>Antinuclear antibody</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Polymyalgia</td>
<td>ESR</td>
<td>Heart</td>
</tr>
<tr>
<td></td>
<td>Metabolic/endocrine disorders</td>
<td>CRP</td>
<td></td>
</tr>
</tbody>
</table>

DM-PM, dermatomyositis-polymyositis; CRP, C-reactive protein; EMG, electromyography; SLE, systemic lupus erythematosus; MRI, magnetic resonance imaging.
**Laboratory Tests**

The laboratory evaluation of blood, urine, and synovial fluid, coupled with radiographic evaluations, history, and physical examination information, can usually help to establish a proper diagnosis. The following initial determinations are made: complete blood count, ESR, SMA 12 (sequential multiple analyzer), RF, and antinuclear antibody (ANA). An HLA B27 determination is performed if spondylitis is suspected. The acute-phase reactants (C-reactive protein [CRP], serum amyloid [SAA], and ESR) should be monitored, as they may be part of the early defense or adaptive mechanisms that precede the immune response. Although nonspecific, moderate elevation in CRP and serum amylase can be seen in systemic diseases.

Joint fluid is easy to obtain in the presence of effusion. Analysis of fluid is essential in the diagnosis of crystal-induced arthritis and infection, and it is helpful in differentiating traumatic from inflammatory arthritis. However, rarely will the diagnosis of RA, OA, PSA, or AS be made on the basis of joint fluid alone. Rather, the fluid helps confirm a diagnosis. Joint taps must be done when a question of infection is raised and should be made before injecting steroid or other material into the joint. Classification of joint fluid into categories will help differentiate inflammatory, noninflammatory, septic, and hemorrhagic arthritis (see Table 40-1).

**Radiographic Assessment**

Radiography is often the most valuable technique for differentiating among arthritides. Carefully selected radiographic series with the proper projections, addition of stress, and weight-bearing views will add valuable information about the extent of soft tissue, articular surface, or bony changes. Marginal erosion of bone with juxtaarticular OP is the hallmark of RA. Nonuniform joint space losses, in association with bony sclerosis and marginal osteophytes, are the characteristic changes in OA. Spondyloarthropathies classically involve the sacroiliac joints, either symmetric, as in AS, or asymmetric, as in RS and PSA. Bony changes include periosteal new bone formation and ankylosis. Gout and pseudogout often involve only a few joints. In gout there are soft-tissue tophaceous deposits and marginal erosion with large bony overhangs, and in pseudogout there is calcinosis in fibrocartilage. Early in joint infection, the x-ray films may be negative, or there may be some joint space widening. If the process continues and osteomyelitis develops, periosteal reaction can occur, which may indicate progressive infection or bony destruction. Table 40-2 presents typical radiographic findings in patients with RDs.

Additional imaging such as computed tomography (CT), which gives good structural definition of soft tissue and bone, is often combined with arthrography to study axial structures in disorders such as sacroiliitis. Magnetic resonance imaging (MRI) allows further differentiation of soft tissue and fluid as well as use of variable imaging planes and combination. Gadolinium for contrast is of great benefit in the evaluation of joint effusions, tendinopathies, and myositis (30). Ultrasound may be useful in evaluating this patient population as well (31).

**Functional Assessment**

Rehabilitation assessment for patients with RDs includes both process and outcome measures. Goniometry, the measurement of joint ROM, is standardized and widely used, as is manual muscle testing (MMT). A new ten-point MMT with specific grade definitions has been devised and offers more sensitivity in the strength range that is most important to know for assessment of capability for independence (32). Quantifiable measures of spine motion are particularly useful for patients with RDs (33). They can help chart progressive loss of spinal mobility, which prompts interventions designed to preserve posture as well as chest expansion programs, as in the management of patients with spondyloarthropathies.

Patients with arthritis often have stiffness rather than pain that limits function. Both symptoms are difficult to measure. However, duration of morning stiffness may be quantified. Pain can be measured in terms of severity in a descriptive way (e.g., mild, moderate, severe) or by use of a visual analog scale (34), which is quite reliable. Measures of degree of joint tenderness, swelling, deformity, relative instability, or crepitus with active and passive movement are also useful in defining RD processes.

Fatigue is a frequent problem for patients with RD. Its cause is multifactorial: medication, chronic inflammation, abnormal posture and gait that are energy inefficient, abnormalities of the sleep cycle, and atrophy of muscle secondary to disease or chronic pain. Fatigue is difficult to quantify. A visual analog scale has been used in the subjective measurement of fatigue but it has an imprecise reference. A multidimensional assessment of fatigue (35) has been devised and validated in this population. The fatigue severity scale has also been used to assess this parameter (36). The human activity profile (HAP) is an instrument designed to measure the amount of activity and those activities that an individual is no longer able to perform. It also has a dyspnea scale. Specific activities have been correlated with metabolic equivalents required for performance (37).

Despite reliable, sensitive indices of strength, ROM, and grip strength, other measures are needed for evaluation of patients with RD. The American Rheumatism Association in 1949 devised a functional scale for patients with RA. This scale, a simple, global assessment that rated patients’ functional status as independent (i.e., class I), able to perform with pain (i.e., class II), able to do some activities (i.e., class III), and unable to perform (i.e., class IV) was revised in 1992 (38).

Two generations of functional assessments have been used in evaluating patients with RD. The first set looked primarily at performance of patients in ambulation, self-care, and other activities of daily living (ADL). Most had some testing of reliability and validity, and were relatively easy to use. The problem with them was that they defined function very narrowly and excluded psychological, social, and vocational functions. The newer functional indices are more comprehensive and offer a broader view of patients’ functioning. These global, multidimensional tools had been designed for the arthritis
population, children and adults, and have demonstrated validity and reliability (39,40).

When rheumatologists were asked which functional measures were important to use in the evaluation of patients with RDs, the consensus was mobility, pain, self-care, and role activity (41). The evaluations needed may vary, because some RDs involve only joints (e.g., OA); others, primarily kidneys, skin, and central nervous system (e.g., SLE); and still others, different organ systems, such as cardiovascular and pulmonary systems. Table 40-6 identifies standard functional measures likely to be needed for each of the RDs. Other useful scales include The Wisconsin Brief Pain Questionnaire, The sickness impact profile (SIP) (42), Stanford health assessment questionnaire (HAQ) (43), Short form 36 (SF36) version-2 (44), arthritis impact measurement scale-2 (AIMS-2) in RA and Osteoarthritis (OA), (45) and Bath and Duogados functional indices in spondylitis (46). These functional assessment tools singly and in combination are valuable tools for measuring both physical and psychosocial health parameters (Tables 40-7 and 40-8).

Functional assessment tools are also used in combination with traditional measures of disease state, including tender and swollen joint counts (number of tender and inflamed joints), and biochemical markers of disease state such as the acute-phase reactants. The combination of these tools into core set outcome measures is used to define clinically significant improvement in diseases such as RA. The ACR has chosen the definition of 20% improvement in tender and swollen joint counts and improvement in three of the five following ACR core set measures: patient or physician global assessments of disease severity, pain, disability, and an acute-phase reactant level (47). This is commonly referred to as the ACR-20. Clinical improvement in the preceding core set measures that constitutes improvement of 50% and 70% (ACR-50 and ACR-70) are often referred to in randomized control trials. The Outcome Measures in Rheumatology Clinical Trials (OMERACT) group has reached consensus on the required core set measures of use in OA (48) and AS (49). Core set outcome measures in other populations, such as those with DM-PM, are currently in development.

Although these tools are easily administered, results are often not immediately available for use because of scoring schemes. Some of these tools (e.g., HAQ) also have limited use because of ceiling effects and often miss subtle changes in the patient’s disease process reflected in level of function; others may have significant flooring effects.

In the clinical setting of RA, a combination of factors including specific clinical and laboratory values are used to produce a disease activity score (DAS28). The DAS28 is derived from the number of tender and swollen joints, patient assessment of disease activity via visual analog scale, ESR, and CRP (50).

---

**TABLE 40.6 Functional Measures in RDs**

<table>
<thead>
<tr>
<th>MMT</th>
<th>Range of Motion</th>
<th>Pain</th>
<th>Fatigue</th>
<th>ADLs</th>
<th>Ambulation</th>
<th>Cognition</th>
<th>Role/Social Interaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>OA</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++++</td>
</tr>
<tr>
<td>RA</td>
<td>+</td>
<td>++</td>
<td>+</td>
<td>+++</td>
<td>++</td>
<td>++</td>
<td>++++</td>
</tr>
<tr>
<td>Spondyloarthropathies</td>
<td>+++</td>
<td>++</td>
<td>+</td>
<td>+++</td>
<td>++</td>
<td>++</td>
<td>++++</td>
</tr>
<tr>
<td>DM-PM</td>
<td>++</td>
<td>+</td>
<td>++</td>
<td>+</td>
<td>++</td>
<td>+++</td>
<td>+++</td>
</tr>
<tr>
<td>PSS</td>
<td>+</td>
<td>+++</td>
<td>+</td>
<td>+</td>
<td>++</td>
<td>++</td>
<td>+++</td>
</tr>
<tr>
<td>SLE</td>
<td>+</td>
<td>+++</td>
<td>++</td>
<td>+</td>
<td>++</td>
<td>++</td>
<td>+++</td>
</tr>
<tr>
<td>Gout (crystals)</td>
<td>+++</td>
<td>+</td>
<td>++</td>
<td>+</td>
<td>++</td>
<td>++</td>
<td>+++</td>
</tr>
<tr>
<td>Fibromyalgia</td>
<td>+++</td>
<td>+++</td>
<td>+</td>
<td>+++</td>
<td>++</td>
<td>++</td>
<td>+++</td>
</tr>
</tbody>
</table>

+, possibly useful evaluation; ++, recommended evaluation; ++++, strongly recommended; ++++, must evaluate.

MMT, manual muscle test; OA, osteoarthritis; PSS, progressive systemic sclerosis; RA, rheumatoid arthritis; DM-PM, dermatomyositis-polymyositis; SLE, systemic lupus erythematosus; OA, osteoarthritis; RA, rheumatoid arthritis; PSS, progressive systemic scoliosis; ADLs, activities of daily living.

---

**TABLE 40.7 Assessments Measuring Physical Health Parameters**

<table>
<thead>
<tr>
<th>Mobility</th>
<th>Self-care Roles</th>
<th>Communication</th>
<th>Pain</th>
</tr>
</thead>
<tbody>
<tr>
<td>American College of Rheumatology (ACR)</td>
<td>Global</td>
<td>Global</td>
<td>0</td>
</tr>
<tr>
<td>Stanford Health Assessment Questionnaire (HAQ)</td>
<td>++</td>
<td>+++</td>
<td>0</td>
</tr>
<tr>
<td>Arthritis Impact Measurement Scale (AIMS II)</td>
<td>+++</td>
<td>++</td>
<td>+</td>
</tr>
<tr>
<td>Sickness Impact Profile (SIP)</td>
<td>+++</td>
<td>+++</td>
<td>+</td>
</tr>
<tr>
<td>Short Form 36 (SF36) version 2</td>
<td>++</td>
<td>+</td>
<td>0</td>
</tr>
</tbody>
</table>


0, no questions in this area; +, few questions in this area; ++, moderate number of questions in this area; +++, many questions in this area.
Compliance

It has long been recognized that a number of factors may influence compliance of patients with treatment in general:

1. Demographic features
2. Nature of the disease
3. Therapeutic regimens
4. Setting in which treatment is given
5. Patient-doctor relationship
6. Sociobehavioral features of the patient

Compliance depends on individual health beliefs, including the importance to the patient of the treatment goal, how likely the treatment is to achieve the goal and benefit the patient (51), and how likely the treatment is to lessen the disability and the physical, psychological, and functional barriers to treatment.

In our clinical experience, patients who have pain are more likely to be compliant with medication, modalities, and techniques that relieve the pain. Education significantly increases adherence to drug regimens (52). Group education enhances self-management strategies (49).

One study of adolescents with chronic diseases (including JIA) indicated that good motivation was likely to result in better treatment compliance. Support from parents, physicians, and friends also predicted good compliance with regimens (53).

Compliance with the use of splints/orthotics has been low due to cosmesis, uncomfortable fit (more likely with noncustomized items), and fear that wear at work may jeopardize job status.

Compliance with unsupervised exercise programs tends to be low. In one study, two out of three patients with some form of arthritis management information used some technical orders and rest daily, and 50% used exercise and heat (25% on a daily basis) (54). Compliance at the 1-year level in patients on a home exercise program was predicted by self-efficacy for exercise, regular ROM before study intervention, and single marital status (55). Perceived benefit of exercise was a significant predictor of participation in an aerobic exercise program. Those who reported exercising in their youth perceived more benefits. Subjects with less formal education, longer arthritis duration, and higher impact of arthritis scores perceived fewer exercise benefits (56).

Compliance with splint use has been shown to increase when patients learn the purposes of use, expectations of use, and precautions for wear and when the splint provider has a positive affective tone and gives encouragement (57). Strategies to improve compliance with exercise can be seen in Table 40-9, and Table 40-10 lists strategies for improving compliance with orthotics and gait aid.

**TREATMENT**

**Pharmacologic Management**

Pharmacologic management of RDs often requires the use of one or more of a variety of medications, the pharmacology and pharmacokinetics of which may influence physical and psychological functioning. The well-known “treatment pyramid” approach to the management of RA and other RDs was the past based on use of initial rest, patient education, joint protection, and nonsteroidal anti-inflammatory drugs (NSAIDs) with progression to steroids and sequential monotherapy, and use of DMARDs later in the course of these diseases. This treatment approach has undergone revision with improved knowledge of the rate of disease progression, prognosis, and
severity at presentation. DMARDs are used earlier and often in combination (Fig. 40-2) (58). At the onset of disease symptoms and diagnosis, the patient would begin early rehabilitation interventions: education, orthotics, physical modalities, joint protection, energy conservation, and strengthening, along with NSAIDs and low-dose steroid therapy in some cases. As we ascend the pyramid, introduction of DMARDs used singly or in combination and earlier use of biologic agents is noted. The time interval at which new agents or dosage changes are made has also been shortened.

**Aspirin**

Aspirin, or acetylsalicylic acid (ASA), has been the foundation of management of rheumatic conditions and the symptoms of pain, fever, and inflammation. It has been shown to block the synthesis of PG in the anterior hypothalamus, which is responsible for the antipyretic effect. The analgesic effect of ASA is not entirely understood. Musculoskeletal pain may be mediated by bradykinin, a synthesizer of PG, which sensitizes nerves to painful stimuli. Aspirin blocks PG synthesis. At doses higher than those used for analgesia (e.g., 5.3 g/day), ASA reduces joint inflammation and swelling. The mechanisms for this action are multifactorial. Aspirin affects leukocyte migration and vascular permeability, both of which may be influenced by PG synthesis. The toxicities of ASA include allergy, tinnitus and hearing loss, GI blood loss, ulcer, chemical hepatitis, and reduced glomerular filtration rate. For patients who have clinically significant GI symptoms, enteric-coated preparations are usually well tolerated. Other forms of salicylate can be used that are often less GI toxic (e.g., choline salicylate).

**Nonsteroidal Anti-inflammatory Drugs**

The agents that form the group of drugs called NSAIDs, which include the cyclooxygenase-2 (COX-2) inhibitors, continue to be used in part as first-line agents in the management of RDs. These drugs also suppress inflammation through the inhibition of synthesis of PG. They inhibit the COX effect on platelets and effects on leukocyte migration. Toxicities include GI bleeding, pancreatitis, hepatotoxicity, decreased renal blood flow, hypertension, peripheral edema, and allergic interstitial nephritis. Some have more GI toxicity than others and cause more sodium retention. A review of the comparative NSAIDs toxicities is available (59).
NSAIDs were widely used as first-line drugs in the treatment of RA, JRA, OA, and the spondyloarthopathies but have more recently been replaced by the early use of DMARDs in some cases. Only tolmetin sulfate, choline magnesium trisalicylate, ibuprofen, and naproxen sodium have been approved for use in children by the U.S. Food and Drug Administration. A list of those commonly used NSAIDs in the United States is provided in Table 40-11.

Aspirin and NSAIDs are likely to provide significant clinical relief for patients with OA, RA, JIA, gout, and spondyloarthritides. These drugs are not usually effective by themselves in controlling RA and vasculitic syndromes.  The newest group of NSAIDs, the COX-2 inhibitors, has found more recent favor. They differ from traditional NSAIDs in controlling RA and vasculitic syndromes. They are associated with fewer untoward effects. Current recommendations to reduce bone loss in patients receiving a prednisone equivalent of more than 5 mg/day include (a) use of supplemental calcium and vitamin D or activated form of vitamin D, (b) use of bisphosphonates, or calcitonin as a second-line agent for those with contraindications or intolerance to bisphosphonates, (c) hormone replacement therapy in those found to be hormone deficient (61). The oral route is usually selected for ease of administration, but glucocorticoids can safely be given intramuscularly, intravenously, or intra-articularly.

Glucocorticoids are used to treat several RDs. In 2002, a consensus conference made the following recommendations for a standardized nomenclature for glucocorticoid dosing. Doses of less than 7.5 mg—low dose, 7.5 to 30 mg/day—medium dose, 30 to 100 mg—high dose, greater than 100 mg/day—very high dose, and pulse-dosed therapies given at greater than 250 mg/day (62). Higher doses are used in treating patients with SLE, vasculitis, and DM-PM (up to 100 mg prednisone every day) (63). The benefit of steroid therapy to patients with AS, PSS, and PSA has not been shown.

Disease-Modifying Antirheumatic Drugs

Methotrexate

This cytotoxic agent has become one of the first line agents in the treatment of moderate to severe RA. It is a structural analog of folic acid and causes the deficiency of intracellular folate by inhibition of dihydrofolate reductase, an enzyme required for DNA synthesis. This may not be the true mechanism of action that makes methotrexate (MTX) an effective anti-inflammatory agent. MTX also may act to promote the extracellular adenosine release which promotes the down regulation of inflammatory pathways by binding surface receptors on lymphocytes, monocytes, and neutrophils and inhibit interleukin production.

MTX has been shown to be effective in RA, JRA, SLE, and PSA and is effective in combination with corticosteroids in the management of some vasculitic conditions. It can be given in an oral form once per week at doses of 7.5 to 15 mg/week. Maximum doses of 25 mg/week may be given. Weekly subcutaneous or intramuscular injections are well tolerated and are associated with less GI side effects. Side effects include increased liver transaminases, movement, leukocyte function, and humoral factors; inhibit recruitment of neutrophils and monocytes into inflammatory sites; cause lymphocytopenia by inducing margination or redistribution of lymphocytes out of the circulation; modify the increased capillary and membrane permeability that occurs at an inflammatory site, reducing edema and antagonizing histamine-induced vasodilation; and inhibit PG synthesis.

Daily high-dose steroid use stimulates Cushing’s syndrome, in which hypertension, hirsutism, acne, striae, obesity, psychiatric symptoms, and wound-healing problems occur. With exogenous doses above 12.5 mg/day, there is an increased incidence of glaucoma, cataracts, avascular necrosis, OP, and pancreatitis. The side effects are, in part, dependent on the particular glucocorticoid used and the dose. Alternate-day steroids are associated with fewer untoward effects. Current recommendations to reduce bone loss in patients receiving a prednisone equivalent of more than 5 mg/day include (a) use of supplemental calcium and vitamin D or activated form of vitamin D, (b) use of bisphosphonates, or calcitonin as a second-line agent for those with contraindications or intolerance to bisphosphonates, (c) hormone replacement therapy in those found to be hormone deficient (61). The oral route is usually selected for ease of administration, but glucocorticoids can safely be given intramuscularly, intravenously, or intra-articularly.

Glucocorticoids

Glucocorticoids and therapeutics for RDs are inseparable and probably have been tried in every RD either systemically or locally. Exogenous glucocorticoids influence leukocyte function and may include increased liver transaminases.
myelosuppression, pneumonitis, cirrhosis, increased risk of infection, and can be highly teratogenic.

**Hydroxychloroquine**

Antimalarials, such as hydroxychloroquine (HCQ), are effective in discoid lupus erythematosus and SLE. Improvements in skin involvement, arthritis, and arthralgias are noted. They can also be effective in patients with RA with improvements noted in joint count, grip strength, walk time, and sedimentation rate. The antimalarials are slow-acting, taking 4 to 6 weeks before a therapeutic effect is observed. They are as effective as some of the other DMARDs, but a lower toxicity profile makes them one of the agents of choice in combination therapy (64,65).

The mechanism of action of these drugs is varied. They have been shown to impair enzymatic reactions, including phospholipase, cholinesterase-hyaluronidase, and proliferation of lymphocytes. They seem to block depolymerization by DNase and interfere with DNA replication. The incidence of side effects and toxicities varies widely. GI disturbance is quite common, and retinopathy is infrequent but of greatest concern; it rarely occurs before a cumulative dose of 300 mg is reached, specifically in chloroquine, but routine ophthalmologic examinations should be performed. The antimalarial agents are most often used in the treatment of RA and SLE.

**Sulfasalazine (SSA)**

Sulfasalazine (SSA) is an agent that combines the antibiotic sulfapyridine with the anti-inflammatory agent 5-aminosalicylic. Its proposed mechanisms of action include inhibition of folate-dependent enzymes similar to MTX, immunomodulatory functions that decrease immunoglobulin and RF production, and several anti-inflammatory properties (66). SSA has been shown to be effective in treating RA, JRA, AS, and PSA.

**Leflunomide**

Leflunomide is an immunosuppressive drug that inhibits de novo pyrimidine synthesis and impairs T-cell proliferation. In the treatment of RA, this drug has been shown to be as effective as MTX (67). This agent has been shown not only to be effective as monotherapy but also to be even more effective when used in combination with MTX; however, concerns remain over the toxicity profile (68). Leflunomide is shown to be effective in the treatment of RA, SLE, and PSA. Monitoring while on this medication includes (a) liver function tests for toxicity and (b) platelet count for signs of thrombocytopenia.

**d-Penicillamine**

d-Penicillamine has been effective in the treatment of seronegative or seropositive RA but has not been used as often due to the associated high incidence of toxicity (69). The patterns of response to penicillamine are similar to those observed with gold. Toxicities include leukopenia, thrombocytopenia, proteinuria, skin rash, stomatitis, GI upset, and a variety of autoimmune syndromes, including Goodpasture disease, PM, and SLE. The mechanism of action is unknown, but it is neither cytotoxic nor anti-inflammatory.

**Immunoregulatory Agents**

Immunoregulatory drugs have been used in the management of RDs in an attempt to restore a balanced immune response by eliminating certain cell subsets. None of these drugs have cured patients with RDs, but they have produced some disease control and long-term remissions (70). Commonly used agents are the alkylation agents (cyclophosphamide and chlorambucil), purine analogs (azathioprine and mercaptopurine), cyclosporine, tacrolimus, sirolimus, mycophenolate mofetil, dapsone, and thalidomide.

**Cyclophosphamide and Chlorambucil**

Cyclophosphamide and Chlorambucil are alkylating agents that form active metabolites that cross-link DNA, preventing replication, and reducing DNA synthesis. Immunoregulatory effects are through the decrease of both T and B-lymphocyte proliferation, antibody production, and suppression of delayed hypersensitivity reactions to new antigens. Cyclophosphamide has been shown to be effective in treating SLE and RA, while chlorambucil is utilized for RA. These agents have toxicity profiles that include myelosuppression, increased risk of infection, and risk of malignancy (70).

**Azathioprine and Mercaptopurine**

The purine analogs, azathioprine and mercaptopurine, are converted to thiopurine, metabolized, and incorporated into cellular DNA, leading to inhibition of nucleic acid synthesis. Through this mechanism, it is believed that these agents function to decrease circulating lymphocyte count, suppress lymphocyte proliferation, inhibit antibody production, and inhibit monocyte production as well as cell mediated and humoral immunity (70). These agents have been used to treat RA and SLE and are well tolerated. The most common side effects include GI symptoms, myelosuppression, and risk of infection.

**Cyclosporine**

Cyclosporine exerts its effects by inhibiting the production of interleukin-2 and other cytokines, leading to reduction in T-cell activation and lymphocyte proliferation. It has been used in the treatment of RA and PSA. Toxicities include GI upset, hypertension, nephrotoxicity, and increased risk of lymphoma and skin cancer. Tacrolimus, a macrolide, functions by binding to an intracellular binding protein (FK-binding protein) and in association with calcineurin suppresses the transcription of cytokines and inhibits the early steps of T-lymphocyte activation (71). Sirolimus also binds to the FK binding protein but functions through blocking the progression of the cell cycle, inhibiting cell signal transduction (72). Mycophenolate mofetil is converted to Mycophenolate acid that reversibly inhibits inosine monophosphate dehydrogenase, an enzyme required for the synthesis of purines, thus inhibiting T- and B-cell lymphocyte proliferation. It has been used in the treatment of SLE and associated nephritis.
<table>
<thead>
<tr>
<th>Diseases</th>
<th>Recommended Medications</th>
<th>Probable Mechanisms</th>
<th>Precautions</th>
</tr>
</thead>
<tbody>
<tr>
<td>RA</td>
<td>ASA, NSAIDs</td>
<td>Inhibition of COX enzyme needed for PG synthesis</td>
<td>Bleeding diathesis (platelet dysfunction)</td>
</tr>
<tr>
<td></td>
<td>Antimalarials</td>
<td>Block lysosomal enzymes</td>
<td>Upper GI toxicity</td>
</tr>
<tr>
<td>Gold</td>
<td></td>
<td>Inhibits phagocytic activity of macrophages</td>
<td>Renal toxicity</td>
</tr>
<tr>
<td></td>
<td>o-penicillamine</td>
<td>Unknown</td>
<td>Retinal toxicity, psoriasis</td>
</tr>
<tr>
<td>Steroids</td>
<td></td>
<td>Interfere with lymphocytic migration; decreases intra articular membrane permeability</td>
<td>Nephritis, rash, marrow suppression</td>
</tr>
<tr>
<td>Azathioprine</td>
<td></td>
<td>Inhibits DNA synthesis</td>
<td>Nephritis, SLE, PM</td>
</tr>
<tr>
<td>Methotrexate</td>
<td></td>
<td>Causes intracellular folate deficiency</td>
<td>Lymphoid tumors</td>
</tr>
<tr>
<td>Cyclophosphamide</td>
<td></td>
<td>Prevents DNA replication</td>
<td>Not used with allopurinol</td>
</tr>
<tr>
<td>Cyclosporine</td>
<td></td>
<td>Inhibits IL-2 production</td>
<td>Cirrhosis, leukopenia</td>
</tr>
<tr>
<td>Biologic agents</td>
<td></td>
<td>Inhibits TNF</td>
<td>Marrow suppression, liver toxicity, lung fibrosis</td>
</tr>
<tr>
<td>Infliximab/etanercept</td>
<td></td>
<td>Inhibits pyrimidine synthesis, cell activation and adhesion</td>
<td>Ovarian cystitis</td>
</tr>
<tr>
<td>Leflunomide</td>
<td></td>
<td>Suppress lymphocyte and leukocyte fxn</td>
<td>Nephrotoxicity, hypertension, ↑ infection risk</td>
</tr>
<tr>
<td>Sulfasalazine</td>
<td></td>
<td>Uregulation of IL-10 (anti-inflammatory cytokine)</td>
<td>Injection site rxn, ↑ infection risk</td>
</tr>
<tr>
<td>Minocycline</td>
<td></td>
<td></td>
<td>Alopecia, stomatitis, abdominal pain, ↑ LFTs, hypertension</td>
</tr>
<tr>
<td>Spondyloarthritis</td>
<td></td>
<td></td>
<td>Hepatitis, marrow suppression, rash diarrhea</td>
</tr>
<tr>
<td></td>
<td>ASA, NSAID</td>
<td>As above</td>
<td>GI toxicity, rash</td>
</tr>
<tr>
<td>Gold</td>
<td></td>
<td>As above</td>
<td>As above</td>
</tr>
<tr>
<td>Methotrexate</td>
<td></td>
<td>As above</td>
<td>As above</td>
</tr>
<tr>
<td>Sulfasalazine</td>
<td></td>
<td>As above</td>
<td>As above</td>
</tr>
<tr>
<td>Gout</td>
<td>NSAID</td>
<td>As above</td>
<td>As above</td>
</tr>
<tr>
<td></td>
<td>Uricosurics (probenecid)</td>
<td>Increases excretion of uric acid</td>
<td>Often need to alkalinize urine</td>
</tr>
<tr>
<td></td>
<td>Allopurinol</td>
<td>Inhibits xanthine oxidase</td>
<td>Do not use with azathioprine</td>
</tr>
<tr>
<td></td>
<td>Colchicine</td>
<td>Inhibits microtubular assembly and inhibits lysosomal enzyme release</td>
<td></td>
</tr>
<tr>
<td>SLE</td>
<td>NSAID</td>
<td>As above</td>
<td>As above</td>
</tr>
<tr>
<td>Steroids</td>
<td></td>
<td>As above</td>
<td>As above</td>
</tr>
<tr>
<td>Antimalarials</td>
<td></td>
<td>As above</td>
<td>As above</td>
</tr>
<tr>
<td>Azathioprine</td>
<td></td>
<td>As above</td>
<td>As above</td>
</tr>
<tr>
<td>Cyclophosphamide</td>
<td></td>
<td>As above</td>
<td>As above</td>
</tr>
<tr>
<td>PSS</td>
<td>o-penicillamine</td>
<td>Unknown</td>
<td>As above</td>
</tr>
<tr>
<td></td>
<td>Colchicine</td>
<td>As above</td>
<td>As above</td>
</tr>
<tr>
<td>DM-PM</td>
<td>Steroids</td>
<td>As above</td>
<td>As above</td>
</tr>
<tr>
<td></td>
<td>Azathioprine</td>
<td>As above</td>
<td>As above</td>
</tr>
<tr>
<td></td>
<td>Methotrexate</td>
<td>As above</td>
<td>As above</td>
</tr>
</tbody>
</table>

GI, gastrointestinal; PG, prostaglandin; RA, rheumatoid arthritis; DNA, deoxyribonucleic acid; TNF, tumor necrosis factor; ASA, acetylsalicylic acid; SLE, systemic lupus erythematosus; COX, cyclooxygenase; PSS, progressive systemic sclerosis; DM-PM, dermatomyositis-polymyositis; LFTs, liver function tests; NSAIDs, nonsteroidal anti-inflammatory drugs.

Thalidomide, a derivative of glutamic acid, is believed to exert its immunosuppressive effects through the inhibition of angiogenesis and tumor necrosis factor (TNF)-α. Its most notable toxicity is as a potent teratogen but it is also associated with peripheral neuropathy. Currently only approved for the treatment of erythema nodosum, studies are looking into its use in RA, SLE, Sjogren's syndrome, and AS (73). Dapsone, an antimicrobial agent, is believed to affect the neutrophil function by decreasing recruitment, chemotaxis, and inhibiting neutrophil function (74). These drugs cause marrow suppression and GI intolerance. A review of the therapeutic application of these drugs is presented in Table 40-12.
Gold
Parenteral gold, and more recently oral gold (auranofin), has been used in the treatment of synovitis in patients with RA. Gold is thought to work by inhibiting lysosomal enzymes or by inhibiting phagocytic activity in macrophages and polymorphonuclear leukocytes. It also inhibits aggregation of human γ globulin in vitro, a phenomenon that is thought to be an inflammatory antigenic stimulus in RA. These events have been observed when parenteral gold is used. The oral preparation alters all mediated immunity, inhibits DNA synthesis in vitro, and suppresses humoral immunity (75).

Adverse effects of gold compounds include rash, stomatitis, proteinuria, and hematologic disorders (e.g., leukopenia, thrombocytopenia). These side effects, combined with lack of long-term clinical efficacy, have led to decreasing use of the agent. This drug is not used in patients with SLE, partly because it may flare the skin involvement. It may be useful in treating patients with peripheral arthritis associated with psoriasis.

Antihyperuricemic Agents
Pain and inflammation of crystal-induced arthritis are frequently adequately controlled with NSAIDs. Although these drugs are effective in controlling symptoms, they do not alter the metabolism of the substances forming crystals, nor do they influence their excretion. Uricosuric agents like probenecid compete with the tubular transport mechanism for uric acid, reduce the reabsorption of uric acid, and hence increase its excretion (76). Their use is widespread, and their toxicities are well known, including nephrolithiasis, which is preventable if the urine is alkalined and fluids are increased. Acute gout can be precipitated as the uric acid levels are lowered and GI symptoms are not infrequently seen. A second approach toward controlling serum urate levels is that of regulating production of uric acid by inhibiting xanthine oxidase. This is done by using allopurinol as an analog of hypoxanthine. It too can precipitate an acute attack of gout and can cause xanthine renal stones. Side effects include rash and, rarely, blood dyscrasia. Allopurinol should not be used with azathioprine. Allopurinol is an inhibitor of the principal pathway for the detoxification of azathioprine.

Biologic Agents-Anticytokine Therapies
Etanercept (administered subcutaneously), infliximab (administered intravenously), and adalimumab (administered subcutaneously) function to inhibit TNF and have gained increasing popularity, as studies have shown these to be as effective, if not more effective, in preventing joint damage than MTX (77). Infliximab and adalimumab are monoclonal antibodies and bind to TNF-α. Etanercept binds both to TNF-α and lymphotoxin α, neutralizing their biologic activity. It is administered in doses of 25 mg twice weekly or 50 mg once weekly. It has been shown to be effective in RA (78), AS (79), and PSA (80). The most significant side effects include injection site reactions and infections, and to a lesser extent development of autoantibodies. There is also some risk of lymphoproliferative disorders and more rarely lupus-like reactions, demyelinating disorders, as well as other malignancies (81).

Anakinra (administered subcutaneously) is an anti-interleukin-1 receptor antagonist, currently approved for the treatment of RA. Doses of 1 and 2 mg/kg were shown to make improvements in joint counts, pain scores, morning stiffness, and physician assessment of disease activity (82).

Other Biologic Agents
Two of the newer biologic agents include rituximab and abatacept. Rituximab is a monoclonal antibody directed against the extracellular domain of the CD20 antigen on B-Cell and initiates compliment-mediated B-cell lysis. Abatacept binds to CD80/CD86 on the surface of antigen presenting cells and inhibits T-cell activation. These agents continue to be studied but have shown significant promise for the treatment of RA when used in combination with MTX (83,84).

Combination Therapies
Along with a shift of the classic treatment paradigm for RDs by using DMARDs earlier in the course of disease, it has become increasingly more common to use these agents in combination. The benefits of using some of these combination therapies include additional therapeutic effects of drugs that may not have been fully effective as monotherapy and possible improvement in the toxicity profile, as lower doses of these potentially toxic agents are needed. MTX is commonly used in combination with other DMARDs. Often used combinations include the following:

1. Cyclosporine and MTX have shown to be very safe and useful in the treatment of severe RA (85).
2. MTX, HCQ, and SSA combination therapy has shown some promise with moderate efficacy and no noted increased toxicity when compared with MTX alone or SSA/HCQ in combination (86).
3. MTX and TNF-α blocking agent infliximab showed good overall drug tolerance and sustained efficacy over a 2-year period (87).
4. MTX and leflunomide as a combination has shown some associated drug toxicity but a good overall efficacy (88).

Many other combinations of these agents have been evaluated. Some have shown promise but others have been marred by intolerable side effects (89).

Complementary and Alternative Medicine
The National Institutes of Health defines complementary and alternative medicine (CAM) as encompassing those treatments and health care practices that are not widely taught in medical schools, not generally used in hospitals, and not usually reimbursed by medical insurance companies. These therapies are sometimes called unconventional therapies, since they are outside the “mainstream of Western medicine.” Often these therapies have not undergone rigorous scientific analysis in randomized controlled trials (RCTs) (89–92). More medical schools have now introduced CAM into their curriculum (93).
The term *alternative* alone has been used to refer to practices used in place of mainstream Western medicine, such as exclusive use of herbs instead of prescription drugs. The Arthritis Foundation (AF) prefers the term *complementary medicine* for the use of unconventional therapies in arthritis that are used to support mainstream Western medicine.

There has been an increasing use of CAM in the general population to treat disease. One recent study (94) indicated that 42% of patients use at least one of 16 CAM therapies, and 50% of these patients had musculoskeletal disease. The study indicated that CAM is used as an adjunct to conventional therapies. Visits to CAM practitioners exceeded those to primary care physicians, and patient expenditure was $27 billion/year, similar to expenditures for all U.S. physician services.

Because of the chronic painful nature of RDs and the fact that they often are not cured by Western medicine, arthritis patients often seek CAM treatments (95). A 1998 review of surveys indicates use by 30% to 100% of arthritis patients. A 2004 study showed 90% of primary care clinic patients with arthritis used CAM, and RA patients used an average of 4.4 CAM therapies (96). A 2004 study revealed 33% of JIA patients use CAM as parents sought it for pain relief for their children (97). However, only 38% to 55% of patients reveal their use of CAM to their physicians (95). A more recent study cited 70% of RA patients revealed their CAM use to physicians (96). Physicians should ask patients about their use of CAM, as some treatments are contraindicated when used with conventional therapies (98). Arthritis patients and physicians have different perceptions about the usefulness of CAM. Arthritis patients generally perceive its usefulness (99). In a group of mixed arthritis patients in a 2005 review, homeopathy and acupuncture were the most used type of CAM (44% and 41%, respectively). Significantly higher self-perceived efficacy scores were seen for CAM use in patients with spondyloarthropathies and OA. The lowest scores were among RA and connective tissue diseases (100). Some rheumatologists do not recommend them (101) and others do (102). The general trend is toward more approval by health care providers.

A number of practitioners and treatments are included under CAM (Table 40-13). The main categories are alternative healing systems; mind, body, and spirit treatments; prayer and spirituality; moving medicine; massage and touch; herbs and supplements; and miscellaneous treatments. In one study, RA patients most commonly used relaxation, glucosamine, and vitamin C. There was less common use of fish oil and gamma-linolenic acid (GLA) containing supplements. Patients used therapies to relieve pain, prevent disease progression, and to feel better (96).

The benefit of some CAM therapies has been researched, while others have been minimally studied. Much of the older research consists of nonrandomized, noncontrolled trials. The issue of efficacy of popular CAM therapies (acupuncture, herbs, or homeopathy) used by the general population has been explored in recent meta-analysis reviews of controlled trials (103–105). The publications listed in references (72) and (87–94) are available for review of CAM benefits. Most of the literature refers to neck and back pain, fibromyalgia (FM), and OA. The literature is sparse on the benefits of CAM in inflammatory arthritis.

There continues to be a growing interest in the use of herbal therapies in RA. A systematic review of RCTs in this area resulted in 14 such trials. There was moderate support for GLA for reducing pain, joint count, and stiffness. Further research is needed to examine the safety and efficacy of herbal remedies (106).

A 2005 review summarizes the efficacy and toxicities of herbal remedies used in CAM therapies for RDs. It elucidated the immune pathways through which they have anti-inflammatory and/or immunomodulatory activity that may provide a scientific basis for efficacy. For instance, GLA acts

### TABLE 40.13 Alternative and Complementary Medicine

<table>
<thead>
<tr>
<th>Alternative healing systems</th>
<th>Ayurveda</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chinese medicine</td>
<td>----------</td>
</tr>
<tr>
<td>Naturopathic medicine</td>
<td>----------</td>
</tr>
<tr>
<td>Chiropractic</td>
<td>----------</td>
</tr>
<tr>
<td>Homeopathy</td>
<td>----------</td>
</tr>
<tr>
<td>Folk and self-help remedies</td>
<td>----------</td>
</tr>
<tr>
<td>Hydrotherapy, balneotherapy</td>
<td>----------</td>
</tr>
<tr>
<td>Mind, body, and spirit</td>
<td>----------</td>
</tr>
<tr>
<td>Meditation</td>
<td>----------</td>
</tr>
<tr>
<td>Biofeedback</td>
<td>----------</td>
</tr>
<tr>
<td>Relaxation exercise</td>
<td>----------</td>
</tr>
<tr>
<td>Stress reduction and relaxation</td>
<td></td>
</tr>
<tr>
<td>Hypnosis</td>
<td>----------</td>
</tr>
<tr>
<td>Visualization and guided imagery</td>
<td></td>
</tr>
<tr>
<td>Prayer and spirituality</td>
<td>----------</td>
</tr>
<tr>
<td>Moving medicine</td>
<td>----------</td>
</tr>
<tr>
<td>Yoga</td>
<td>----------</td>
</tr>
<tr>
<td>T’ai chi chuan</td>
<td>----------</td>
</tr>
<tr>
<td>Qigong</td>
<td>----------</td>
</tr>
<tr>
<td>Alexander, Feldenkrais, Thager</td>
<td></td>
</tr>
<tr>
<td>Massage, touch</td>
<td>----------</td>
</tr>
<tr>
<td>Myofascial release, trigger point massage therapy</td>
<td></td>
</tr>
<tr>
<td>Deep-tissue massage, acupressure</td>
<td></td>
</tr>
<tr>
<td>Rolfing, polarity therapy</td>
<td>----------</td>
</tr>
<tr>
<td>Therapeutic touch, Reike</td>
<td>----------</td>
</tr>
<tr>
<td>Acupuncture/acupressure</td>
<td>----------</td>
</tr>
<tr>
<td>Miscellaneous therapies</td>
<td>Bee venom, copper bracelets, laser, magnets, pulsed electromagnetic therapy (PEMF)</td>
</tr>
<tr>
<td>Diet</td>
<td>----------</td>
</tr>
<tr>
<td>Weight reduction</td>
<td>----------</td>
</tr>
<tr>
<td>Fasting</td>
<td>----------</td>
</tr>
<tr>
<td>Vegetarian, vegan</td>
<td>----------</td>
</tr>
<tr>
<td>Fish oils</td>
<td>----------</td>
</tr>
<tr>
<td>Herbs and supplements</td>
<td>----------</td>
</tr>
</tbody>
</table>
as a competitive inhibitor of PG E₂ and leukotrienes (LTs). It appears to be efficacious in RA (107).

Proven benefits in noninflammatory arthritis include the following: (a) glucosamine/chondroitin decreases knee pain and slows disease progression in OA (108) and (b) acupuncture significantly decreases knee pain in OA (109).

Analysis of 17 acupuncture trials in 1997 (110) failed to show benefits of acupuncture in RA, SLE, AS, and PSS. AF reviews confirm this (91). In a 2005 systematic review of studies on acupuncture and electroacupuncture for RA, the reviews concluded that although electroacupuncture showed significant knee pain reduction 24 hours and 4 months posttreatment, the poor quality of the trial, including small sample size, precluded its recommendation. They further conclude from the studies reviewed that acupuncture has no significant effect on ESR, CRP, pain, patient global assessment, number of tender joints, general health, disease activity, and reduction of pain medications (111). A recent pilot study of two treatment groups of SLE patients (acupuncture or minimal needling) versus usual care revealed a more than 30% improvement in pain measures in the acupuncture and minimal needling groups with no improvement in the control-usual care group (112).

Other CAM modalities/treatments have also been studied. Low-level laser treatment decreased pain in RA in a Dutch study (113). A cognitive behavioral intervention (biofeedback) resulted in a decreased number of clinic visits and hospital days, and medical costs in RA (114).

A yoga program based on upper-body posture flexibility; correct hand, wrist, arm, and shoulder alignment; and stretching provided significant reduction in pain and increased grip strength in Carpal Tunnel syndrome (a common problem in RA) (115). An NIH consensus conference in 1998 concluded that acupuncture was useful for Raynaud’s syndrome (a common problem in SLE, MCTD, PSS) (116). A study by Yocum indicated that biofeedback increased the fingertip temperature in Raynaud’s syndrome in SLE and PSS (117). Intercessory prayer in one uncontrolled study significantly decreased the number of swollen joints, pain, and disability in RA patients (118). Tai chi is safe in RA but benefits are not proven. A 2007 systematic review of 45 studies of tai chi for RA, found only two RCTs and three non-randomized controlled clinical trials (CCTs) meeting the Jadad score for methodological quality. The RCTs demonstrated some positive findings for tai chi on disability index quality of life, depression, and mood but not on pain reduction. It was concluded by the author that collectively, at the current time, evidence is not convincing enough to suggest tai chi is an effective treatment for RA (119). A 1997 study indicated massage decreased pain and joint stiffness in JIA (120).

In general, diet can influence gout, but other than that, there is no definite evidence that diet can cure arthritis (72). There is some suggestive evidence that a decrease of omega fatty acids and substituting omega-3 oils may decrease pain and inflammation. Sources of omega-3 oils include cold-water oily fish, sardines, green soybeans, tofu, canola, and olive oils (89).

Most CAM therapies are low risk, but some do involve risks (89). Herbs can interfere with prescription medication. The following increased sensitivity to anticoagulants: bromelain, chondroitin, fish oil, GLA, garlic, ginkgo, ginger, ginseng, evening primrose oil. Folic acid interferes with MTX. Ginger can increase NSAID effects. Ginseng may increase the effects of glucocorticoids and estrogens, and should not be used in diabetes mellitus or with monoamine oxidase (MAO) inhibitors. Kava Kava increases the effects of alcohol, sedatives, and tranquilizers. Magnesium may interfere with blood pressure medication. St. John’s Wort enhances the effects of narcotics, alcohol, and antidepressants, and increases the risk of sunburn and can interfere with iron absorption. Valerian increases sedative effects. Zinc interferes with glucocorticoid and immuno-suppressive drugs (89).

Caution should be noted in patients with inflammatory arthritis receiving manipulation therapy. These patients often have damaged joints that can sublux (RA, JIA) (121) and ligamentous laxity (122). Patients with significant AS have rigid spines and can fracture. Those with moderate to severe OP from disease and steroids can also fracture. Spinal fractures can result in neurologic compromise. Patients with RA have C1-2 laxity or instability, and can sublux with neurologic compromise.

Precautionary advice in CAM administration includes use of only sterile disposable needles, and prohibiting the use of pulse electromagnetic field therapy in pregnant or cancer patients, and magnets in those with implanted electronic devices or electric blankets.

**Surgery: Soft-Tissue and Reconstructive Procedures**

Indications for surgery in the RDs include the restoration or preservation of joint mechanics and function, and relief of pain. In general, pain relief is a more predictable outcome in arthritis surgery than is functional restoration. Contributors to functional outcome include motor strength, motivation, postoperative complications, and participation in rehabilitation, all of which are highly variable.

The decision to operate on a patient with RD requires a thorough preoperative evaluation. This must include the overall health status of the individual and identification of the medications that might increase the risk of surgical complications. This population is likely to be more than 55 years of age, have an altered immune system, and be receiving medication that could influence healing. Examples of these medications include steroids, NSAIDs, MTX, and the newer cytokine inhibitors. Information about disease status and medication will help illuminate any additional risks for surgery.

Equally important is to perform a comprehensive physical examination and review of x-rays. The decision about which procedure to perform depends on properly identifying the cause of symptoms and the ability to determine whether they are likely to be correctable by the proposed procedure. Efforts to realign joints or soft tissue, although important, should be considered within the context of functional needs.
and symptom control. For example, pain is usually the result of joint deformity and its sequelae, but nerve entrapment, referred pain, and myopathy might need to be ruled out before surgery.

Proper identification of potential risk factors for anesthesia, surgery, or postoperative course is also important. For example, C1-2 subluxation poses a significant risk for intubation. Carious teeth may increase the likelihood of developing postoperative infection. Obesity may make rehabilitation difficult and compromise the long-term outcome of surgery. These problems would need to be treated or accommodated before surgery.

Timing of surgery may be critical to outcome. For example, shoulder replacement in a patient who has had longstanding ROM deficits, in whom rotator cuff function is limited or absent, will have a poorer functional outcome than someone whose rotator cuff is still working. When possible, surgery should be considered before the development of significant joint contracture, muscle atrophy, and instability.

The surgical procedures relevant for joint and soft-tissue management include synovectomy and joint debridement, tendon repair and realignment, osteotomy, arthrodesis, and arthroplasty. Each procedure has specific indications. Much has been written about the success of these procedures, their life expectancy, and long- and short-term complications. Rehabilitation professionals can assist in preparing patients for optimal outcomes by assisting patients to achieve a higher preoperative level of fitness, helping set realistic expectations, and educating them about health and function-promoting behaviors. These might include developing a strengthening program for the hip girdle muscles for those undergoing hip surgeries, encouraging the use of a cane until the hip abductor is adequately strengthened and Trendelenburg gait is eliminated, and eliminating leg length discrepancies (123).

**Synovectomies**

Synovectomies were first performed by Volkmann in 1877 for tuberculosis of the knee. Today they are sometimes performed on RA patients, most commonly to relieve pain and inflammation associated with chronic swelling uncontrolled by medication; to retard the progression of joint destruction, which is a controversial issue; and to prevent and retard tendon rupture. Other indications include the alleviation of decreased ROM caused by hypertrophied synovial tissue and denervation effect with reduced pain and inflammatory response.

Synovectomies are usually performed on the knee and wrist, and may be performed by arthroscopy or arthroscopically. They are most frequently done in hemophilic arthropathy, for pigmented villonodular synovitis or early RA.

Tenosynovectomy is most frequently performed for the extensor tendons of the hand. Regrowth of synovium commonly occurs postoperatively, so the procedure is not a curative one. Often local management using intra-articular injections of long-acting corticosteroids is tried, along with splinting of the joint when feasible and education to help develop alternative strategies that prevent overuse (124,125).

**Tendon Surgery**

Tendon surgery is common in inflammatory disease. Frequent indications for surgery include repair of ruptured extensor tendons, realignment of tendons of the hand, synovectomies for tendons with severe tenosynovitis, and reanastomosis following tendon ruptures (Achilles and patellar tendons) and tendon releases for intrinsic tightness (126).

**Arthrodesis**

Arthrodesis is performed less often today than in the past because of the popularity and success of joint replacement. It may still be the best procedure to eradicate resistant infection that has destroyed significant bone. The stability provided by an arthrodesis should be permanent. Adolescents and young adults with many more years of activity might well be considered for an arthrodesis in selected instances rather than a joint replacement, which often does not stand the stress placed on it by a young, vigorous patient. Arthrodesis for patients with arthritis is usually limited to the wrist, interphalangeal (IP) joints of the hand, first metacarpal phalangeal (MCP) joint, subtalar joints, and vertebral bodies. The triple arthrodesis remains one of the best procedures for reconstructing the hind-foot and restoring a pain-free, functional foot. Postoperative rehabilitation for this procedure requires 4 to 6 weeks of non-weight bearing, for which a rollabout can be prescribed for mobility. The rollabout is an ambulation aid, similar to a scooter, mounted 22 in. above the floor on four small wheels. It has a handle, a hand brake, and a padded shelf on which the leg is placed at 90 degrees of knee flexion. It is propelled by the nonoperative lower extremity and permits a reasonably rapid ambulation speed. The next level of independence is the cast boot and then a shoe with a custom insert and a rocker sole that assists in push off. Somewhat more controversial is knee arthrodesis, a procedure that is rarely done but is occasionally suggested for the very young and highly mobile patient.

Common indications for arthrodesis of a joint are to relieve persistent pain, to provide stability where there is mechanical destruction of a joint, and to halt progress of the disease (e.g., infection, RA). Joints should be fused in optimal functional position (127). Contraindications for arthrodesis include significant bilateral joint disease. Joint replacement is indicated more in this instance and arthrodesis of the same joint on the contralateral side.

**Joint Replacements**

**Upper Extremity**

Upper extremity joint replacements have become more common today (128). Patients with RA, JIA, OA, and AN in SLE may require joint replacement. Common indications for replacement are persistent pain despite adequate medical and rehabilitative management, loss of critical motion in the involved joint, and loss of functional status.

The main contraindications for joint replacement are inadequate bone stock and periarticular support, serious
medical risk factors, and presence of significant infection. Other contraindications include lack of patient motivation to cooperate in a postoperative rehabilitation program and inability of the procedure to increase the patient’s total functional level.

Wrist arthroplasty is recommended for those with adequate bone stock, who have relatively low use requirements. Loosening over time is common (129). MCP arthroplasty is performed for the preservation of function due to the relative frequency of subluxation and progression to dislocation in RA. Surgery performed before MCP dislocation usually has a better outcome (130).

Elbow surgery is usually restricted to radial head excision and arthroscopic synovectomy. Results from these procedures deteriorate over time (131). Elbow replacements have been shown to be effective in reducing pain and in improving ROM in pronation/supination, though improvement in flexion/extension is modest (132). Total elbow replacement has been recently used for patients with inflammatory arthritis with good success. Patients have noted substantial reduction in pain and improved functional ROM (133). Excision of the radial head, however, remains one of the best procedures for pain reduction and improvement in elbow ROM.

Shoulder arthroplasty has been shown to be beneficial in relieving pain. Older patients have better function and longer-lasting results than younger patients. Those with rotator cuff tears have 33% to 50% return of ROM following surgery, which is approximately half of those without significant tears that undergo arthroplasty (134). Long-term reports suggest that it is also associated with good functional outcome when rotator cuff function is intact (135). The shoulder is stiff, but flexion and abduction can be performed (<50 degrees) using scapulothoracic movement.

Indication for fusion of the cervical spine in patients with RA remains somewhat controversial. There is agreement that pain, unresponsive to nonsurgical treatment, cord compression, or peripheral sensory and/or motor loss are indications for cervical spine stabilization and/or cord decompression. Some studies suggest that early intervention is associated with better neurologic outcome. Instability of more than 10 mm at the atlantoaxial joint, or greater than 4 mm of basilar invagination suggests the need for spinal stabilization.

Lower Extremity

Total hip replacement surgery has been performed in the United States for more than 30 years. More than 120,000 hip replacements are done annually, and function remains good 25 years after surgery (136). Hip surgeries are no longer limited to patients more than 60 years of age. Infection rates have been dramatically reduced to less than 1% (137). Loosening of the prosthesis is the reason for long-term failure. The acetabular component is more likely to loosen than the femoral component, even in younger patients (138). Uncommented acetabular and femoral components may offer greater prosthetic durability.

Hip replacement surgery offers patients with RA, SLE (avascular necrosis), and AS pain relief and improved function. The decision about whether to offer a cemented or noncemented prosthesis is usually based on the age and the functional requirements of the population to be treated. The older (>70 years of age) patients most frequently receive a cemented prosthesis, which provides good, immediate stability. The noncemented prosthesis is associated with better preservation of bone but may be associated with persistent thigh pain.

Patients are given prophylactic antibiotics before surgery and low–molecular-weight heparin the night of surgery and for the duration of the hospital stay. Antibiotic prophylaxis is recommended for dental work. A good discussion by Sledge of the operative and postoperative course is recommended to the reader (139).

Total knee replacement surgery is commonly used for patients with bi- and unicompartmental joint space destruction, persistent pain from poor joint mechanics, and functional loss. Long-term complications tend to result from uneven patellar surface wear and loosening. Total knee arthroplasty has provided excellent pain relief and good functional outcomes for arthritics. Studies report prosthetic longevity with sustained, excellent function for more than 12 years (140). Problems with the patellar components are the most significant cause of knee joint replacement failure.

Ankle replacement arthroplasties have not been demonstrated to be effective over time. Loosening remains the most serious complication (141). Those patients with very limited mobility who require ambulatory function may be reasonable candidates for this procedure. Forefoot arthroplasty with total resection of the metatarsal heads is an excellent pain-relieving procedure. This enables patients to walk on a pain-free foot, although the toes become floppy, the foot size is smaller, and the mechanics of pushoff, which are usually improved from the painful state, are not returned to normal. Use of a roller sole helps correct the dynamic abnormality (142).

Preoperative Rehabilitation Management

To maximize postoperative gains, preoperative rehabilitation interventions are desirable. These interventions include teaching the patient crutch walking with the appropriate type of crutch; weight reduction for the obese patient; and strengthening of the quadriceps before knee replacement and the hip abductors before hip surgery. Orienting patients to the types of pain they may experience postoperatively—such as acute, incisional; muscle strain or fatigue; and nerve root irritation—may help allay fears about the stability of the hip. Descriptions of the usual course of recovery may also prepare them for what to expect.

Postoperative Rehabilitation Management

The rehabilitation management goals of a total joint replacement program are to relieve pain, to redevelop comfortable musculoskeletal function, and to use joint protection techniques to avoid overstressing the prosthetic joint. The postoperative management of a hip replacement is individualized...
based on the preferences of the orthopedic surgeon and the needs of the patient. However, the program usually includes bed mobility. ROM is started immediately with ankle pumps, and isometric exercise to the quadriceps. The patients are usually made to stand by the bedside with full weight bearing crutches if the hip is cemented, or partial weight bearing if it is uncemented. Patients are placed in an abduction sling and told to restrict hip flexion to less than 90 degrees and to limit adduction and internal rotation (IR). While in bed, they are to sleep in a supine position with a pillow between their knees for a month. They need to be carefully monitored for signs of deep venous thrombosis, fever, excessive wound drainage, and/or infection. Patients should be instructed to use an elevated toilet seat and an elevated chair seat to minimize hip flexion. Discharge from the hospital is usually on the fifth day, provided they can get in and out of bed independently, walk independently with crutches or walker, and manage stairs. Key exercises include quadriceps, hip abductor, and hip flexor strengthening. Patients should expect to use a cane until hip abductor strength is in the four (based on the medical research council [MRC] scale) range and there is no Trendelenburg sign. Many orthopedists permit return to full activity, including recreational tennis, cycling, and gardening (143).

The postoperative management of knee replacement includes the following procedures: Begin knee ROM immediately postoperatively, often with the aid of a continuous passive motion machine. Total weight bearing (to tolerance) with crutches and ad lib ambulation is started on the first postoperative day using crutches or a walker. Active assistive flexion is the cornerstone of management and usually needs to be done under supervision of the physical therapist. Knee replacement patients, unlike those undergoing hip replacement, frequently need some additional rehabilitation requiring admission to a rehabilitation center (skilled nursing level or if accompanied by significant comorbidities, an inpatient rehabilitation facility [IRF]).

The extent of rotator cuff repair and function in part dictates the nature of the post-shoulder arthroplasty rehabilitation program, but the postoperative management of total shoulder replacement usually includes the following: Immobilization of the shoulder for 2 to 8 weeks in an airplane splint with the shoulder in 80 degrees of flexion, 70 degrees of abduction, and 5 degrees of IR. Passive motion through range, in excess of where the limb is in the splint, is performed in the supine position. At the eighth to tenth postoperative day, active, assistive shoulder exercise is begun in the sitting position to 110 degrees of flexion and 20 degrees of external rotation (ER). At 6 weeks after surgery, active, unrestricted ROM is permitted, sometimes using an overhead pulley to assist in end range. Lifting up to 10 lb is permitted (144).

**REHABILITATION INTERVENTIONS**

Rehabilitation treatment plans must be individualized for the patient’s needs; they should be practical, economical, and valued by the patient to enhance compliance. Treatment should begin early in the disease process to help prevent impairment and functional decline so that the patient identifies this as part of the overall management plan. There is scientific and clinical rationale for the use of some specific rehabilitation treatments: others are based on clinical judgment. Rehabilitative rheumatology treatments and techniques must be monitored carefully, and periodic reevaluation of the patient with adjustments in treatment should be made.

**Rest**

Three forms of rest have been used by persons with arthritis: complete bed rest, local rest of a joint or joints with splints or casts, and short rest periods of 15 to 30 minutes dispersed throughout the day.

**Systemic Rest**

In the 1960s and 1970s, the literature revealed studies supporting bed rest for up to 4 weeks for persons with RA to decrease the number of inflamed joints, joint stiffness, and disease activity. However, systemic rest has many adverse effects, including muscle weakness and bone loss. Currently, the approach to the management of RA has changed. Much more adequate pharmacologic management of disease activity exists, such as early treatment with DMARDs. In addition, the literature clearly supports mobilizing and exercising patients with inactive and subacute inflammatory arthritis and encouraging them to be proactive in maintaining fitness and healthy lifestyle behaviors through exercise and symptom control throughout the disease course.

Depending on the disease severity, shorter periods of complete bed rest than formerly used may be needed. Several days of bed rest may be indicated for severe new-onset RA, SLE, and PM and for severe acute flares during the course of these diseases when intravenous medication may be needed.

**Local Rest**

Local rest of acutely or subacutely inflamed joints at night with nonfunctional resting splints and during the day with functional splints reduces inflammation and pain, and may help prevent contracture. Immobilization of the wrist for painful periarticular syndromes (e.g., de Quervain’s syndrome and carpal tunnel syndrome) is useful to relieve pain. One ROM exercise daily for joints during rest of 2 weeks’ duration is not noted to cause an adverse effect. In terms of muscle effect, 4 weeks of knee immobilization in nonarthritic population causes a 21% decrease in muscle mass determined by computerized tomography and biopsy (145).

**Short Rest Periods**

Provision of short rest periods during the day of 20 to 30 minutes along with appropriate local splinting is considered the appropriate way to manage patients with inflammatory arthritides to help control joint inflammation and fatigue. Some workplace sites now provide rest areas for persons, complete with cots. Negotiation with employers may also result in a person with an office being allowed to keep a small cot or sofa for napping.
Exercise

Arthritis commonly produces decreased biomechanical integrity of joints and their surrounding structures, which results in decreased joint motion, muscle atrophy, weakness, joint effusion, pain, instability, energy inefficient gait patterns, and altered joint loading responses (146,147).

Arthritis patients may lose muscle strength and bulk because of inactivity. A muscle can lose 30% of its bulk in a week and up to 5% of its strength a day when maintained at strict bed rest (148). Other factors contributing to loss of strength are myositis, myopathy secondary to steroids (149), inhibition of muscle contraction due to joint effusion (150), and direct effects of the disease itself on the muscle. For example, in RA, some destruction of muscle fibers occurs as well as intermuscular and perimuscular adhesions, which may impair blood flow. Muscle fascicles adhere to one another, and the entire muscle may adhere to the intermuscular septum and perimuscular fascia, causing inhibition of muscle contraction and normal movement (146) with RA, PM, SLE, and PSS. This results in weak, painful, and easily fatigable muscle. Reduced strength, as determined by isometric testing, has been documented in the quadriceps even in mild RA (151,152), PM (153), and JIA (154). Isokinetic strength testing has shown deficits in the quadriceps in patients with RA (151) and DM-PM (155). Patients with RA, AS, SLE, JIA, DM, and JDM have been found to have decreased aerobic capacity (156–161).

The biomechanical advantage of joints is compromised by weak muscles. Normally, muscles function to provide postural stability and distribute forces of impact and stress across joints during activity. Normal joint function requires that muscles contract and relax synchronously. Atrophic muscles around joints do not coordinate well and are deficient in both static endurance and strength. There is decreased tone and increased spasm in muscle surrounding arthritic joints, resulting in less coordinated motion of the joint (146). Exercise programs for patients with arthritis have been shown to produce a variety of benefits that include:

- Increase and maintain ROM (162).
- Reeducate and strengthen muscles (163).
- Increase static and dynamic endurance (156,164–166).
- Decrease the number of swollen joints (166).
- Enable joints to function better biomechanically (146).
- Increase locomotor ability (165,167).
- Increase bone density (165).
- Decrease pain (168).
- Increase the patient’s overall function and well-being (168–170).
- Decrease inflammation (171–173).
- Increase aerobic capacity (159,169,174,175).
- Reduce body weight (176).
- Limit rheumatoid cachexia (177,178).

Exercise prescriptions should be designed to improve function that patients value. Once these goals are set, limits need to be established that preserve joint function and do not unduly fatigue inflamed muscle and joint structures. Exercise should be performed with proper joint support, after reduction of joint effusion is accomplished, and attention should be given to the level of aerobic capacity (158). Patients with collagen diseases often have cardiac abnormalities, which should be evaluated before initiating an exercise program and on an intermittent basis while on a maintenance exercise program (179). Exercise programs for arthritis patients should specify whether the exercise should focus on aerobic activities, strengthening, and/or be aimed at building muscle endurance. Prescriptions should specify the muscles that need strengthening; the type, intensity, duration, and frequency of exercise to be used; and the specific precautions (158). It is helpful and generally felt to improve compliance if the patient is provided with a written exercise program and verbal directions if the program is to be performed at home. Patients who exercise as part of a group or with a partner are more likely to adhere to a given program. The patient or family, or both, should be told the purpose of the exercises (see Table 40-9).

An exercise program should be progressive. It should start with relieving pain of the involved joints with appropriate modalities, and/or pain medications. Once comfortable, a combination of stretching, strengthening, and fitness training should be initiated in a progressive manner. A program that shows results in terms of functional activity is best. Isometric exercise is usually the initial approach, with the addition of an isokinetic exercise program for muscle endurance and for strengthening if joints permit. Isometric low-resistive and low-weight progressive resistive exercise, as well as low-force isokinetic exercise can be used without joint damage. Incorporating exercise into recreational activities may be the treatment of choice for the patient because of its inherent appeal and the fact that it is done outside of a medical setting. Depending on joint integrity, a variety of options (i.e., gardening, swimming, mall walking, low-impact dancing, table tennis) are safe and effective for those with RDs (147).

Passive Exercise

Passive exercise is beneficial for patients with severe muscle weakness due to PM or neuropathic disease associated with stroke, peripheral neuropathy, and vasculitis. Patients with acute joint flares should passively or actively move the acute joint through the tolerable range once or twice a day, to prevent motion loss. Caution should be applied as passive exercise may also increase intra-articular pressure in the presence of joint effusion and has been associated with rupture of the joint capsule with large effusions (147).

Active Exercise

Active resistive exercise uses three types of muscle contraction:

1. Isometric or static contraction, which is highly suited for arthritis patients with mechanically deranged joints.
2. Isotonic or dynamic contraction, which is most suited for patients without acutely inflamed or biomechanically deranged joints, because it stresses the joint throughout its ROM.
3. Isokinetic dynamic contraction, which in most cases is not recommended for arthritis patients.
Strengthening Exercise

Strengthening of a muscle in arthritis patients may be achieved via isometric, isotonic, or isokinetic exercise. The degree of inflammation and disease stage should be evaluated before making an exercise choice. Isometric (i.e., static) exercise is ideally suited for restoring and maintaining strength in patients with muscle atrophy from RDs and steroids, and for the recovery phase of DM-PM, and for patients with significantly biomechanically compromised joints. Machover and Sapecy demonstrated a significant increase (27%) in quadriceps strength in patients with RA on an isometric strengthening program (180). This program consisted of three daily maximal contractions held for 6 seconds, with 20 seconds of rest between each. The knee was in 90 degrees of flexion. The opposite quadriceps had a cross-over effect with a 17% increase in strength. A similar program for a PM patient increased strength (1). To decrease force across a joint, a less than maximal contraction should be used. To effect strengthening this should not be less than two thirds of the maximal contraction. The patient should be instructed how to do this.

An advantage of isometrics is that muscle tension can be generated with minimal joint stress. Pain has been reported with maximal contraction. One study of adults by Gnoortved et al. indicates that isometric quadriceps exercise of inflamed knee joints in RA yields increased oxidative damage to hyaluronate and glucose, determined by analyzing synovial fluid 1 hour following exercise (181). Therefore, isometric exercise in an inflamed joint is not recommended. However, pediatric rheumatologists often recommend a few isometric contractions a day in children because loss of strength can be rapid around inflamed joints.

With isometric exercise, muscle strength is only achieved at the angle at which the muscle is trained. With isotonics, strength is achieved throughout the range. In addition, strength obtained by isometric training is not fully transferable to isotonics tasks. Therefore, the addition of isotonic exercise to the arthritis program is warranted where appropriate. DeLorme progressive resistive exercises with high loads and low repetitions build strength but put considerable stress across joints. A low-load resistive muscle training program in functional class II and III RA (12-week circuit weight bearing with light loads/high repetitions three times a week) resulted in significant improvement in self-reported joint count, the HAQ, grip strength, and knee extension strength (163). A 6-week high-intensity, progressive resistive strength training program (183) and home program (184). Strength increases were seen in PM patients on an isotonic resistive machine program (183) and home program (184).

Dynamic isotonic high-resistive exercise has the potential of exacerbating inflammation in general, which can increase muscle fatigue and joint pain and secondarily decreases joint ROM. Progressive resistive high-intensity isotonic exercise and isokinetic exercise are not recommended for those with arthritis. Moderate-intensity isotonic programs can be used in selected cases. Joints should be nonactive and biomechanically intact. Strength gains with isokinetic exercise do not exceed those obtained with low-weight, isotonic strengthening programs, and equipment is expensive and only available in a clinic (185).

Isokinetic testing of strength has been done in a 1990 study on RA patients with mild joint disease (186), and in PM patients (187) without deleterious effects. An isokinetic strengthening program used in RA patients increased strength. Complications included several joint flares and a ruptured Baker cyst (188). A 1994 study on a small group of RA patients showed that isokinetic strength training at four speeds for 3 weeks significantly increased strength without joint flares (189). A 1999 study revealed increased knee flexion torques at 60 and 90 degrees/second in RA patients on a knee flexion/extensor training program. The patients had nonacute RA (190). An isokinetic program with medium (120 degrees/seconds) velocity is most likely used in RA patients. Low velocities (30 to 90 degrees/seconds) produce high torque around joints and are best avoided. Isokinetic exercise should not be used in arthritic patients with joint effusion, Baker cyst, ligamentous laxity, acute joints, or joint replacements. A study of isokinetic exercise in six PM-DM patients demonstrated a significant increase in strength without significant creatine phosphokinase (CPK) increases (191). See Table 40-14 for specific exercise regimens.

Endurance

Patients with systemic RD have an overall limited endurance, and their ability to continue static or dynamic tasks is impaired. Endurance exercise can lead to an increased functional level in RA patients (192,193). In 1981, Nordemar described RA patients who were trained for 4 to 8 years on a bicycle ergometer at home and a self-suited exercise program consisting of jogging, skiing, swimming, and cycling (194). He found improved ADL performance in the exercised group as well as less progression of X-ray changes in arthritis, more improvement in hamstring strength, and less sick leave (195). Harkcom et al. (196) also reported benefit from aerobic exercise in RA patients. Minor et al. showed that aerobic exercises increase aerobic capacity in both RA and OA patients. Decreased joint counts for pain and swelling also have been associated with aerobic programs for RA patients (166,170,182). Exercise programs using different combinations of ROM, strengthening exercises, and aerobic exercises on land and in the pool have been beneficial in RA (147).
<table>
<thead>
<tr>
<th>Type of Exercise</th>
<th>Study</th>
<th>Patient Population</th>
<th>Duration</th>
<th>Frequency</th>
<th>Intensity</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strengthening (progressive resistive)</td>
<td>Rail et al. (182)</td>
<td>8 well-controlled RA 8 controls</td>
<td>12 wk</td>
<td>2×/wk</td>
<td>80% of maximum adjusted upward weekly</td>
<td>Significant increase in strength with decreased pain and fatigue</td>
</tr>
<tr>
<td>Moderate resistive (controls—ROM and stretch)</td>
<td>Hakkinen et al. (183)</td>
<td>35 recent onset RA &gt;2 yr RCS 35 controls</td>
<td>2 y</td>
<td>2×/wk</td>
<td>Loads of 50%–70% rep max (2 sets with 8–12 reps; 45 min each person)</td>
<td>Significant increase (19%–59%) in strength with decreased disease activity (HAQ) with increase in walk speed</td>
</tr>
<tr>
<td>Progressive resistive (machine)</td>
<td>Spector et al. (185)</td>
<td>10 RA with mild knee involvement 10 controls</td>
<td>12 wk</td>
<td>3×/wk</td>
<td>3 sets (10, 15, and 20 reps) 90-s rest each set 120° s 90° s 60° s PRE (3–4)</td>
<td>Significant torque increase in lower-extremity muscles (except RT quad)</td>
</tr>
<tr>
<td>Isokinetic (Lido) (controls—no exercise)</td>
<td>Leventhal et al. (188)</td>
<td>JIA (class II–III) RCS</td>
<td>6 wk</td>
<td>3×/wk</td>
<td>12–15 reps set in 30 s with 30-s rest period between sets 27–30-min session with 1–3-lb weights</td>
<td>Significant increase in extension, strength/(0.03) decreased joint count, decreased HAQ score; significant decrease in number of painful joints</td>
</tr>
<tr>
<td>Low load resistive weight-bearing exercise</td>
<td>Kumati Reddy et al. (163)</td>
<td>JIA (class II–III) RCS</td>
<td>12 wk</td>
<td>3×/wk</td>
<td>High to moderate muscle function (0.25–2 kg) weights; low muscle function (anti-gravity) ten reps per muscle group, total 15 min</td>
<td>Significantly improved function by (function index for myositis) SF-36</td>
</tr>
<tr>
<td>Isokinetic resistive plus walking</td>
<td>Alexanderson et al. (184)</td>
<td>11 active PM UC</td>
<td>12 wk</td>
<td>5×/wk</td>
<td>Alternating every 2 wk ROM with isokinetic exercise</td>
<td>60% increase in isokinetic peak torque</td>
</tr>
<tr>
<td>Isokinetic and ROM</td>
<td>Escalante et al. (191)</td>
<td>5 PM (active) No controls</td>
<td>2–7 mo</td>
<td>2×/wk</td>
<td>60% increase in isokinetic peak torque</td>
<td>Increase in knee flexion torque at 60 and 90 degrees; significant decrease in pain (P.O.C.) increase flexion (HAQ)</td>
</tr>
<tr>
<td>Isokinetic knee extension/flexion</td>
<td>McMeeken et al. (190)</td>
<td>18 nonacute RA 18 controls RCS</td>
<td>6 wk</td>
<td>3×/wk</td>
<td>70% increase of max speed four sets of five reps (1-min rest after each set) controls not stated</td>
<td>Significant strength increase; decrease in pain; no increase in disease activity or joint pain</td>
</tr>
<tr>
<td>Progressive resistive (high intensity isokinetic machines)</td>
<td>Rall et al. (182)</td>
<td>Exercise group 8 RA (well controlled) 8 healthy young 8 healthy elderly CS</td>
<td>12 wk</td>
<td>2×/wk</td>
<td>Concentric/eccentric contraction all MAS or muscle group 80% repetition; max two sets of 8 reps each</td>
<td>Significant strength increase; decrease in pain; no increase in disease activity or joint pain</td>
</tr>
</tbody>
</table>
### Table 40.14 Exercise Summary Chart (continued)

<table>
<thead>
<tr>
<th>Type of Exercise</th>
<th>Study</th>
<th>Patient Population</th>
<th>Duration</th>
<th>Frequency</th>
<th>Intensity</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Aerobic</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Walk aquatics Controls—ROM</td>
<td>Minor et al. (168)</td>
<td>40 RA</td>
<td>12 wk</td>
<td>3×/wk</td>
<td>1 h (30 min aerobic and warm-up and cool-down) at 60%–80% VO₂ max</td>
<td>Significant increase in aerobic capacity and walk time; decreased anxiety and depression</td>
</tr>
<tr>
<td></td>
<td>Wiesinger et al. (175)</td>
<td>14 PM (controls)</td>
<td>6 wk</td>
<td>2×/wk (first wk) 4×/wk (next 4 wk)</td>
<td>1 h 60% max heart rate (5-min warm-up, 5-min cool-down)</td>
<td>Significant increase in ADL and VO₂ max in study group; no increase in disease activity</td>
</tr>
<tr>
<td></td>
<td>Van den Ende et al. (197)</td>
<td>25 RA each of four groups</td>
<td>12 wk</td>
<td>High intensity 3×/wk</td>
<td>High intensity—70%–80% max heart rate—1 h</td>
<td>Intensive aerobic training with significant increase in aerobic capacity; strength and ROM no increase in disease activity; loss of effect 12 wk postexercise</td>
</tr>
<tr>
<td><strong>Group 1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full weight-bearing stationary bike</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Group 2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ROM/isometrics in a group</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Group 3</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ROM/isometrics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Group 4</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ROM/isometrics at home</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Dance-based aerobics</strong></td>
<td>Novreau et al. (211)</td>
<td>19 RA (class I–II)</td>
<td>12 wk</td>
<td>2×/wk</td>
<td>50% max heart rate (first 3 wk) 70% max heart rate (9 wk); (10-min warm-ups and 15–30 aerobic exercise)</td>
<td>Study group significant increase in aerobic capacity and walk time and peak torque</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10 controls</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Dance-based aerobic program</strong></td>
<td>Novreau et al. (212)</td>
<td>10 RA (class II–III)</td>
<td>8 wk</td>
<td>2×/wk</td>
<td>50% max heart rate (first 3 wk) 60%–70% max heart rate (last 2 wk) warm-up 10 min aerobic exercise (at 6 wk 25 min aerobic exercise)</td>
<td>No significant group aerobic capacity gain (four of ten subjects had 10%–20% aerobic gain); significant decrease in anxiety and depression; significant decrease in swollen joints; significant increase in walking, speech</td>
</tr>
</tbody>
</table>
### Low-impact aerobic dance

Moffet et al. (167)  
10 RA (class III) UCS  
50% VO\(_2\) max (first 3 wk)  
60%–70% max VO\(_2\) rest of program  
Moderate intensity aerobic  
7-min warm-up with 10-min aerobic and 2-min walk  
each of program up to 20-min aerobic recovery  
and cool-down (8 min)  
Significant improvements in isomotor ability

### Aerobic (supervised) walking

DeCarvalho et al. (304)  
41 SLE (training)  
19 SLE (controls)  
(SLEDAI scores eight or above)  
12 wk  
3×/wk (one supervised session two at home sessions)  
1 h (40 min aerobic walk at HR 10 min warm-up 10 min cool-down)  
Significant increase in VO\(_2\) max  
(P≤0.0001)  
Inter-group comparison:  
Significant difference in VO\(_2\) max (P≤0.01) and aerobic threshold (P= 0.0001)  
favoring aerobic group  
Significant improvement in Back (P≤0.001) and HAQ (P<0.01)  
Significant improvement in CHAQ  
Both groups (P≤0.0001) (no inter-group difference)

### Aerobic (high intensity) Dance and cardio-karate

Singh-Grewal et al. (352)  
40 JIA each of two groups (stable active)  
12 wk  
3×/wk (one supervised—two at home sessions)  
50 minute High Intensity  
Group 1 (30 min 75% MHR and 10 min warm-up and 10 min cool-down)  
Significant improvement in VO\(_2\) max  
Both groups (P≤0.0001) (no inter-group difference)

### Qigong (control group II)

Karapolat et al. (324)  
22 (AS) Group I exercise  
Group I exercise  
16 AS  
Group II home exercise  
8 wk  
7×/wk  
Control group II (below 75% MHR)  
45 min  
No significant change in VO\(_2\) submax in  
or between groups  
Significant improvement in BASDAI,  
BASMI, and NHP energy, pain, sleep  
and emotional sub-scores (P≤0.05) in both groups  
Inter-group comparison  
Significant difference in favor  
Group in the sleep NHP scores (P≤0.05)  
With-in groups:  
Group I (supervision)  
Significant improvement at 6 week and  
3 mo (P>0.05) of mobility CE (cm), MS,  
TWD (cm), MLS (cm) IMD (cm), FFD (cm), mean VO\(_2\) max BASFI  
Beck scores (6 wk only) (controls—no  
significant differences)

### Combined

Analay et al. (325)  
27 AS  
24 AS (controls)  
6 wk (plus  
3 mo follow-up)  
3×/wk  
50 min sessions  
Within groups:  
Group I (supervision)  
Significant improvement at 6 week and  
3 mo (P>0.05) of mobility CE (cm), MS,  
TWD (cm), MLS (cm) IMD (cm), FFD (cm), mean VO\(_2\) max BASFI  
Beck scores (6 wk only) (controls—no  
significant differences)

(continued)
<table>
<thead>
<tr>
<th>Type of Exercise</th>
<th>Study</th>
<th>Patient Population</th>
<th>Duration</th>
<th>Frequency</th>
<th>Intensity</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group II (at home)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Inter-group difference</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Significant improvement in Beck and BASDAI in favor of group I at 6 week</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>and 3 mo ($P = 0.05$)</td>
</tr>
<tr>
<td></td>
<td>Durmas et al. (326)</td>
<td>25 AS (exercise)</td>
<td>12wk</td>
<td>7×/wk</td>
<td>20 exercise</td>
<td>Significant improvement in BASFI, BASDAI, SF 36, MAF and BDI in both</td>
</tr>
<tr>
<td></td>
<td></td>
<td>19 AS (controls)</td>
<td></td>
<td></td>
<td></td>
<td>groups and between both groups in favor of the exercise group</td>
</tr>
<tr>
<td></td>
<td>RCS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Home-based exercise (HEP)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Within the exercise group</td>
</tr>
<tr>
<td>Combined (muscle flexibility and</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Significant increase in joint mobility, spine flexibility ($P \leq 0.0001$)</td>
</tr>
<tr>
<td>strength, breathing posture)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Control group</td>
</tr>
<tr>
<td>Control group (regular daily</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Significant decrease in knee flexion ($P = 0.05$)</td>
</tr>
<tr>
<td>activities)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Inter-group comparison:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Significant increase in FFD and joint mobility in exercise group ($P \leq 0.0001$) and function, pain and depression ($P \leq 0.001$) favoring the exercise group</td>
</tr>
<tr>
<td></td>
<td>Lim et al. (327)</td>
<td>25 AS (exercise)</td>
<td>8wk</td>
<td>7×/wk</td>
<td>30 min</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>25 AS (control)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>RCS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Strengthening/mobility</td>
<td>Group I (supervised-GPR method)</td>
<td>Fernandaz-de-las-Penas et al. (E15)</td>
<td>20 AS</td>
<td>4 mo</td>
<td>1×/wk</td>
<td>1 h sessions</td>
</tr>
<tr>
<td>------------------------</td>
<td>-------------------------------</td>
<td>--------------------------------------</td>
<td>-------</td>
<td>------</td>
<td>-------</td>
<td>-------------</td>
</tr>
<tr>
<td>Group II (conventional exercise, spinal flexibility)</td>
<td>RCS</td>
<td>ACR grade I–III 20 As controls</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>One year follow-up of strengthening/mobility</td>
<td>RCS</td>
<td>Fernandaz-de-las-Penas et al. (E16)</td>
<td>20 AS 20 AS (control)</td>
<td>4 mo</td>
<td>1×/wk</td>
<td>1 h sessions</td>
</tr>
<tr>
<td>Group I (supervised GPR method)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group II (conventional spinal flexibility exercise)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

RA, rheumatoid arthritis; AS, ankylosing spondylitis; RCS, random controlled study; NCS, noncontrolled study; CS, controlled study; Max, maximal; OA, osteoarthritis; PM, polymyositis; Reps, repetitions; SF, standard form; ADL, activities of daily living; IBM, inclusion body myositis; JIA, juvenile idiopathic arthritis; MEN, minute; ROM, range of motion; HAQ, health assessment questionnaire; Vs, versus; SLE, systemic lupus erythematosus; ACR, American College of Rheumatology; SLEDAI, systemic lupus erythematosus disease activity index; BDQ, back depression index; NHP, Notthingham health profile; MAS, myositis-specific antibodies; CHAQ, childhood health assessment questionnaire; SSA, spondylitis association of America; CE, chest expansion (cm); MS, BASMI, bath ankylosing spondylitis metrology index; BASFI, bath ankylosing spondylitis functional index; BASDAI, bath ankylosing spondylitis disease activity index; TWD, tragus-to-wall-distance (cm); FS, modified Schober; IMD, intermalleolar distance (cm); FFD, finger-to-floor distance; LF, lumbar spinal flexion; MLS, modified lumbar Schober; UC, uncontrolled; UCS, Uncontrolled Study; RC, randomized controlled; RT, right; PRE, Progressive Resistive Exercise; DAI, Disease Activity Index; HR, Heart Rate; MHR, Maximal Heart Rate; MAF, multidimensional assessment of fatigue.
Van den Ende in a 2000 Cochrane database study reviewed dynamic aerobic training and its effect on improving joint mobility, muscle strength, aerobic capacity, and function in RA patients. Negative effects such as increased pain, disease activity, and radiologic progression were also assessed. Selection criteria included randomized control trials on the effect of dynamic exercise therapy, an exercise program of at least 60% maximal heart rate for 20 minutes at least two times a week for 6 weeks. Only six of 30 controlled trials met criteria. It was concluded that dynamic exercise has a positive effect on physical capacity as measured by aerobic capacity and strength. Further research is needed on the positive effect on functional ability and radiologic progression (197). A RCT of dynamic training revealed RA women on low-dose steroid on a dynamic weight-bearing program showed positive effects on physical function and fitness level, and BMD with no disease exacerbation (198). Another trial reports the effect of intensive exercise in patients with active RA (161). A 2008 Cochrane database review summarizes dynamic exercise effect in RA (199).

Wiesinger reported improved physical fitness and muscle strength with short-term 6 week and long-term aerobic programs in DM-PM (175,186). Children with JIA can improve their aerobic endurance with a weight-bearing fitness program without disease exacerbation (171). SLE patients have been shown to increase their aerobic capacity by 20% on a bike aerobic program (156). A 12-week, 3×/week multimodal exercise program (aerobic stretching and pulmonary exercises) versus a medication only control group showed significant improvement in spinal mobility, work capacity, and chest expansion in the exercise group (200).

Bone mineralization is thought to be partially dependent on muscle contraction. Exercise has been shown to have a positive effect on bone mineralization in postmenopausal women (201). Patients with RD develop osteopenia from disuse, medication, and calcium and collagen metabolism abnormalities. Most studies that support the positive effects of exercise in these areas cite the use of isotonic and some resistive exercises. Sinali and Grubbs showed in a study that back extensor exercise can increase spinal bone density in postmenopausal women (202). This type of exercise may be useful for RD patients.

**Stretching Exercises**

Stretching may be used to prevent contractures and maintain or restore ROM by breaking capsular adhesions. These exercises must be graded according to the degree of inflammation, pain present, and pain tolerance of the patient. Heat may be used to increase collagen extensibility and cold to decrease pain before stretching exercises. Stretching to preserve or increase ROM should not be performed if there is acute inflammation, because it may increase it.

Active-assistive stretching can be used for maintaining or increasing ROM when the problem is subacute and pain is decreased. The patient initiates muscle contraction, and the therapist or an assistive device serves as an aid. Forceful stretching should be avoided in the presence of a large joint effusion because capsular rupture may occur.

Active stretching is performed in the absence of pain and inflammation to maintain ROM. It may be facilitated by the use of pulleys. Active stretching exercises in a pool are excellent. Devices may be needed to facilitate stretching for hip flexion contractures in JIA and knee flexion contractures in RA. Passive stretching has been found useful in increasing hip and shoulder motion in AS patients (203) (see Table 40-14).

**Aquatic Therapy**

The benefits of performing exercise in a pool include elimination of gravity and the positive effect of water buoyancy, which may result in decreased joint compression and pain (204,205). This may further result in increased muscle relaxation. In addition, a greater level of aerobic exercise may be tolerated in the water than on land. Therefore, therapeutic pool therapy may be most useful for those with moderate to severe arthritis, with recent joint replacement, with AS, and with any cardiopulmonary compromise.

Specifically, Danneskiold-Samsoe et al. have shown in a study on RA patients that isometric and isokinetic quadriceps strength can be increased by 38% and 16%, respectively, by adhering to a 2-month pool exercise program when compared with pretreatment values. A significant increase in aerobic capacity also can be obtained in RA patients on a pool program (206). AS patients with reduced vital capacity have been shown to be capable of undergoing pool therapy programs without untoward effect (207). In a RCT it was demonstrated that combined spa-exercise therapy is effective in patients with AS (208). Exercise tolerance appears to be related to pulmonary function in these patients (209).

**Recreational Exercise**

Patients with RDs often want to participate in recreational exercise programs (170). Care must be taken to advise the patient which activities or programs would be beneficial for him or her, and to relate the use of recreational exercise to the condition of the joints (i.e., inflamed, subacute, chronic mechanical derangement problems). In exercise gyms, the use of preset rate-limited devices at high torque speeds of muscle contraction against high resistive forces on machines should be avoided. Light weights (3 lb or less) and minimal repetitions (ten or less) on isotonic machines are permitted for patients with RA with no inflammation, minimal x-ray changes, and no ligamentous laxity. If isotonic weight lifting is performed, it generally should be with light weights (1 to 3 lb), and ten repetitions. A short arc of motion can be used to decrease joint pain. Swimming is an excellent form of isotonic exercise for arthritis patients because gravity is eliminated and ROM of the joints is less painful. ROM and stretching exercises, and pool jogging or walking are good. Local chapters of the AF have aquatic courses for arthritis patients and often make heated pools available. The Young Men’s Christian Association also has special pool exercise programs. Adaptive devices and special handgrips are available to help patients in specific sports (e.g., table tennis, golf, gardening, and bowling).
**Dance**
Dance has become a popular recreational activity for patients with arthritis. It can help increase joint motion, muscle strength, and aerobic capacity. Van Deusen and Harlowe describe the efficacy of a ROM dance program for adults with RA (210). Other, more formal and therapeutic dance programs, such as Educize, have shown increased strength, flexibility, and aerobic capacity along with decreased joint pain and depression (193). Formal dance-based aerobic programs for RA have shown positive effects: (a) changes in fatigue, tension, and aerobic capacity (211); (b) positive changes in fatigue, depression, and anxiety with increased aerobic capacity in some patients with class III disease (212); and (c) significantly improved locomotor ability by gait test in class III RA (213).

**Jogging**
Dry land jogging, which involves repetitive joint motion and offers little chance for increase in strength, is not recommended if arthritis of the knee, hip, or ankle is present. It is a good rule that a patient should be made as strong as possible by isometrics, and strength and local muscle endurance increased by light isotonic exercises, before recreational exercise is begun. Indications of excess therapeutic and recreational exercise include postexercise pain at two or more hours, undue fatigue, increased weakness, decreased ROM, and joint swelling. If these occur, the program should be adjusted.

**Treatment with Heat and Cold Modalities**
The use of superficial and deep heat and cold modalities have produced conflicting data. Recent formal literature reviews utilizing the Cochrane Databases and specific study selection criteria have been done for treating RA with theromotherapy (heat and cold) (214), balneotherapy (bathing/spa therapy) (215), and therapeutic ultrasound (US) (216). Only RCTs were selected for data analysis. Problems with poor methodological quality and inadequate statistical analysis and outcome measures often hindered pooling data and made conclusions about treatment efficacy difficult. The general conclusions from the thermotherapy reviews indicate no effect on objective measures of disease activity (inflammation, pain, x-ray measured joint destruction) of either ice versus control or heat versus control. Ninety-four percent of patients reported they preferred heat therapy to no therapy. There was no difference in patient preference for heat or cold, and no harmful side effects were noted (213). The balneotherapy review indicates ten studies report positive benefits, but the findings should be viewed with caution because of methodological flaws (215). The review on US suggested that ultrasound alone on the hand increases grip strength. It did not conclude that the combination of US with exercise, faradic current, or wax therapy was beneficial (216).

Recent data have helped clarify some dilemmas in this area. Therefore, we have some data (both human and animal) of the effects of these modalities on pain relief and disease activity. With the use of these data and clinical observation, recommendations can be made as to the appropriate use of heat and cold alone or in combination with other modalities for acute or chronic arthritis.

**Heat**
Superficial heat has been used for hundreds of years for pain relief in patients with arthritis. Patients report that warm baths, heated pools, hot packs, and warm mineral springs provide relief of joint pain and stiffness. Studies (217) indicate that superficial heat applied to patients with arthritis increases both skin and joint temperature in inflammatory arthritis. Painful stimuli, apprehension, alarm, or smoking lowers skin temperature and elevates knee joint temperature, as do active and passive exercise.

When joint temperature is increased from 30.5°C to 36°C, as it is in active RA, collagenase found in rheumatoid synovium is four times as active, resulting in lysis of cartilage (218). Increasing joint temperatures increases the metabolic rate and may increase inflammation and joint destruction (218). Mainardi et al. (219) found no increase and no decrease in joint destruction and inflammatory activity in the hand in RA with the use of superficial heat.

Heat affects the viscoelastic properties of collagen. As tension is applied, stretch is affected, and an increase of creep (i.e., the plastic stretch of ligamentous structures placed under tension) occurs. Heat may enhance the efficacy of stretching if applied to appropriately chosen joints.

Both superficial and deep heat can raise the threshold for pain, producing sedation and analgesia by acting on free nerve endings of both peripheral nerves and fibers of muscle spindles.

**Cold**
Cold modalities such as air or ice decrease skin, muscle, and joint temperature in arthritis patients (217). The application of cold to rheumatoid joints may, therefore, inhibit collagenase activity in the synovium. Some clinical studies have shown greater and more prolonged relief of pain with ice than deep or superficial heat in patients with RA (220). Other investigations found the increase of knee joint ROM to be the same with either ice or superficial heat applied daily for 5 days, with a 9-day interval between the two treatments (221).

Cold decreases muscle spasticity by direct action on the muscle spindle activity (222) and raises the pain threshold. Cold should not be used in patients with Raynaud’s phenomenon, cold hypersensitivity, cryoglobulinemia, or paroxysmal cold hemoglobinuria. The abrupt application of cold causes discomfort and produces a stressful response.

In treating the acutely inflamed or subacute joint, the goal is pain relief. One is careful not to use interventions that may increase the metabolic rate and secondarily increase inflammation. The use of cold seems most logical because it can decrease the pain threshold, can relax surrounding spastic muscles, and is associated with decreased joint temperature, collagenase, and cell count in the joint fluid (223).

---

**CHAPTER 40 | REHABILITATION OF THE PATIENT WITH RHEUMATIC DISEASES**

---

---
Later in the subacute period, when inflammatory pain is subsiding and stiffness is present, and the patient may have lost some ROM, either cold or superficial heat is appropriate and the patient should decide based on what is most convenient, economical, and effective.

**Orthotics**

Splints and orthotics are used to unweight joints, stabilize joints, decrease joint motion, support joints in a position of maximal function, and increase joint motion (i.e., dynamic splint). Splints may be prefabricated but are best when molded to fit the individual patient.

**Upper Extremities**

Orthotics for the upper extremities are mainly confined to the wrist and hand and include resting splints, functional wrist splints, thumb post splints, ring splints, and dynamic splints. Resting splints immobilize the hand and wrist and are used at night for patients with active RA, carpal tunnel syndrome, or extensor tendinitis (Fig. 40-3). The role of splints in preventing deformity in RA has not yet been scientifically proven. The clinical recommendation is to use both resting and functional splints in early RA. In JIA they probably help in delaying ulnar deviation and in reducing pain, synovitis, and edema. Functional wrist splints extend to the mid-palm crease, permit finger function, block wrist flexion, and are used for activities during periods of inflammation. They provide wrist and ligament support. A 2008 study using prefabricated working wrist splints revealed a 32% decrease in wrist pain (control group had a 17% increase in pain) (224). A functional thumb post splint may be used to relieve carpometacarpal (CMC) and IP pain associated with OA. The same type of splint with a longer wrist extension is useful for de Quervain extensor tendinitis of the thumb. A functional wrist cock-up splint can help relieve pain in carpal tunnel syndrome (Fig. 40-4).

Small ring splints (e.g., Bunnell orthoses, boutonniere orthoses) can reduce swan neck or boutonniere deformity. Cosmetic splints made of silver or gold and highlighted with semiprecious stones are available (Fig. 40-5).

For patients who have had MCP replacements or who have a radial nerve neuropraxia, a dynamic outrigger splint pulls the fingers into extension, from which patients must actively work to pull into flexion (Fig. 40-6). They provide gentle stretch through limited range while supporting the wrist and MCP joints. Splints that realign digits to help reduce ulnar deviation are also available (Fig. 40-7).

Elbow orthotics are rarely used. They may be useful in children with JIA and PM. Resting night splints may help to contain the advancement of an elbow flexion contracture. Braces with dial locks are used to increase extension during the day.

Compliance with splints has been assessed in a number of studies. Compliance is best when family members expect the patient to be compliant and when the patient uses splints to relieve pain. Cosmesis is a major factor for nonuse, as is fear of discrimination in the workplace.
Lower Extremities

The most useful orthotics in arthritis are those for the foot and ankle. Those for the knee have been less successful, and there are none generally used for the hip. The advances in lighter materials for orthotics assist in reducing energy consumption with their use.

Foot/ankle. Excess pronation at the subtalar joint, loss of the medial arch, and subtalar movement commonly seen in RA can cause pain, contribute to tarsal tunnel syndrome, and cause strain on the knee and hip. Control of pronation by bringing the calcaneus perpendicular to the floor often relieves pain and helps to balance the weight-bearing column (Figs. 40-8 and 40-9). The first step toward control is to fit the patient with a shoe with a good heel counter and a soft or rigid orthotic insert lined with Spenco (AliMed, Dedham, MA). The sole should not be too soft. This will minimize the flotation effect on heel-strike and stance during gait and decrease stress, hypermobility, or instability at the ankle or a higher joint level. If pronation is not controlled by a shoe, a hindfoot orthotic has been shown to improve gait and reduce pain (Fig. 40-10). A beveled heel that makes a 20-degree angle with the floor can decrease ankle motion and pain. For the very painful or arthritically involved RA ankle, a short-leg patellar tendon-bearing orthosis that shifts weight away from the ankle to the patellar tendon is useful. Plantar fasciitis may be relieved with a cup insert or an insert with a depression in the area of the tender fascia.

Appropriate wide toe-box shoes should be used to accommodate a wide forefoot, cocked toes, and hallux valgus seen in RA or JIA, as well as the hallux valgus deformity seen in OA. A soft insert is added, as are metatarsal reliefs, whether in the form of a cookie inside the shoe or an external bar on the sole of the shoe. We prefer the former because we believe it to be safer. A rocker bottom shoe can facilitate rollover in the presence of a painful ankle. Sometimes it is best coupled with a stirrup type brace to provide subtalar support (Fig. 40-11).
Knee. Braces for the knee may be used for pain relief, instability caused by ligamentous laxity, significant quadriceps weakness, or excess recurvatum. A useful brace for quadriceps weakness is a double upright Klenzak (Pel Supply, Cleveland, OH) set at 5-degree plantar flexion at the ankle to put the knee in extension during heel-strike and stance. The Klenzak can be used for a unilateral problem or for the weaker side when the problem is bilateral, as in PM. Success can also be achieved with a plastic molded ankle-foot arthrosis (AFO) cast in 5-degree plantar flexion with a small added 3/8-in. heel incorporated into the orthosis, provided the patient is not overweight. A Lenox Hill orthosis (3M, Long Island, NY) may be used to control mediolateral or rotational instability. These braces are rarely used in RA but are used for younger athletic individuals.

A knee, ankle, foot orthosis (KAFO) with ischial weight bearing and dial lock at the knee can be used to reduce knee pressure and may be adjusted to relieve medial or lateral compartmental stresses in OA or RA. This orthosis is difficult to fit with severe valgus deformities and in the obese patient. Compliance in the use of KAFOS is poor.

Smaller knee orthoses, such as hinged orthoses, the Swedish knee cage, or Lerhman orthoses (Pel Supply, Cleveland, OH), may be used to help control sagittal and frontal knee plane motion. A knee orthosis to help prevent dislocation of the patella is available and often effective. A shoe with a beveled heel at 20 degrees also decreases knee flexion and promotes a more stable extended knee. Orthoses with a dial lock turned 1 or 2 degrees daily can be used to reduce knee flexion contracture. Elasticized knee supports may help control swelling and often provide patients with a sense of control of the quadriceps.

Spinal Orthoses
Spinal orthoses are used primarily to relieve pain, limit motion, or support an unstable spine. A lumbar spinal orthosis or thoracic orthosis with mold and form insert will often relieve a painful back due to compression fracture or disc disease. This type of orthosis may reduce lordosis, reinforce abdominal muscles, and unload the spine. For thoracic compression fractures prone to gibbous deformity or an unstable lumbar or thoracic spine, a Jewett orthosis (Florida Brace, Winter Park, FL) or a molded polypropylene body jacket is required. A lumbosacral corset does not limit motion but provides some abdominal support and relieves painful lower lumbar musculature.

The cervical spine is involved in RA, OA, JIA, and spondyloarthropathies. Various collars provide different levels of support. A soft cervical collar only minimally limits motion but provides some pain relief. A Philadelphia collar (Pel Supply, Cleveland, OH) offers slightly more support and some limitation of extension. A two-poster, four-poster, or sternococciptal mandibular immobilizer (SOMI) brace substantially limits flexion and extension, particularly at C1, at C2, and also at C4 and C6. A halo is needed to completely control C1, C2 instability.

Assistive Devices and Adaptive Aids
Assistive devices and adaptive aids compensate for limited ROM and pain, and help promote independence and lesser impairment, and disability for arthritic patients. To help ensure patient acceptance, the appliance should be affordable, be easy to use, and improve patient function.

Ambulation and transfer skills are extremely important for persons with arthritis, and gait aids and devices may also be needed.

Gait Aids
If joint pain is a problem, secondary to loss of cartilage, effusion, or active synovitis, the painful joint needs to be unloaded.
Weight reduction is encouraged, because a 1-kg weight loss results in a 3- to 4-kg decrease in load across the hip joint. A straight cane when used properly unloads the limb by 25% (2), which is good for balance but is not as efficient in unloading the limb as compared to a forearm crutch. The elbow should be in 30 degrees of flexion when such a device is in use.

Custom handgrip pieces can be made by making a mold of the patient’s hand in a functional position of weight bearing, or commercially made hand pieces on canes can be used. Platform crutches distribute weight on the forearm, reducing the need for wrist extension and eliminating weight-bearing forces through the wrist and hand. Forearm attachments for walkers and wheelchairs are available.

For significant loss of strength or endurance, a small, lightweight wheelchair is recommended. There are also small motorized scooters, such as the sporty Amigo Chair (Amigo Mobility Intl. Inc., Bridgeport, MI); however, power wheelchairs will offer more seating options for those with spinal/pelvic involvement.

**Adaptive Devices for Transfer**

Chronic hip or knee pain, limited motion, and proximal muscle weakness make transfers from low-level chairs, toilets, and beds difficult. Upper extremities may be needed for pushoff, but when these are incapacitated by RA, such simple motion becomes impossible. Independence in making transfers can be restored by elevating the seat with a cushion or using a levered seat or placing 3- or 4-in. blocks under each leg of chairs, tables, and beds. Chairs with elevating motorized seats, elevated toilet seats, and clamp-on tub seats are also helpful.

Transfers in and out of the car are facilitated by the use of an extra-thick seat cushion, levered seat, or twist-about plastic disk and a mounted grab bar to increase leverage. In the car, use of the side, rearview, and wide-angle mirrors for patients with limited cervical ROM as a result of spondyloarthropathies becomes essential. A spinner bar for the steering wheel and a large-handled door opener and ignition piece are adaptations for the patient with significant hand problems. Patients with back pain benefit from a firm seat and back cushion, such as a PCP Champion Sacro cushion (OTC Professional Appliances, Ripley, OH). Those with neck pain need adjustable neck support or pillows.

**Self-care**

Dressing, undressing, and other daily self-care activities can be time- and energy-consuming tasks for those with RA, SLE, and PM. Adaptive and self-care aids such as long-handled reachers, shoe horns, elastic shoelaces, long-handled sponges, brushes and toothbrushes, Stirex scissors (North Coast Medical, San Jose, CA), button hooks, zipper hooks, toilet paper holders, and large-handled items are all helpful devices that also conserve energy. Clothing made with elastic and Velcro is easier to don than that with buttons and hooks. Wrinkle-resistant fabrics that do not require ironing and lightweight fabrics and woolens (e.g., mohair, alpaca) are useful. Large buttons and partial zipping before putting on the garment may facilitate dressing, as will stretch straps and waists, garments with large raglan sleeves, and those with smooth linings. Capes, ponchos, and down jackets are easy to put on, are warm, and lightweight.

**Devices in the Kitchen**

Useful kitchen devices include food processors, long-handled reachers, built-up handles on utensils, electric knives, and vegetable peelers, mounted wedge-shaped jar openers, and lightweight aluminum pans. Lining pans with aluminum foil saves scrubbing. Bringing together items involved in a work area (e.g., kitchen stove, work area, sink, refrigerator) is helpful. A microwave oven cuts down on food preparation time. A kitchen cart loaded with often-used utensils cuts down on walking. These are all excellent examples of tools for work simplification and energy conservation.

**Environmental Design**

Slopes, stairs with deep steps, high curbs, and buses or cars may be difficult to negotiate for someone with disease of the hips and knees. Appropriate placement of steps, lowered curbs, suitably graded inclines, and ramps are helpful. Buses that kneel to accept passengers are available in many communities. Indoors, thick carpets increase friction and are difficult to walk on or negotiate in a wheelchair. In the bathroom, guard rails are best for safety. The bathtub should have nonskid strips or an entire nonskid surface. A wheelchair shower with a wall seat is excellent for wheelchair use. Door openings should be wide enough to accept a wheelchair. Chest-high storage cabinets and waist-high work surfaces are best; special door handles are available. For those patients in wheelchairs, proper positioning of doorknobs, light switches, and kitchen equipment is necessary. Large-handled pencils and eating utensils are helpful. Devices to help with spray cans are available. The AF provides a catalog of available assistive devices (www.arthritis.org).

**Education**

One of the most important aspects of care for patients with inflammatory disease, and perhaps all chronic illness, is a good rapport between physician and patient. This begins with clear communication between patient and health care providers so that patients can be assisted in habilitation or adjustment to the chronic disease process. In addition, educational groups consisting of arthritis patients who gather to hear experts in the field and talk about aspects of medical and rehabilitative management are very informational and supportive. They also provide a social context for coping with medical and psychosocial problems. Local hospitals, the YMCA, and the AF and other disease foundations and local interest groups offer such group activities.

It has been thought that patients with RA, and chronic disease in general, have had significant disability based on inactivity and lack of motivation. Approaches to solving this problem and promoting better patient function have included techniques to empower patients and to promote self-efficacy, thereby overcoming learned helplessness (225).
Educational intervention, using disease self-management programs, has been employed in the treatment of RA, SLE, AS, and fibromyalgia. These have reduced pain, improved function, and delayed the onset of disability. They are also under-used (226–228). Studies show that information is helpful in empowering patients and it often helps to relieve anxiety. However, educational interventions that only promote increased knowledge have little impact on disease activity and behavior. Educational models that establish approaches to problem solving, support groups, and healthy lifestyle/fitness behaviors are usually more effective in promoting improved function.

Patient education should include a discussion about the natural history of the disease and the likely impact that it will have on lifestyle, job, and leisure activities. Many systemic RDs are chronic and have periods of remission and exacerbation that affect function. Planning activities so that they are done at the optimal time of day, or are broken up into smaller intervals for sustained activity, may permit completion of important activities without undue fatigue.

A full discussion of the impact that medication may have will help alert patients to medication-related problems. For example, corticosteroids may cause muscle atrophy, upset sleep patterns, and alter appearance. These changes may have profound effects on mood, energy, and strength, all of which may influence function.

**Joint Protection**

It is well known since the observations of Cordery that arthritic patients change their lifestyle and habits to protect their joints and use adaptive strategies and devices that make work easier rather than persist with activity that causes pain (229). However, patients need formal education in this area to learn techniques to reduce force across their joints. Principles of this program include the following: promoting good sitting posture; using adequate back contact, arm rests, and proper height such that the feet touch the floor; changing position frequently to avoid gel phenomenon. Additional examples of ways to protect/preserve joints include: using the largest possible joints to support activity, unweight painful lower-extremity joints, avoid overuse by interrupting sustained activity with rest periods, using adaptive equipment and strategies for efficient use of joints, using splints when limbs need to be supported in functional positions.

**Energy Conservation**

RD is most often accompanied by fatigue, so conserving energy to maximize function is an important part of the arthritic patient's lifestyle. Some of the mechanisms for conserving energy include the proper orthotics and assistive devices to effect energy-efficient ambulation and hand function, appropriate adaptive aids and clothing, proper environmental design, rest periods throughout the day, maintenance of ROM and strength, and maintenance of proper posture.

Good body posture, whether sitting or standing, balances the weight of the head and limbs on the body framework so that gravity helps to maintain joint position with minimal muscle activity. Significant changes in posture cause muscles to exert more energy to pull against the force of gravity. For example, standing takes 25% more energy than sitting to perform activities. Ideal posture cannot be maintained unless care is taken to preserve ROM and strength of muscles around the joints.

**Psychosocial Interventions**

RD has a major impact on the patient's mobility, ADLs, general lifestyle, self-image, family life, sexuality, and work. Stress and depression play major roles in influencing RD patients' reporting of their symptoms and illness.

Two forms of stress play a role in RD: major life events, such as death of a spouse, and persistent negative events (230), such as conflicts with people. The latter stressors can influence immune function and increase global ratings of disease activity in RA (231), increase psychological distress in SLE (232), and increase symptoms in JIA (233).

These studies support the association between stress and symptom flares and indicate that interventions to help coping may be useful. Major patient reactions ranging from denial and repression to depression occur, as do other components of chronic illness, such as anger, bargaining, and acceptance.

Depression is common in RD. It has been reported that 37% of patients with RA have depression (234). Predictors of depression include scores in HAQ, pain intensity (235), and, in women, loss of valued activities (230). A history of depression in RA makes patients vulnerable to higher levels of pain, fatigue, and disability, even if they are not currently depressed, but just in a dysphoric mood (236). The belief that there is a particular personality type prone to developing RA has been dropped. However, premorbid personality is an important determinant of the patient's reaction to illness. Patients with arthritis, in addition to coping with pain, may have to deal with losses in function and physical attractiveness because of the disease and side effects of medication, and they have to deal with the reactions of friends, spouse, and family. Some families adjust well, maintaining good communication, support, and flexibility in family routines. Other families are not able to reorganize and adjust to the needs of a person with arthritis. Psychiatric support for both patient and family is often needed. Group therapy, including discussion about body image, job status, family relationships, and coping mechanisms, is effective and has been used for a variety of arthritic groups.

The unpredictability of systemic disease complicates the coping process. Patients with RD often experience learned helplessness, believing no solutions exist to reduce stressful events, which leads to anxiety, depression, increased pain, decreased activities, and decreased ability to adapt to distress and disability (230). High levels of helplessness are associated with increased pain, depression, and functional impairment in RA and poorer physical health in SLE (237). Self-efficacy in patients with RD may vary for different actions to achieve health-related goals. High self-efficacy for pain is correlated with low frequency of pain behavior in RA and is related to physical and mental...
health in SLE (238,239). Coping strategies involve motor and cognitive actions to control the impact of stresses. Passive coping strategies are associated with high pain levels, disability, and poor illness adaptation (230). Psychosocial interventions help improve pain, affect, and function (239–241). They may also increase self-efficacy and health status. RA patients who use stress management programs increase self-efficacy and improve ratings of depression, pain, and walking speed (242). Treatment offered through telephone-based interventions has been associated with psychological improvement in SLE. Audiotapes and videos used at home yield improvements in physical function, pain, joint count, and global scores in RA (230). In summary, psychological and behavioral approaches are very useful in RD (243,244).

**Sexual Adjustment**

Many normal, well-educated, financially successful, and maritaly stable U.S. adults admit to sexual dysfunction or difficulties. However, arthritis may impose additional limitations or alterations that influence the sexual life as they do other ADL (245).

Particular sexual problems in the arthritic patient arise as a result of mechanical problems associated with decreased ROM, pain, and stiffness; depression, with decreased self-image and interest; drug therapy resulting in decreased libido; psychosocial problems in the family unit related to the patient's arthritis; and fatigue. Studies indicate that impairments occur in various indices of sexual function in a number of autoimmune diseases. In PSS, vaginal dryness, ulcers, and dyspareunia are significantly more common after disease onset in comparison with SLE and RA controls (246). Dyspareunia is also common in Sjögren's (247). There is a 50% decrease in the number and intensity of orgasms. Sexual satisfaction index is impaired in all three groups. Skin tightness, reflex, heartburn, and muscle weakness adversely affect sexual relations, more in PSS than in SLE and RA (246). Impotence occurs in males as a result of neurovascular problems in PSS and related syndromes (248).

**Joint stiffness and mechanical problems.** Arthritic involvement of the hips, knees, lumbar spine, hands, and shoulders commonly causes joint stiffness and painful mechanical problems that interfere with sexual performance. Hip involvement commonly causes mechanical problems. Analgesics and warm baths before intercourse may help. In women with more advanced disease, a posterior approach by the man for intercourse may be successful. With severe limitation of motion, either unilateral or bilateral hip replacements may be a requisite to achieve intercourse. After hip surgery, intercourse should not be resumed at all for 6 weeks, and hip flexion of more than 90 degrees should be avoided. Attention to sexual adjustments after joint replacement is important (249). Neither pain, nor stiffness, nor limitation of motion of the knees should mechanically limit intercourse, but a change in position may be required for more comfort. For those patients with back pain (e.g., spondyloarthropathies, disc disease), a lateral position for both the man and the woman is preferable. Patients with significant muscle weakness, as in DM-PM, may need to assume a more passive sexual role.

Significant problems in the joints of the hands and arms, as in RA, are more restrictive for the man than for the woman; a side-lying position may alleviate this. In both men and women, arthritis of these joints may interfere with the early stages of lovemaking involving caressing and manual stimulation.

**Self-image.** Often arthritic patients experience a decrease in self-image, a feeling of helplessness, and ultimately depression, which in turn is associated with decreased libido. Chronic pain may reduce a woman's efforts to make herself more attractive. A man may abstain from sexual relations rather than cause his arthritic spouse pain, and the spontaneity is reduced. In comparison to RA patients with a low degree of morning stiffness, RA patients with a high degree of morning stiffness worried more about self-image and reported significantly more problems in sexuality (250). Appropriate counseling may help alleviate these problems.

**Medications.** Certain medications are associated with decreased libido. High-dose steroids used in SLE and DM-PM may affect the physical appearance and contribute to decreased self-image or acceptance by one's spouse. Immunosuppressive medications may interfere with conception. Reports have been published regarding gynecomastia and sexual impotence associated with MTX (251). Many of the medications used to manage the RDs may cause significant fatigue. Education about energy conservation and the practice of aerobic exercise may be generally useful in managing this symptom.

**Psychosocial problems.** Psychosocial problems in the family unit may lead to decreased sexual relationships between husband and wife. Such problems include inability to work, inadequate finances, limited acceptance by friends, and limited participation in social or recreational events. Misunderstandings and anxieties of the children about the chronic disease of the parents also contribute to tension and adversely influence sexual relationships.

Patients with JIA had similar attitudes toward sexual activity, conception, wish for children, age at first child, and duration of lactation. However, fecundity is significantly decreased and an increased miscarriage rate has been reported. Males had greater problems than females in establishing relationships (252). Successful treatment of musculoskeletal disease does not necessarily serve as a sexual restorative; rather, understanding, counseling, alterations in sexual positions, or appropriate joint replacements are necessary.

**Vocational Aspects**

A vocational assessment should be part of the workup for RD patients and should cover educational level, work history and achievements, physical functional level, and social and psychological adjustment. The capacity of individuals with RD to work is decreased when compared with those without RD (253,254), and those with RA are particularly impacted. Three years after diagnosis with SLE, 40% patients are no longer working (255). For those patients who work as homemaker, the RD has a negative impact on family activities and function (248).
Determinants of work disability exist at societal and individual levels (256). Societal factors include economic conditions, attitudinal and architectural barriers, types of jobs available, employers’ practices, and types of disability pension plans available (253,254). Individual determinates are disease, social factors, and ability to maintain locus of control and anatomy on the job (252,256). Functional status is the most important determinant of household work performance (257,258).

A recent study indicates that women have higher degree of work disability in RA (259). Remaining in the work force or being a homemaker is an important goal for the person with arthritis, and rehabilitative care and the support of the family help ensure success. The first step in vocational counseling is to see if adjustments can be made in the job or home setting so that the patient can continue working. If changes cannot be made, then the patient may have to receive formal vocational rehabilitation, or training for another vocation.

**Specific Diseases**

**Rheumatoid Arthritis**

RA is a systemic RD characterized by an acute inflammatory response. It has different patterns of onset and can be divided into different types (Table 40-15) (260).

Depending on the severity of the progression of the disease, the pathologic process of RA has impact on joints, their surrounding structures, and other major organ systems. This process can produce a number of physical (anatomic and physiologic) and psychological (mental and emotional) impairments. These impairments cause functional problems that can result in disability. Table 40-16 lists the potential variables that can influence the extent of impairment, functional decline, and disability.

**Course and Disease Outcome**

Fewer than 25% of RA patients have a true remission in 5 years (260), as measured by ACR criteria. Most have progressive disease of varying degrees. The rate of progression to joint destruction and disability is proportional to the intensity of the inflammation, proliferative reaction in the joints, and persistence of disease over time. Low-grade attacks of synovitis, separated in time, are less likely to cause joint deformity than continuous active synovitis. Radiographic progression is highest in the first few years of the disease but can progress over decades. Once the synovium begins to destroy and invade cartilage, it is at risk for irreversible damage even if disease activity decreases (260,261).

In the 1980s, studies indicated that RA patients had a progressive decline in functional capacity, ability to work, and increase in mortality rate of 5 to 20 years (260). Work disability after 5 years is 60% to 70% in RA patients under 65 years of age (262). Other studies have developed models using baseline indices to predict outcome in RA (263,264). This information has been useful to clinicians in planning treatment (Table 40-17). One study predicted increased disability by both disease factors and nondisease variables. Several comorbid conditions may contribute to disease mortality (Table 40-18). Disease-related factors explained 33% of total disability, and nondisease factors (depression, education, psychological status) explained 26%. Unexplained were 41% (264). Similarly, disease mortality has been predicted and 5-year survival has been reported as 85% to 95% in patients who had favorable baseline variables versus those who had unfavorable baseline variables and who had a mortality rate of 45% to 55% (265).

### Table 40.15 Onset and Types of RA

<table>
<thead>
<tr>
<th>Onset</th>
<th>Types</th>
</tr>
</thead>
<tbody>
<tr>
<td>Insidious over weeks to months, 65%</td>
<td>I  A self-limited polyarthritis</td>
</tr>
<tr>
<td>Acute over a few days, 8%—15%</td>
<td>Meets ACR criteria for RA on presentation</td>
</tr>
<tr>
<td>Intermediate over days to weeks with more systemic complaints, 15%—20%</td>
<td>No patients meet criteria 2—5 y later</td>
</tr>
<tr>
<td></td>
<td>II Persistent disease</td>
</tr>
<tr>
<td></td>
<td>Management by conservative treatment</td>
</tr>
<tr>
<td></td>
<td>Few significant long-term consequences</td>
</tr>
<tr>
<td></td>
<td>III Progressive disease with radiologic damage</td>
</tr>
<tr>
<td></td>
<td>Functional decline</td>
</tr>
<tr>
<td></td>
<td>Premature morbidity</td>
</tr>
</tbody>
</table>

ACR, American College of Rheumatology; RA, rheumatoid arthritis.

### Table 40.16 Variables Influencing Impairment, Function, and Disability in RA

<table>
<thead>
<tr>
<th>Impairment</th>
<th>Function</th>
<th>Disability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disease severity</td>
<td>Number of impairments</td>
<td>Number and severity of impairments</td>
</tr>
<tr>
<td>Disease rate of progression</td>
<td>Severity of impairments</td>
<td>Number and degree of functional problems</td>
</tr>
<tr>
<td>Parameters of disease activity</td>
<td>Level of coping strategy</td>
<td>Economic and impairment issues</td>
</tr>
<tr>
<td>Persistence of disease (disease duration)</td>
<td>Level of self-efficacy</td>
<td>Availability of disability income</td>
</tr>
<tr>
<td>Access to early diagnosis and treatment</td>
<td>Tolerance to pain</td>
<td></td>
</tr>
<tr>
<td>Compliance with treatment</td>
<td>Social and economic support system</td>
<td></td>
</tr>
<tr>
<td>Significant radiographic erosions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medication side effects</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extra-articular manifestations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comorbid illness</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

RA, rheumatoid arthritis.
CHAPTER 40 | REHABILITATION OF THE PATIENT WITH RHEUMATIC DISEASES

Treatment of RA

Early aggressive pharmacologic treatment by rheumatologists has been shown to be associated with improved function in early RA (4). Early functional assessment and a multidisciplinary team approach to patient management and rehabilitation have shown benefit in improving clinical outcomes (266).

To best treat patients with RA, the physiatrist should be aware of a number of entities (Table 40-19). All these are important because they influence the type, intensity, and frequency of prescribed treatments. They help in setting up appropriate precautions and in determining if other medical or rehabilitation consultation needs to be made before treatment initiation.

RA is associated with a high degree of disability, the causes of which have been frequently attributed to impairments and the biology of the disease. Recent literature suggests that the radiologic appearance in those with RA correlates highly with measures of functional ability (HAQ and AIMS) (267,277).

Critical to the management of RA should be treatment aimed at preventing disability and preserving or improving function. Usually these include muscle strengthening while protecting joint integrity (isometric exercise may be preferred), aerobic conditioning using low-impact methods, and encouragement of leisure activity that promotes socialization and well-being (268,269).

Impairments in RA and Treatments

See Table 40-20 for a listing of rehabilitative treatments and functional deficits and disabilities in RA.

Pain. Among patients with RA, 70% seek help for pain (270). Pain in RA is caused by active synovitis, stretch of the joint capsule by effusion, mechanical joint destruction, and moving and loading joints. Pain is complex and has physiologic and subjective components. It contributes significantly to functional disability in early RA (271). Treatment includes education, exercise, stress management, medication, physical modalities, orthotics, assistive devices, and joint protection.

Fatigue. Fatigue is one of the most common impairments in RA. Characteristics of fatigue are varied and may include inability to sustain daily routines or leisure activities, sleepiness, inability to wake up refreshed, depression, and muscle weakness. Treatment is targeted at the most likely organ system contributing to the symptom. For example, sleepiness or lack of refreshing sleep may best be treated with medication for...
sleep. Muscle weakness, on the other hand, is best treated with a strengthening program, and endurance problems are likely to respond to aerobic conditioning. Interrupting activity with periodic rest (20-minute naps) during the day has been shown to be effective in increasing physical activity (PA) levels (272). Also to be included are energy conservation techniques and depression management.

**Decreased aerobic capacity.** Deconditioning contributes to decreased aerobic capacity. Patients tend to limit activities because of joint pain fatigue and decreased motivation. Cardiovascular problems due to RA or other diseases may contribute. A 2006 internet-based PA program for RA with individualized guidance (IT) versus general information (GT) showed the IT program was more effective with respect to the number of patients who report meeting PA recommendations than the GT program. This program is inexpensive and readily accessible and should be useful for RA patients who have limited activity levels (273).

Aerobic capacity in RA is commonly evaluated by submaximal aerobic testing. A recent 2008 study identified that a HAP is useful in estimating fitness levels when aerobic standard testing is not feasible (274). Aerobic training has proven to be of benefit for class I through class III RA. It is generally of low intensity in class III or of moderate intensity in classes I and II (Table 40-21). Appropriate forms include cycle ergometer, land (low-impact) dance or aerobic routines, or water aerobics. The latter is appropriate for class III deranged joints or joint replacements. Aerobic exercise can improve function and increase aerobic capacity without increasing inflammation, if chosen appropriately.

**Sleep disorder.** Several different types of sleep disorders exist in RA: (a) fragmented with nocturnal clonus, (b) frequent prolonged wakefulness after sleep onset, and (c) sleep apnea. Insomnia or early AM awaking is not common. Treatment consists of modalities and appropriate pain medication before bed, and provision of a good mattress or pillow. Medications such as clonazepam, L-dopa, short-acting benzo-diazepines, and zolpidem can be used to improve sleep quality and quantity.

**Depression.** Depression can often accompany RA but is no more common in RA than in other chronic disorders (275). Depression and its subtypes in RA are correlated with health status measurements. Stress and anxiety are common. Appropriate treatment for depression includes treating the underlying cause and medication. Depression, stress, and anxiety may benefit from stress management, relaxation techniques, and coping strategies.

**Decreased strength.** Strength deficits occur very early in RA and are due to atrophy around painful joints, reflex inhibition by pain and joint effusion, inactivity, and myositis, which may be associated with RA. Muscles around key joints (particularly the knee, hip, shoulder) should be strengthened. The literature provides evidence that isometric, isotonic (low- and moderate-resistance exercise) (163,199), and isokinetic exercises all strengthen muscles in RA. In addition, exercise as

---

**TABLE 40.20  Rehabilitative Treatment of Impairment, Functional Deficits, and Disabilities in RA**

<table>
<thead>
<tr>
<th>Deficit</th>
<th>Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Impairment</td>
<td>Pain</td>
</tr>
<tr>
<td>Fatigue</td>
<td>Exercise (isometric/isotonic/isokinetic)</td>
</tr>
<tr>
<td>Joint swelling (acute inflammation) (subacute/chronic)</td>
<td>Cold pack, ROM, heat, exercise (strengthening/aerobic)</td>
</tr>
<tr>
<td>Weakness</td>
<td>Aerobic exercise, energy conservation, short rests</td>
</tr>
<tr>
<td>Deconditioning</td>
<td>Remove before exercise program (if clinically significant)</td>
</tr>
<tr>
<td>Joint effusion</td>
<td>ROM/strengthening (if biomechanically stable, not ankylosed and 0 to mild effusion), modalities</td>
</tr>
<tr>
<td>Limited joint motion</td>
<td>Medication, counseling, relaxation techniques, coping, strategies</td>
</tr>
<tr>
<td>Depression/anxiety/stress</td>
<td>Medication, modalities before sleep, naps, stress reduction, mattress, and pillow changes</td>
</tr>
<tr>
<td>Sleep disturbance</td>
<td>Education, counseling, address course</td>
</tr>
<tr>
<td>Sexual dysfunction</td>
<td>Mobility</td>
</tr>
<tr>
<td>ADL</td>
<td>Gait aides, scooter</td>
</tr>
<tr>
<td>Gait</td>
<td>Assistive devices, joint protection, education environmental strategies</td>
</tr>
<tr>
<td>Homemaking</td>
<td>Assess cause of deviation, gait aids, appropriate shoes, and orthotics</td>
</tr>
<tr>
<td>Disability</td>
<td>Environmental adjustments, adaptive devices, energy conservation joint protection</td>
</tr>
<tr>
<td>Can’t work</td>
<td>Disability Can’t do housework</td>
</tr>
<tr>
<td>Can’t care for self</td>
<td>Disability</td>
</tr>
</tbody>
</table>

RA, rheumatoid arthritis; ADL, activities of daily living; ROM, range of motion.
part of aerobic and aquatic programs also yields strength gains. Isotonic resistive exercise yields more gains than isometric exercise. The types of exercise, the intensity, and the frequency depend on ACR functional class, presence or absence of joint effusion, biomechanical joint integrity, inflammation stage, presence of moderate to severe OP, and presence of joint replacements (see Table 40-21). In general, keep the exercise simple, meet needed patient goals, and follow exercise guidelines for compliance (see Table 40-9). RA is a chronic disease that needs continued physical exercise with sufficient intensity to prevent decreases in muscle strength and function loss (276).

Decreased ROM. In early RA it is important to prevent contracture. In later stages contractures may be present. Patients should be on a basic ROM program daily for all affected joints. In patients with limited ROM, due to contracture, ROM exercise should be in the available arc of motion.

Functional problems. Such problems as decreased mobility, compromised ADL, and impaired gait are common in RA. Determining the impairments that produce these problems is necessary. The impairments should be managed appropriately. This will help reduce the functional limitation. Education on use of gait aids and assistive devices, orthotics, and environmental strategies is needed (see Table 40-20).

Disability. Although past studies indicate a significant number of persons with RA experience work and household disability, newer pharmacologic agents and early rheumatologic and rehabilitative care have resulted in improved function and less disability (277). Management includes investigating the impairment and functional problem contributors to disability. Recently, subclinical disability, a need to modify task performance or frequency without reported difficulty with the task, has been identified as part of the disability continuum. Seventy-five percent of a group of RA patients identified difficulties with valued life activities (VLAs). They were significantly more likely to experience functional limitations over a 2 year period. Subclinical disability may be a valuable marker of those in a disability transition phase where intervention may help maintenance of function (278). Intervention by vocational rehabilitation with patient interview and workplace assessment is

| TABLE 40.21 Exercise Type Recommendations for Specific Joint Conditions and Functional Level |
|----------------------------------|------------------------------------------------|-------------------------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| Joint Activity                  | ROM Activity | Stretch Activity | Isometric Activity | Isokinetic Activity | Isokinetic Activity | Aerobic Activity |
|                                 | Full | Partial | Gentle | Full | Low R | Moderate R | Low R | Moderate R | Low I | Moderate I |
| Acute                           | +   | −     | −      | −     | −     | −         | −     | −         | −     | −         |
| Subacute                        | +   | −     | +      | −     | +     | −         | −     | +         | +     | −         |
| Chronic (inactive)              | −   | −     | −      | +     | +     | −         | −     | +         | +     | +         |
| Functional level                |     |       |        |       |       |           |       |           |       |           |
| I                               | +   | −     | +      | +     | +     | +         | +     | +         | +     | +         |
| II                              | +   | −     | +      | +     | +     | +         | +     | +         | +     | +         |
| III                             | +   | +     | +      | −     | −     | −         | −     | −         | −     | −         |
| IV                              | +   | +     | +      | −     | −     | −         | −     | −         | −     | −         |
| Biomechanical level joint       |     |       |        |       |       |           |       |           |       |           |
| Normal                          | +   | −     | +      | +     | +     | +         | +     | +         | +     | +         |
| Mild                            | +   | −     | +      | +     | +     | +         | +     | +         | +     | +         |
| Moderate                        | +   | +     | +      | −     | −     | −         | −     | −         | −     | −         |
| Severe                          | −   | +     | −      | −     | −     | −         | −     | −         | −     | −         |
| Joint effusion                  |     |       |        |       |       |           |       |           |       |           |
| 0                               | +   | −     | +      | +     | +     | +         | +     | +         | +     | +         |
| Mild                            | +   | −     | +      | +     | +     | +         | +     | +         | +     | +         |
| Moderate                        | −   | +     | −      | −     | −     | −         | −     | −         | −     | −         |
| Severe                          | −   | +     | −      | −     | −     | −         | −     | −         | −     | −         |
| Joint replacement               |     |       |        |       |       |           |       |           |       |           |
| Osteoporosis                    |     |       |        |       |       |           |       |           |       |           |
| Mild                            | +   | −     | +      | +     | +     | +         | +     | +         | +     | +         |
| Moderate                        | +   | +     | −      | −     | −     | −         | −     | −         | −     | −         |
| Severe                          | −   | +     | −      | −     | −     | −         | −     | −         | −     | −         |

AROM, active range of motion; R, resistance; I, intensity.
essential. Ergonomic assessment may help people with arthritis maintain employment. An ergonomic assessment tool was developed and was shown to be feasible and comprehensive in identifying job issues that need accommodation strategies (279). All factors influencing disability need to be addressed, since workplace accommodations alone may not be associated with improvement in employment rate (280).

Specific joints. RA may cause a variety of specific joint impairments that primarily affect the synovial lining of diarthrodial joints. It can affect almost any and all the peripheral joints, with relative sparing of the axial spine, except for the upper cervical (i.e., atlantoaxial) joints. However, degenerative change frequently accompanies it and involves C4-5 and C5-6 and the hands, knee, hips, and metatarsal phalangeal (MTP) joints. The end result of the process is joint pain, swelling, and malfunction. The muscles that surround the swollen, inflamed joints, or biomechanically compromised joints, are painful and are often atrophied or myositic. Ankylosis may occur, but subluxation is more common. Malalignment and pain result in impairments that primarily affect the synovial lining of diarthrodial joints. It can affect almost any and all the peripheral joints, with relative sparing of the axial spine, except for the upper cervical (i.e., atlantoaxial) joints. However, degenerative change frequently accompanies it and involves C4-5 and C5-6 and the hands, knee, hips, and metatarsal phalangeal (MTP) joints. The end result of the process is joint pain, swelling, and malfunction. The muscles that surround the swollen, inflamed joints, or biomechanically compromised joints, are painful and are often atrophied or myositic. Ankylosis may occur, but subluxation is more common. Malalignment and pain result in increased problems with ADLs, mobility, and energy expenditure. The muscles that surround the swollen, inflamed joints, or biomechanically compromised joints, are painful and are often atrophied or myositic. Ankylosis may occur, but subluxation is more common. Malalignment and pain result in increased problems with ADLs, mobility, and energy expenditure. Pain and deformity may cause problems with self-image and sexual activity (260).

Joint-specific problems. Shoulder RA affects the glenohumeral joint, distal third of the clavicle and surrounding bursae, capsule, and ligaments. Arthritis is associated with pain in the shoulder girdle, which is referred to the neck, back, and upper arm. Decreased motion of the joint, soft-tissue contracture, and muscle atrophy follow. Because the capsule lies deep under the rotator cuff, effusion is difficult to detect on physical examination. Static and dynamic rotator cuff exercises in mild RA decrease arm pain, help swollen joints, and improve SIP scores (281).

Pain causes decreased ROM. Limitation of IR is seen early. Proximal subluxation of the humeral head occurs late in the disease. Weakness of the rotator cuff may cause superior subluxation in about 33% of RA patients; about 21% of patients develop rotator cuff tears; and an additional 24% have fraying of the tendons (260). The insertion of the rotator cuff tendon into the greater tuberosity makes it vulnerable to erosion by synovitis. Adhesive capsulitis, with anterolateral shoulder swelling, subacromial and subdeltoid bursitis, and bicipital tendinitis are associated problems. Anterolateral soft-tissue shoulder swelling can indicate subacromial bursitis.

Adhesive capsulitis with motion loss can occur quickly. Pain is worse at night, when sleep movements stretch the capsule. Full motion is not needed for most daily activities, so pain may be less during the day. In addition to pain control with heat and cold, local steroid injection into the specific affected area is often useful. A ROM program to increase and prevent loss of mobility is crucial. For functional activities, the shoulder must have 30 to 45 degrees of flexion and 10 degrees of IR. Care must be taken to assess the degree of radiographic involvement and joint stability when prescribing mobilization so as not to injure a compromised joint. When pain and inflammation subside, Codman exercises and the use of a cane or wand can increase flexion and IR and ER. Wall walking is good for chronic capsulitis. In the presence of adhesive capsulitis, a technique of abduction-ER-flexion traction for 1 hour a day, in conjunction with transcutaneous electrical nerve stimulation (TENS), has been successful in decreasing pain and increasing ROM (147). Isometric strengthening should first focus on the deltoid with the shoulder adducted, then wrist-restricted isometrics in IR and ER; finally, triceps and biceps isometrics are added. Instruction in joint protection is essential to avoid overusing the shoulder. Arthroplasty should be considered before end-stage erosion and soft-tissue contraction occur.

Elbow involvement is common in RA (20% to 65%), depending on disease severity, with loss of full extension being an early problem. Loss of lateral stability can cause significant impairment. Preservation of flexion is needed for ADL. In severe disease, lateral stability may be lost, which may cause significant pain and disability in ADL function. Olecranon bursitis and accumulation of RA nodules that may breakdown easily are annoying to the patient (260).

Bursitis may be caused by Staphylococcus infection, and care must be taken not to inject the bursae of the elbow with steroid before a culture is performed. Wearing a padded Heelbo (Heelbo, Niles, IL) is useful to relieve pressure.

Lateral and medial epicondylitis is common. Acute epicondylitis is managed with cold modalities. Steroid injection may be necessary. Stretching exercise should not be forceful because articular damage is often present in the arthritic elbow.

The hand and wrist function as a unit. With weakness of the extensor carpi ulnaris, the carpal bones rotate (i.e., the proximal row in an ulnar direction and the distal ones radially), resulting in ulnar deviation of the MCPs. A power grasp and weakened intrinsics accentuate these problems.

Synovial proliferation increases pressures in the wrist joint, so that ligaments, tendons, and cartilage may begin to be destroyed. When the ulnar collateral ligament is stretched or ruptures, the ulnar head springs up dorsally and floats. In advanced disease, the carpus becomes significantly compacted. Carpal tunnel syndrome may occur bilaterally. Progressive wrist disease results in decreased motion or ankylosis if the disease is severe and of long duration.

In the hand, muscle weakness and contraction occur and grip strength decreases. One study of male and female RA patients (disease duration mean of 7.5 years) followed for 5 years, showed at the 5-year mark that grip strengths, Keitel function test, and HAQ scores were significantly worse in women. One fourth of patients in both groups needed more ADL assistance (282).

Swan neck deformity flexion of the distal interphalangeal joint (DIP) and hyperextension of the proximal interphalangeal joint (PIP) occur, as does Boutonniere deformity when the extensor hood of the PIP is stretched, causing flexion of the PIP and hyperextension of the IP. With incomplete profundus contraction, limitation of full flexion occurs at the DIP joints. Similarly, tight intrinsics prevent full flexion of the PIP joints with the MCPs in extension.
Three types of deformity occur at the thumb:

_Type I:_ A boutonniere-type deformity at the IP joint (i.e., Nalebuff)
_Type II:_ Volar subluxation at the CMC joint during contraction of the pollicis adductor
_Type III:_ In severe disease, exaggerated adduction of the first carpometacarpal joint and flexion at the MCP, and hyperextension at the DIP joint

Flexor tenosynovitis and de Quervain thumb extensor synovitis are common (260).

Rehabilitative hand care involves stretching of tight intrinsic and exercise. A controlled study of hand exercises versus no exercise for 48 months in female RA patients showed a significant increase in grip strength and pincer grip strength in the exercise group. The control group showed a significant decline in these parameters (283). A 3-week randomized control trial of combined ice massage and wax treatment and thermal and faradic baths with exercise revealed significantly improved Richie articular index, hand pain, ADL score and grip strength, and ROM in the treatment group. All parameters slightly declined in the control group (284). Use of functional wrist splints and finger ring-type splints to help reduce hyperextension or fixed flexion deformities, joint protection techniques, and postoperative care are important. These splints may help decrease synovitis, relieve pain and edema, and, when worn, reduce deformity and possibly retard its progression. A study of 273 RA patients indicated that in those with disease of less than 5 years, 269 needed help from a person in ADL activities. The highest amount of disability resulted from impairment of the wrists. This indicates that early attention should be paid to treatment and stabilization of the wrist (285).

_Lower-extremity joints._ The hip, knee, and foot-ankle are weight-bearing joints and cause greater pain and dysfunction than upper-extremity joints.

About 50% of patients with RA have radiographic hip involvement (260). Since this is a deep joint, early inflammation is not apparent. Palpation of effusion is difficult. Synovitis of the hip can cause pain radiating to the groin, whereas trochanteric bursitis causes pain radiating over the lateral thigh but can be in the buttock, anterior thigh, low back, and knee. Collapse of the femoral head and remodeling of the acetabulum, which is pushed medially (i.e., protrusio), occur in 5% of RA patients. Reduction in IR is an early finding with hip involvement. Synovial cysts can develop around the hip joint and communicate with the trochanteric bursae. Hip effusion can inhibit contraction of the gluteus medius muscle.

ROM exercises are important, first to maintain at least the crucial 30 degrees of hip flexion. A tight tensor fascia lata should be stretched. Stretching in abduction helps to relieve pain. Stretching of the internal and external rotators, extensors, and abductors should be followed by isometric strengthening exercise for the hip abductors and extensors. Ultrasound is best avoided in RA of the hip because it is somewhat difficult to assess the state of the inflammatory process in this deep joint. Because ultrasound increases joint temperature, it may aggravate an existing acute or subacute process.

The knees are commonly involved in RA, and synovial inflammation and proliferation and effusion are easily seen. Quadriceps atrophy occurs within weeks of the onset of the disease and leads to increased forces through the patella to the femoral surface. Loss of full knee extension also occurs early, and fixed contractions may ensue. Patients with lower limb involvement have been found to have quadriceps sensorimotor dysfunction that was associated with lower-limb disability. A clinically applicable rehabilitation program increased quadriceps sensorimotor function and decreased lower-limb disability without exacerbating pain or disease activity in well-controlled RA (286).

One study describes kinematic gait analysis difference in two groups of RA patients: (a) knee joint involvement and severe inflammation without progressive destruction and (b) knee joint involvement with progressive destruction. This analysis, which showed limitation in knee angle changes in swing and stance phases as well as a shortened swing phase duration in inflamed knee joints, can provide practical information about functional joint integrity in RA that could aid in therapeutic decision making (287).

In the presence of moderate to large joint effusion, knee flexion in excess of 20 degrees is associated with significant increased articular pressure, and caution must be observed in performing ROM exercises on a knee with significant fluid, as outpouching of the posterior joint space may occur, creating a popliteal or Baker cyst. Fluid from this popliteal portion does not readily return to the anterior joint space, but rather adds increased pressure to the popliteal space. Forceful ROM may also cause rupture of the capsule. There may be uncomfortable fullness or pain in the popliteal space, and rupture into the calf may simulate thrombophlebitis. Ultrasound can define a cyst. If rupture occurs, a hematoma may be seen below the malleoli. Observe the patient from the rear while he or she is standing to check for a popliteal cyst.

Meniscal cartilage and cruciate ligaments can be easily destroyed by proliferative synovitis. Collateral ligaments become stretched, causing valgus and varus deformity. Tests for knee stability are always indicated in an examination. X-rays should be taken in the standing position to assess the cartilage and joint space.

Treatment is directed to the particular focal intra- or periarticular problem. The patient should be instructed in early ROM exercise to preserve knee extension and flexion. Ninety degrees of knee flexion is needed to kneel, and 100 degrees is needed to climb stairs. A pillow under the knee at night is to be avoided because this will encourage a knee flexion contracture. Stretching of the hamstrings is important. Strengthening of the quadriceps mechanism in nonacute joints by isometrics and in nonacute or subacute joints by low- and moderate-intensity isotonics in 30 degrees flexion, if performed early in the disease process, helps to maintain the biomechanical advantage of the knee. Higher-intensity isotonic exercise should follow for the noninflamed joint.
Moderate to large effusions that inhibit contraction of the quadriceps and contribute to knee pain are best removed. Weight-bearing restrictions regarding the knee are indicated with acute flares. Bracing the knee for instability is possible (see the section "Orthotics"), but patient compliance is low with bulky braces. Inflamed periarticular or articular structures respond favorably to ice massage. When the joint is subacute or chronic, moist hot packs and TENS can be used.

Arthritis affects the MTP, talonavicular, and ankle joints in descending order. Ankles are less frequently involved than knees. Ankle damage is usually present in severe or progressive RA. Synovial involvement can be prominent and is seen anterior and posterior to the malleoli. The tarsal tunnel, posterior and inferior to the medial malleolus, contains the posterior tibial nerve, which is often painfully compressed by synovitis. In acute disease, stretching and erosion of collateral ligaments around the ankle occur, resulting in incongruity and usually pronation of the hindfoot. Subtalar joint involvement is common, and patients experience more pain walking on uneven ground. About 85% of patients with RA have forefoot problems, such as widening at the metatarsal area, prominent MTP joints caused by subluxed metatarsal heads, hammer toe deformities, and hallux valgus of the great toe. Areas of skin breakdown are common on the dorsum of the toes (i.e., hammer toes), and callus is seen under the MTP heads. Planter fasciitis and sub-Achilles bursitis may occur. Gait is typically flat-footed with little heel-strike or toe-off and a shuffling gait. Pronation of the hindfoot can be prominent. Appropriate footwear is extremely important. RA of the hindfoot and ankle can produce considerable dysfunction. A variety of nonoperative treatments may slow the progression of deformities, improve function, and provide symptomatic relief (288).

**PM and Inflammatory Myopathies**

DM-PM, IBM, and their subsets are systemic RDs that affect skeletal muscle. The clinical picture is predominantly one of profound weakness of the shoulder and hip girdle muscle, as well as of the neck and pharynx. In severe cases the diaphragm, intercostals, and abdominal muscles are involved. Twenty percent of patients with DM-PM also have some distal muscle weakness, and 50% of IBM patients have distal weakness. Some have weakness of the respiratory muscles. The muscle weakness is often compounded by steroid myopathy. A fair amount of muscle pain may be experienced when the inflammation is active. There may be complete remission of the disease, but chronic weakness is more common. Episodic periods of remission and exacerbation, which are often unpredictable, pose problems with functional activities and maintaining work status.

The inflammatory myopathies (18) can be divided into six major types:

- **Type I (PM):** insidious onset, beginning in the pelvic girdle and later progressing to the shoulder girdle and neck muscles. Weakened posterior pharyngeal and laryngeal muscles result in dysphagia and dysphonia. Remission and exacerbations are quite common. Moderate to severe arthritis as well as Raynaud phenomenon may be present. The skin over the knuckles and elbows is often atrophic.

- **Type II (DM):** acute onset. Proximal muscle weakness and an erythematous heliotropic rash on the skin of the eyelids and the dorsum of the hands are seen. Muscle tenderness is encountered in 25% of cases; subacute joint findings are common, as are systemic manifestations of malaise, fever, and weight loss.

- **Type III (cancer-associated myositis):** associated with malignancy and most common in men more than 40 years of age. Often, muscle weakness precedes the diagnosis of malignancy by 1 to 2 years. The muscle weakness is usually progressive and does not respond well to steroids. Dysphagia and respiratory muscle weakness are common events. The mortality rate is high, and death is often the result of respiratory failure and pneumonia.

- **Type IV (JDM):** involves children. The muscle weakness is rapidly progressive, and problems with dysphagia, dysphonia, and respiratory weakness are quite common. It is important to remember that late exacerbations occur after 7 years of remission. The propensity for the development of severe joint contractures and muscle atrophy is high. Skin problems in the form of calcinosis universalis (i.e., cutaneous and muscle calcification), particularly over bony prominences, contribute to skin breakdown, draining lesions, and joint contracture (289).

- **Type V (myositis associated with other collagen vascular diseases):** namely, RA, SLE, and PSS. The functional problems associated with the individual collagen disease often dominate the clinical picture.

- **Type VI (IBM):** most commonly involves men more than 40 years of age and has a slowly progressive course of muscle weakness. The quadriceps, hip, and shoulder girdle muscles often develop profound atrophy. In addition to proximal weakness, 50% of patients have significant distal weakness. Each subtype has specific problems and impairments. Rehabilitation of the patients with DM-PM is influenced by the specific subtypes. However, there are some common impairments in the subgroups on which rehabilitation goals and treatment can be formulated (289–292).

Newer studies have found that clinical and autoantibody subsets can determine 5-year survival. Existing clinical subsets and 5-year survival rate include cancer-associated myositis (55%), DM (80%), IBM (95%), and connective-tissue-related myositis (85%). Autoantibody subsets include anti-SRP (30%), antisynthetase group (65%), and anti-M2 (95%). These subsets also predict response to treatment. Subtypes based on autoantibody status present with a different set of problems (293). Antisynthetase syndrome consists of interstitial lung diseases, fever, arthritis, Raynaud phenomenon, and mechanic’s hands. The disease often has a rapid onset and aggressive course. The arthritis can affect the hands, knees, elbows, and shoulders, and be chronic and deforming. The lung disease can be severe and can substantially limit ADLs and mobility.

Anti-SRP syndrome is associated with initial severe muscle weakness, myalgias, and cardiac involvement that significantly
impacts function. Anti-M2 presents with the rash of DM and cuticular overgrowth, and responds well to treatment.

Rehabilitation intervention must be tailored to suit the needs of each patient, depending on the disease type and associated impairments, functional problems, and disability present. Patients with type V PM and associated collagen vascular disease (e.g., PSS, RA, SLE) have muscle weakness plus the added problems of the additional disease, which needs to be addressed from a rehabilitation standpoint.

**Survival and Prognosis**

Much has been learned about medical management and the efficacy of rehabilitation of patients with DM-PM in the past 10 years. In earlier years, DM-PM had a high mortality/morbidity rate, diagnosis was often made late, and steroids were the sole available treatment. Today better understanding of the types of PM, early diagnosis, expanded therapeutic regimens, and recognition of the importance of early rehabilitation intervention has helped decrease the mortality-morbidity rates.

General factors associated with poor survival are listed as follows: older age, malignancy, delayed initiation of corticosteroid therapy, myocardial involvement, pharyngeal dysphagia with aspiration pneumonia, steroid and immunosuppressive drug complications, and GI vasculitis (children).

Adults with type I and II disease can recover completely or be left with residual muscle weakness and fatigue, which can respond to rehabilitation management. The initial disease onset may be acute and require acute-care hospitalization for significant weakness with respiratory and swallowing difficulty. Rehabilitation must be provided from day 1 in a progressive manner. Transfer to a rehabilitation unit after acute-care management is appropriate (294). Patients with PM and associated incurable malignancy are not expected to recover from the disease. The rehabilitation goals are short term. Ambulation mobility and self-care functions progressively decline. Preserving ROM and strength will aid in keeping the person functional as long as possible. Not being able to continue in the work force for long is difficult for a middle-aged person to accept. Disability support payments and community support efforts need to be mobilized early on. Psychological support in coping with a chronic illness is needed. Good medical backup to contend with the problems of respiratory compromise and infection will be needed. Children with type IV disease need to be watched carefully for contractions.

Patients with type IV IBM have a slow, progressive course of proximal and distal weakness. Significant atrophy of the deltoid and quadriceps muscles is seen. Patients often have frequent falls resulting in fractures. The ability to increase the strength of significantly atrophied muscle has been shown to be poor (153). Lower-extremity braking to support a weak quadriceps mechanism is often needed.

**Problems and Interventions**

Scientifically proven impairments in isometric (153), isokinetic strength (155), and endurance exist. Also studies document decreased aerobic capacity in DM-PM adults (159) and children (160). The functional problems that arise depend on the muscle groups involved and the extent of the weakness. For example, weakness of the pelvic girdle muscles is associated with difficulty in rising from a chair or a prone position, difficulty going up stairs, difficulty getting in and out of a bathtub, frequent falls with difficulty returning to the standing position, waddling gait, and toe walking caused by heel cord tightness, which is common in children. Shoulder girdle weakness causes functional problems with dressing (e.g., difficulty pulling on shirt, hooking a bra) and grooming (e.g., combing hair, showering, shaving, brushing teeth), difficulty picking up heavy objects on a shelf, and difficulty eating. Neck weakness causes difficulty lifting and holding the head off a pillow and holding the head up while in a sitting position. Respiratory muscle weakness (e.g., intercostals, diaphragm) results in difficulty with respiration, causing shallow and sometimes inefficient respiration. Distal muscle weakness (seen in 10% to 50%) may cause foot drop and related ambulation problems and difficulties with hand function and activities (291).

Rehabilitation goals in the acute phase consist of maintaining ROM of the joints and preventing joint contractures (294,295). In the recovery phase, the goals are to increase and regain muscle strength, maintain ROM, return to functional ADL and ambulatory activities, and restore previous lifestyle activities as much as possible. UE ROM deficits are common in patients who have less than antigravity strength. The shoulder is most often involved in this case. Often knee flexion contracture occurs because of over pull of the stronger hamstrings in PM over the weaker quadriceps. Elbow, knee, and plantar flexion contractures are common in children, particularly those with skin/muscle calcinosis.

Patients with DM-PM/JDM should be on a basic ROM/stretching exercise program. These may have to be adjusted in patients with active calcinosis, as stretching could increase muscle inflammation and calcinosis. In JDM patients without acute calcinosis or rash, stretching exercise, casting, or dynamarSplints and aquatic exercise are appropriate. A recent text on the rehabilitation of persons with juvenile idiopathic inflammatory myopathy gives an excellent comprehensive overview of management of this patient group (292).

Isometric exercise is appropriate for DM-PM with inactive, mildly, and moderately active disease. Rest periods between contractions should be increased (20 seconds is appropriate) (153). Patients with significant atrophy are less likely to benefit. Therefore, early intervention is important. Clinically, low-weight concentric isotonic exercise is appropriate after a ROM and isometric program is in place in inactive and mildly active disease (173). (Both short- and long-term aerobic exercise are appropriate for inactive and mildly active myositis (174,175).) The efficacy of programs to increase the muscle strength by isometric (153), isotonic (185), and isokinetic exercise (191) without exacerbation of muscle inflammation, as determined by CPK, has been documented. The ability to increase aerobic capacity by short-term (6-week) and long-term (6-month) aerobic programs (176,186) has been documented.
Patients with significant quadriceps weakness often trip and fall. A study of an IBM patient yielded some reasons for this (296). Clinically bracing a patient with significant isolated quadriceps weakness on one side can help stabilize the gait pattern, make it safer, and reduce falls. The appropriate bracing in this instance is a short leg brace—either polypropylene or Klenzac—positioned in 5-degree plantar flexion to create a hyperextension moment at the knee that will stabilize the limb. In patients with bilateral quadriceps weakness, it is appropriate to brace the weakest knee. Both knees cannot be braced in this manner, as the center of gravity would be shifted posteriorly and compromise the patient’s balance. In the DM-PM patient who has both significant proximal and distal weakness, use of a 90-degree brace at the ankle or with a 10-degree dorsiflexion assist increases knee buckling. A 5-degree plantar flexion brace would stabilize the knee and basically correct the ankle foot drop. For those patients (less common) with only significant foot drop a 90-degree or a 10-degree dorsiflexion brace is appropriate.

If pharyngeal and laryngeal weakness is present, referral to a speech pathologist to teach the patient techniques to avoid aspiration of food and prevent respiratory infection is needed (292). Neck flexor muscle weakness has been shown to correlate directly with swallowing dysfunction in patients with DM-PM, and its presence may cue the clinician to obtain a speech and language consultation (297). If respiratory muscle weakness is present, chest physical therapy breathing techniques, proper positioning, suctioning, postural drainage, and breathing exercises, if the patient is not in the acute stage, are indicated. Tidal volume should be checked daily with a bedside spirometer. A collar may be provided to support the neck when neck flexor or extensor weakness is present.

For muscle pain, gentle muscle massage may produce a sedative relaxing effect on the muscle. The use of heat therapy for this is poorly described in the literature.

Most patients with PM do not have arthritis. PM associated with anti-JO-1 antibody is characterized by arthritis that may become chronic and deforming without bony erosions. It occurs in the wrists, hands, elbows, knees, and shoulders. DM-PM with associated RD (e.g., RA, SLE, PSS) is frequently associated with arthritis, which may be deforming; therefore, the use of modalities, splints, and joint conservation techniques is needed.

Iatrogenic steroid problems cause vertebral compression fractures in the thoracic and lumbar spine, resulting in back pain and muscle spasm. OP is also common in the long bones and joints. Interventions include a corset to decrease spinal mobility, heat modalities to relieve pain, and a long-handled reacher and shoehorn. If avascular necrosis of the femoral head causes pain in the hip and groin on weight bearing, unweighting the hip is indicated. Deep heat (i.e., ultrasound) to the hip may relieve pain. Isotonic and isokinetic strength exercise should be available with moderate or severe OP.

Steroid myopathy presents with atrophy of muscle and increased muscle weakness. Repeated enzyme tests, electromyography (EMG), and muscle biopsy should be performed. If the enzymes have remained normal, biopsy shows no active inflammation, and EMG is normal (only abnormal in rare cases of very severe steroid myopathy), the steroid should be reduced and the exercise program continued.

A number of problems can occur with the respiratory system. In addition to respiratory insufficiency, aspiration pneumonia may result because of weak pharyngeal and laryngeal muscles. Patients with the myositis-specific antibodies (MAS) anti-JO-1 commonly have interstitial lung disease. Primary interstitial fibrosis in JO-1 patients with PM associated with PM associated with SLE, RA, and PSS may have lung disease associated with these diseases. Pulmonary rehabilitation may be indicated.

Cardiovascular complications with DM-PM include congestive heart failure (3.3%), cardiomyopathy (1.3%), cot pulmonary (0.7%), and electrocardiogram abnormalities (50%) (17). They are most common in type I and type V disease. Rehabilitation includes cardiac precautions, energy conservation techniques, and an endurance program.

Dermatologic problems include pressure sores over bony prominences (e.g., sacrum, elbows, heels). Extensive calcinosis seen in childhood DM-PM causes breakdown of the skin over bony prominences of joints with drainage of calcium oxalate. Vasculitis with ulcerations of the fingertips and toes may occur in the overlap syndromes. Preventive measures include proper positioning, good nutrition with adequate protein intake, use of an egg crate mattress pad or waterbed, and padded support over elbows, knees, and heels. Restoration measures include appropriate treatment if deep pressure sores are present.

Raynaud’s syndrome precipitated by cold and stress is usually mild when it occurs, unless associated with collagen vascular disease. Symptoms include painful cold fingers and color changes from white to blue. Wearing gloves and using biofeedback have been useful.

Systemic Lupus Erythematosus
SLE is a chronic, autoimmune inflammatory disease that can affect any organ in the body. The most frequently involved sites are the skin, joints, pleuropericardium, kidneys, and central nervous system. Its course is varied in severity and duration.

The rehabilitation team is often consulted with respect to a number of functional problems, and fatigue is one of the most common. The causes of fatigue are multifactorial and may result from deconditioning resulting from decreased activity, and release of inflammatory cytokines known to cause muscle atrophy and joint inflammation. These inflammatory cytokines may also influence mood and behavior. Interferon, TNF, and IL2 are associated with fatigue and lack of motivation. Medication used in treating SLE, such as prednisone, is known to cause type 2 muscle fiber atrophy, which may be associated with muscle weakness. Several studies have shown that fatigue is more prevalent in those with SLE than in those without, and that characteristically, it is least on awakening and steadily increases as the day progresses (298,299). This pattern is clearly different from morning stiffness, characteristic of RA and may be a result of...
deconditioning rather than an inflammatory process. Physical findings in this patient group include muscle weakness, disruption of sleep/wake cycle, cognitive difficulty, and psychomotor slowing (300). Depression is reported in about 30% of patients with SLE (301). Fatigue is the best predictor of disease activity (300). Patients with SLE have reduced peak VO2 and exercise duration. They also have reduced muscle strength and reduced forced expiratory volume (FEV) (302). A pilot study of 15 SLE patients aged 12 to 19 showed moderate impairment in aerobic fitness compared to referenced controls. Sixty-seven percent reported significant fatigue by questionnaire. There was no significant correlation between fitness and fatigue. Neither fitness nor fatigue is correlated with disease activity or damage, or quality of life (303).

The symptoms prevalent in the SLE population (fatigue, deconditioning, and disability) respond to rehabilitation interventions. Central nervous system manifestations, such as poor concentration and memory impairment, do not. Isometric and aerobic conditioning programs are feasible in this population and effective in increasing fitness, strength, and functional capacity. Patients with SLE have been shown to have decreased aerobic capacity (300,302). An aerobic exercise program has been shown to increase endurance by 20% (156). A RCT of adult SLE patients showed significant improvement in exercise tolerance and aerobic capacity, quality of life, and depression after a supervised cardiovascular training program when compared to non-exercising controls (304). Treatment may include an energy conservation training program teaching that PA should be combined with frequent periods of rest. Naps are taken during the day, and sleep can be promoted by the use of relaxation tapes.

Pain is common in the small joints of the hands and feet because of arthralgias and arthritis. Joint pain can also result from avascular necrosis of bone. Joint deformity is also seen. Control of joint pain has been successful with acupuncture and acupressure techniques, heat, cold, and TENS (305). These techniques are more effective in treating arthralgias than in treating avascular necrosis, which requires unweighting of the lower extremity. When unsuccessful, joint replacement to control symptoms may be required, but should be undertaken with great caution because of the high degree of joint laxity (Jaccoud arthropathy) which is often difficult to correct surgically.

The rashes associated with lupus are usually not responsive to nonpharmacologic treatments, but the skin ulceration that can occur as a result of active Raynaud’s syndrome responds to hand-warming techniques (306). Temperature biofeedback has been used to relieve vasospastic disease. Patients with renal disease often have diminished stamina and fatigue. Improvement will occur with good blood pressure control and management of edema. In patients with nephrotic syndrome and significant edema, care must be taken to position the limb in the most functional way to minimize contracture. Compression pumps and garments can be used to help with and maintain the reduction of limb edema. Precautions must be taken for patients with cardiac failure. They must be compressed slowly or not at all because they may not tolerate any additional fluid load resulting from compression.

One of the major challenges to the rehabilitation team is the request to evaluate and treat patients with central nervous system involvement. Stroke, psychosis, depression, and memory deficits all have a significant impact on function. Treatment is aimed at the underlying problem. An uncommon but well-described problem is transverse myelitis. Management may require mobility aids, including a wheelchair. Reeducation in self-care skills and exercise to help promote stamina is needed. Spasticity may be controlled with oral antispasticity agents, or local neurolytic nerve blocks/chemodenervation. Treatment of flaccidity with braces or adaptive equipment should be offered. Speech therapy can provide strategies to enhance memory with the use of lists and cues.

In a study, in 109 SLE patients, of the impact of hand function on daily activities, 73% of them reported interference with ADL. Reduced grip strength, fumbling, and pain most commonly interfered with productivity (307).

The patient with SLE has to overcome major obstacles to successfully cope with this multifaceted illness. Support groups and family have been shown to be helpful in increasing compliance and are an essential component in the rehabilitation process.

Systemic Sclerosis
Management of PSS presents a significant challenge for rehabilitation specialists. The primary organ affected is the skin, but the pathogenesis is microvascular damage that can affect the kidney, the lung, and occasionally the heart (308). Whereas internal organ involvement usually requires pharmacologic treatment (blood pressure control, possible treatment of pulmonary fibrosis), musculoskeletal involvement leads to soft-tissue contractures and loss of ROM for which rehabilitation interventions are often effective.

The skin in those with SS is often shiny and bound down, and is associated with loss of soft tissue. This is most obvious with the hands, where fibrosis of the skin (with or without calcinosis) causes contractures, depigmentation, telangiectasias, ulceration, and loss of prehension. Raynaud phenomenon often accompanies these findings.

The rehabilitation approaches include use of heat modalities. Care needs to be taken to avoid overheating relatively underperfused soft tissue that results from the microvascular damage. Paraffin and moist heat are usually best tolerated. Treatment of Raynaud phenomenon includes topical creams, such as nitroglycerine paste, and temperature biofeedback, as well as electric mittens. In general, heat and stretch programs are recommended, but little data is available to establish efficacy.

Several studies have reported beneficial effects of TENS and acupuncture. Both techniques are thought to increase skin blood flow through increased vasculation (TENS) or activation of vasoactive peptides (acupuncture) (309).

Joint contracture is a frequent result of skin thickening and contraction at the joint. Disuse leads to muscle atrophy.
Muscle may also be involved from an inflammatory process that begins as a cellular muscle infiltrate causing myositis and may lead to fibrotic change in muscle. This may be painful and may result in weakness accompanied by a rise in CPK and other muscle enzymes.

Physical medicine and rehabilitative treatment relies on a heat and stretch program of active or active assistive ROM. Active contraction of muscle is usually needed to restore or preserve functional ROM. Usually isometric exercise has been employed, but isotonic and isokinetic exercise appear to be well tolerated in many with RD. Slow-rate isokinetic exercise (<60 degrees/seconds) may be very stressful on inflamed joints. Adjunctively, splinting of the hand or wrist can be added to the heat and stretch program. Dynamic splinting appears to be ineffective and poorly tolerated (310).

The peripheral nervous system is commonly involved in SS, with a motor-sensory neuropathy (311), and occasionally autonomic dysfunction (312). Carpal tunnel syndrome is most commonly seen, although weakness of the long flexors may be seen as well (312). Proper joint positioning, avoidance of compression at bony prominences, and short-term functional splinting may be useful.

The impact of SS on function is often profound and may reduce work capacity. Job modification, protection from cold, toxic exposures (urea formaldehyde, benzene, silica, etc.) and work-site modification may help preserve function and reduce risk of exacerbation. Adjustment to illness often depends on psychosocial support and preservation of function (313,314). Provision of treatment improves the disease outcome (314). A review of clinical management is available for the reader (315).

**Spondyloarthopathies**

AS and other related diseases (e.g., PSA, reactive arthritis, and IBD-related arthritis) constitute a group of disorders involving the axial skeleton, enthesopathy, and extra-articular manifestations (e.g., skin rashes/psoriasis, uveitis, and aortitis). They are known as spondyloarthopathies. The disability resulting from these diseases is usually attributable to the loss of spinal mobility and reduced pulmonary function. The impairments include pain at the enthesis, which is often worst at night and makes sitting difficult, and intervertebral fusion, producing substantial postural change.

Treatment for spondyloarthritis had typically consisted of NSAIDs, corticosteroids, and x-ray therapy, the latter two contributed to the development of OP. More recently, MTX, SSA and now TNF blockers such as etanercept have been used successfully.

Rehabilitation therapies have been aimed at reducing pain, preserving joint alignment and posture, and promoting independence in functional activities. Fitness training is essential to maintain independent function.

At disease onset the patient should be educated about maintaining mobility and preserving alignment. A ROM program designed to reduce hip flexor contractures and kyphosis should be initiated. Prone lying and sleeping without a pillow are encouraged. Some degree of muscle strengthening and especially abdominal breathing should be instituted (316–319). Often therapies are suggested in conjunction with therapeutic pool/spa treatment, which seem to provide some symptomatic relief (320). One report describes the use of a Jewett spinal orthotic as an effective treatment for increasing spinal mobility and reinstating the lumbar curve (321).

Convincing data has been published for treatment efficacy of an intensive 3-week inpatient hospitalization (319). Using the Bath AS global assessment, disease activity index, and visual analog scale for stiffness, improvements were demonstrated that were significantly better than those not treated for a similar 3-week period.

Patients with AS must be counseled about potential risk of cervical spine fractures that may result in tetraplegia. Cervical fracture is much more common than thoracic or lumbar. The likelihood of spinal cord injury is 11.4 times higher in this group than in those without AS (322). The cause of injury was usually falling (53% of cases). Patients must be educated about self-protection to try to minimize injury, especially in the house, and to avoid contact sports, drinking and driving, and to keep homes well lit in order to prevent injury.

Hip arthroplasty has been performed for patients with AS, as has wedge thoracic osteotomy to correct thoracic kyphosis (323). More recent RCTs have looked at the benefits of various combined exercise programs (combined flexibility, strengthening, breathing, and posture) and the latter further combined with aerobic exercise in supervised outpatient versus home settings. Control groups have been: the same exercise, conventional or no exercise. Results in the nonaerobic studies have indicated significant benefits in the areas of flexibility and function in both settings. In the aerobic incorporated study, only the supervised outpatient group had significant improvement when compared to the home group during similar exercises (324–329).

A recent 2008 Cochrane Database Review of eleven AS trials, summarizes the scientific evidence on the effectiveness of physiotherapy interventions (individualized home exercise vs. supervised no intervention; supervised group therapy vs. individualized home exercise; initial inpatient spa exercise followed by outpatient group therapy vs. outpatient group therapy, and experimental exercise vs. conventional exercise). The author concludes that individualized home-based or supervised programs are better than no intervention, supervised group physiotherapy is better than home exercise, and combined inpatient spa exercise followed by outpatient group exercise is better than group exercise alone (330).

Evaluation of occupational status and conditions is important in AS patients. A study of 397 AS patients showed that those with jobs that required dynamic flexibility had more functional limitations than those without this requirement (bath ankylosing spondylitis functional index [BASFI] 48.3 vs. 38.1). A significantly increased BASFI and radiographic damage score was seen with dynamic flexibility and exposure to whole body vibration jobs (331). Reviews of treatment for AS are available (323,332).
Juvenile Idiopathic Arthritis

JIA is an umbrella term now used to refer to a heterogeneous group of disorders of childhood onset that have in common chronic arthritis. The term JIA now replaces the older classification for JRA and juvenile chronic arthritis. JIA is defined by an onset before age 16 with arthritis affecting one joint or more joints for more than 6 weeks with no other known etiology (333). There are eight subtypes of JIA: systemic onset, oligoarticular persistent, oligoarticular extended, RF negative polyarthritis, RF positive polyarthritis, PSA, enthesitis-related, and undifferentiated. In general, the classification criteria for JIA are based on clinically presenting symptoms but also include exclusion criteria in defining each disease process. These exclusions include the following: (a) psoriasis or history of psoriasis in the patient or a first degree family member, (b) arthritis in an HLA-B27 positive male beginning after the sixth birthday, (c) AS, enthesitis-related arthritis, sacroilitis with IBD, RS, or an acute anterior uveitis, or a history of one of these disorders in a first degree relative, (d) the presence of IgM RF on at least two occasions in at least 3 months, (e) the presence of systemic JIA in the patient (Table 40-22). Each of these subtypes has an adult counterpart which may differ somewhat in presentation. JIA is more often manifested in large joints such as the ankles, knees, or wrists, unlike adult forms. Rheumatoid or subcutaneous nodules and RF are less likely seen in this population as well.

Early identification of the disease, advances in drug therapy, appropriate and well-timed surgical intervention, and early active ongoing team rehabilitation programs have contributed to better quality of life and functional outcome (334–336). Many patients who in the past were wheelchair mobile are now functionally ambulatory, with early diagnosis and intervention (337). A recent 2008 study of 106 JIA patients showed 66% have low PA levels (not meeting public health standards for moderate to vigorous exercise). Those with low PA were more likely to have decreased bone mass, strength, and function. Increased levels of moderate to vigorous PA may improve exercise capacity, function, and general quality of life (338). Adults who have had JIA have significantly lower rates of employment and exercise tolerance than age-matched unaffected controls (339).

JIA Categories

Systemic onset JIA is defined as arthritis in one or more joints with or preceded by fever of at least 2-week duration with at least 3 days of daily documented temperatures greater than 39°C, evanescent rash, generalized lymphadenopathy, hepatosplenomegaly or splenomegaly, or both, and serositis. In 50% of cases, five or more joints may be involved during some phase of the disease. This group comprises 2% to 17% of children with JIA and has a M:F ratio 1:1 with peak onset 1 to 6 years of age. They often have growth delays, osteopenia, anemia, leukocytosis, thrombocytosis, and elevated acute phase reactants while RF positivity is rare. In milder cases patient will respond well to NSAIDs but in more involved case will require initial intravenous steroids followed by tapering oral regimens of DMARD therapy, with cyclosporine and MTX commonly being given. The newer biologic agents, such as anakinra, that are being studied, may show more usefulness in systemic onset JIA (340). The long-term prognosis is better for those with less severe arthritis. Patients who have systemic features; more than 6 months of polyarthritis with hip involvement, and thrombocytosis, have a poorer prognosis overall, and more of these children have long-term disability (341).

Oligoarticular JIA presents in two ways: (a) persistent oligoarthritis affecting one to four joints during the first 6 months of the disease; (b) extended oligoarticular arthritis affecting greater than four joint in the first 6 months. Females are affected at a 4:1 ratio to males, with peak age onset around 6 years of age. It is the most common of the JIA subtypes, affecting 50% to 60% of all children with JIA. The knee is the most commonly involved joint, followed by the ankle. Oligoarticular JIA may also involve the small joints of the hand and temporomandibular joint (TMJ). Chronic anterior uveitis is seen in 20% to 30% of those with oligoarticular JIA (342). ANA positivity is seen in 50% to 70% of these children and has a high degree of association with chronic anterior uveitis. Treatment algorithms for these patients may begin with NSAIDs but due to a low level of responders in this group, intra-articular steroid injections are more commonly used. Those with little or no response in these cases are treated with disease-modifying regimens such as MTX or anti-tumor necrosis antibody agents. In most cases, these patients respond well to pharmacological therapies, as 68% of those with the persistent type, and up to 70% of those in the extended oligoarticular group show improvement when treated with MTX (343).

Polyarticular JIA is divided into two subgroups which include: (a) JIA RF positive and (b) JIA RF negative. Both of these processes present with five or more joints involved during the first 6 months of the disease. The first of these groups differ with the presence of a positive immunoglobulin-M RF on at least two occasions, at least 3 months apart. Female predominance is seen with ratios of 9:1 and comprises 10% of total JIA cases. It is usually more aggressive, symmetrically affecting small joints of the hands, wrists and some larger joints similar to the adult form of the disease. Ten percent of those affected may have rheumatoid nodules. RF negative polyarticular JIA has the greatest risk for chronic, severe arthritis. It accounts for 30% of all cases of polyarticular JIA with a 3:1 female predominance. The joint distribution can be symmetric or asymmetric, affecting large and small joints that may include the TMJ and cervical spine. Positive ANA may be seen in up to 40% of cases and a higher risk for anterior uveitis is present.
<table>
<thead>
<tr>
<th>Type Classification</th>
<th>JIA Diagnostic Criteria</th>
<th>Peak Age of Onset</th>
<th>Sex Ratio (F:M)</th>
<th>% JIA</th>
<th>Systemic Manifestations</th>
<th>Articular Involvement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Systemic</td>
<td>Arthritis with or preceded by daily fever of at least 2 wk duration, documented to be quotidian for at least 3 d, and accompanied by at least one of the following: rheumatoid rash, generalized lymphadenopathy, hepatosplenomegaly, anemia</td>
<td>2–4 y/o</td>
<td>1:1</td>
<td>10%</td>
<td>Fever, rash pericarditis, pleuritis, myalgias,</td>
<td>Knees, wrist, ankles &gt; fingers, C-spine</td>
</tr>
<tr>
<td></td>
<td>Exclusions: a, b, c, d</td>
<td></td>
<td></td>
<td></td>
<td>hepatosplenomegaly, anemia</td>
<td></td>
</tr>
<tr>
<td>Oligoarthritis, subcategory persistent</td>
<td>Arthritis in ≤ four joints at any time during the onset or course of the disease.</td>
<td>&lt;5 y/o</td>
<td>4:1</td>
<td>24%–58%</td>
<td>Anterior uveitis</td>
<td>Knees, ankles &gt; small joints of hand, TMJ, C-spine</td>
</tr>
<tr>
<td></td>
<td>Exclusions: a, b, c, d, e</td>
<td></td>
<td></td>
<td></td>
<td>(Same as persistent type) + wrist involvement</td>
<td></td>
</tr>
<tr>
<td>Oligoarthritis, subcategory extended</td>
<td>Arthritis ≤ four joints in the first 6 mo of disease but affecting cumulative total of ≥ five joints after the first 6 mo.</td>
<td>&lt;5 y/o</td>
<td>4:1</td>
<td>12%–29%</td>
<td>Anterior uveitis</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Exclusions: a, b, c, d, e</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Polyarthritis RF negative</td>
<td>Arthritis affecting ≥ five joints during the first 6 mo with negative tests for RF.</td>
<td>&lt;6 y/o (+ ANA) 7–9 y/o (− ANA)</td>
<td>3:1</td>
<td>10%–28%</td>
<td>Uveitis</td>
<td>Small and large joints, symmetric or asymmetric C-spine, TMJ</td>
</tr>
<tr>
<td></td>
<td>Exclusions: a, b, c, d, e</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Polyarthritis RF positive</td>
<td>Arthritis affecting ≥ five joints during the first 6 mo and positive test for RF at least twice 3 mo apart.</td>
<td>&gt;8 y/o</td>
<td>9:1</td>
<td>2%–10%</td>
<td>Rheumatoid nodules</td>
<td>Small joints (PIP, MCP, wrist, symmetric)</td>
</tr>
<tr>
<td></td>
<td>Exclusions: a, b, c, e</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enthesitis-related arthritis</td>
<td>Arthritis and enthesitis or arthritis or enthesitis plus any two of the following:</td>
<td>&gt;6 y/o</td>
<td>1:7</td>
<td>10%</td>
<td>Anterior uveitis, IBD</td>
<td>Lower limbs and some spine involvement</td>
</tr>
<tr>
<td></td>
<td>• Sacroiliac joint tenderness and/or inflammatory lumbosacral pain</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Positive HLA-B27</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Physician-diagnosed HLA-B27–associated disease in first- or second-degree relative</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Symptomatic anterior uveitis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Male &gt;6 y old at onset of arthritis or enthesitis.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Exclusions: b, c, d, e</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psoriatic arthritis</td>
<td>Arthritis and psoriasis or arthritis and at least two of the following:</td>
<td>7–10 y/o</td>
<td>2:1</td>
<td>2%–15%</td>
<td>Uveitis, psoriasis</td>
<td>Asymmetric small and medium-sized joints (knees, hands and feet) also large joints—knees and ankles</td>
</tr>
<tr>
<td></td>
<td>• Physician diagnosed psoriasis in first-degree relatives</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Dactylitis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Nail abnormalities (pitting or onycholysis).</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Exclusions: b, c, d, e</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Undifferentiated</td>
<td>Arthritis, but does not fulfill any of the above categories or fits in more than one category.</td>
<td>Variable</td>
<td>1:1</td>
<td>2%–23%</td>
<td></td>
<td>Variable</td>
</tr>
<tr>
<td></td>
<td>Exclusions: Not applicable</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Exclusions:
(a) Psoriasis or a history of psoriasis in the patient or a first-degree relative.
(b) Arthritis in an HLA-B27+ male beginning after the sixth birthday.
(c) AS, enthesitis-related arthritis, sacroiliitis with IBD, RS, or acute anterior uveitis, or a history of one of these disorders in a first-degree relative.
(d) Presence of IgM RF on at least two occasions at least 3 mo apart.
(e) Presence of systemic JIA in the patient.

TMJ, temporomandibular joint; PIP, proximal interphalangeal; MCP, metacarpophalangeal; JIA, juvenile idiopathic arthritis; ANA, antinuclear antibody; IBD, inflammatory bowel disease; RF, rheumatoid factor.
In both types of polyarticular JIA early treatment with MTX is recommended. In RF positive patients with signs of progressive disease, MTX in combination with anti-TNF antibody agents is recommended, as this may play a role in preventing bony erosions (344).

Psoriatic JIA, which is a combination of arthritis and psoriasis, and accounts for about 10% of the total JIA cases with a 2:1 female predominance. Diagnostic criteria include: (a) dactylitis, (b) nail pitting and onycholysis, and (c) psoriasis in a first-degree relative. The arthritis typically involves peripheral, asymmetric joints of the hands, feet, knees, and ankles. When joint involvement is limited, NSAIDs and intra-articular injections can be used initially. MTX is effective in treating skin involvement and anti-TNF antibody therapy is useful when the disease is most aggressive.

Enthesopathic JIA represents about 10% of all JIA, and is characterized by arthritis, enthesis, or both, in addition to any two of the following: (a) sacroiliac joint tenderness and/or inflammatory lumbosacral pain, (b) positive HLA-B27 antigen, (c) acute symmetrical anterior uveitis, (d) onset of arthritis or enthesitis in a male greater than 6 years old, or (e) physician-diagnosed HLA-B27–associated disease in first- or second-degree relative. The lower limbs and axial skeleton are most often affected. Other clinical manifestations include anterior uveitis, IBD, and a reactive arthritis. Treatment with NSAIDs for symptomatic relief of enthesitis is often given initially. Those with specific articular involvement may be treated with intra-articular corticosteroids, but may require DMARD therapy. MTX, SSA and more recently anti-TNF agents have been used as well.

Unclassified JIA are those subjects who do not fulfill any of the ILAR classification categories. This group comprises 2% to 23% of all cases at initial presentation of JIA. Of those classified as undifferentiated JIA, 60% showed no characteristic eligibility criteria for any of the other JIA categories while 40% had criteria from more than one specific category (345).

Special problem areas and impairments exist that deserve the attention of the rehabilitation specialist. Growth retardation, in general, may limit full stature. Abnormalities in growth related to specific joints result in a number of physical problems, cause impairment, and impact on function: short toes and fingers, leg length discrepancies, and micrognathia (i.e., small mandible). These abnormalities are due to premature closure of epiphysial plates caused by intra-articular inflammation disturbing the development of the growth plate. Iritis and blindness are other major physical problems. It is important to remember that impairments of joint contractures and loss of ROM and muscle strength occur rapidly in JIA and must be managed quickly and efficiently. Impaired aerobic capacity is also present early in the course of the disease (346–349). Treatment goals should be determined as soon as possible, and treatments implemented early in the disease course.

Specific Problems

Upper limbs. The wrist is often involved in JIA, and wrist flexion contracture can occur rapidly. A cock-up resting splint should be used at night. If the wrist is inflamed or forearm muscles are weak, contributing to wrist flexion, a functional splint should be used for activities. If a wrist flexion contracture is present, serial casting may be needed. If there is PIP involvement, the resting splint should include the hand as well as the wrist. With IP joint contractures, a dynamic outrigger splint should be used during the day. If the elbow is acute inflamed, an adjustable hinge splint can be used. ROM exercises to maintain extension, pronation, and supination are important. If a contracture exists, serial casting can be performed.

Spine. Every effort should be made to avoid flexion contracture at the neck. Proper positioning at night with the use of a single thin pillow, like a pediatric Wal-Pil-O (Roloke, Culver City, CA), is recommended. Special attention to the cervical spine disease in systemic type and polyarticular RF negative JIA should be noted. When there is acute pain, a soft cervical collar is worn. Sometimes torticollis becomes a problem, and a firmer Plastazote collar may help. A collar is recommended for desk work. A desk with a tilt top may reduce pain and help maintain more ideal spinal posture.

Lower limbs. Knee involvement should be managed promptly. If the joint is acute, a posterior resting splint should be used at night to prevent flexion contracture. If contracture already exists, a posterior splint may increase the danger of tibial subluxation and should not be used. Rather, the contracture should be reduced with a skin traction device or by serial casting. Occasionally, soft-tissue release is necessary.

Valgus deformity frequently occurs, and a supracondylar osteotomy may be needed to achieve realignment if conservative measures fail. A hip flexion deformity contributes to knee flexion problems, and care must be taken to maintain hip extension. In severe disease, knee joint replacement can be considered if pain management and function cannot be achieved by medical and rehabilitative means.

An acutely inflamed hip joint is most often associated with acute muscle spasm and rapid formation of a flexion contracture. Often skin traction is used during acute hip pain to prevent contracture, with the patient lying supine in bed; 1 kg of weight for each 10 kg of body weight is used. In the child with a tendency toward knee flexion contracture, use of light hip traction during the night reduces the chance of the formation of hip and knee contractures. There should be periods of lying prone during the day to encourage maintenance of hip extension. A prone lying board also may be used in bed. If hip contracture is not responsive to conservative treatment, a soft-tissue release may be needed. Joint replacement may be needed in severe disease and is usually performed before 16 years of age.

Ankle/foot. Particular attention should be directed to management of the foot. Use of the proper shoe type and orthoses, as well as ROM exercise, is important. Leg length discrepancy should be corrected with a shoe insert or built-up shoe if a greater than 3/8-in. correction is needed.

Since patients with JIA have been demonstrated to have decreased ROM, muscle strength, and aerobic capacity, exercise
in the form of ROM stretching and isometric, isotonic, and aerobic exercise are recommended early in the treatment program. Since most children lose strength rapidly around inflamed joints (particularly the knee), a few isometric contractions daily are recommended even with an inflamed joint. Both the parents and the child should understand treatment regimens to ensure compliance. Many treatments are performed at home with parent’s supervision.

Aquatic exercise has been shown to significantly improve hip motion (350). An 8-week weight-bearing conditioning program has been shown to increase the aerobic capacity in polyarticular JIA without disease exacerbation or increased pain (351). A RCT of 80 JIA children consisted of a low to moderate (as tolerated) aerobic experimental group, and a Qigong control group, both exercising three times a week for 12 weeks. Both groups showed significant improvement in childhood health assessment questionnaire (CHAQ) but not in aerobic capacity. Adherence was higher in the Qigong group (352). Children with JIA often have OP resulting from glucocorticoids, and weight-bearing activity has a positive influence on BMD (353).

The efficacy of exercise in RA has been reviewed but not in JIA. A recent 2008 Cochrane Review, up to October 2007, of RCTs on the effects of exercise on function, quality of life (QOL), and aerobic capacity in JIA was done. Three out of 16 identified studies met the selection criteria. Data was pooled for 212 JIA patients. The outcome measure favored a positive exercise effect but was not significant. Exercise in the excluded and included studies did not exacerbate arthritis. The author concludes that more “silver level” studies are needed to ascertain the short- and long-term effects of exercise for JIA patients (354).

Psychosocial factors. The disease has an impact on the child’s self-image, socialization, sexuality, and integration into school activities. Efforts should be made to keep acute admissions to hospitals at a minimum so the child can participate as fully as possible in school, family, and social activities. Home schooling is not advocated. It is best to keep the child active in school. He or she should be allowed to participate as much as desired but should be given clear guidelines about limitations. Particular advice should be given in regard to sports activities. Body contact and high-impact sports are to be avoided (football, soccer, running, and ballet jumps). Cycling and swimming are to be encouraged. Guidance and support are needed during adolescence to deal with issues of vocation and sexuality. Both children and parents should be educated in the disease and the benefits of treatment. One study shows parents in an education program improved significantly concerning their self-reported competencies on medical, exercise, pain, and social support issues (355).

Geriatric Arthritis
Care of the aged patient with arthritis is a challenge (356). Older people value the preservation of their function and independence. They are often retired or working less than full time, so that they may enjoy family, travel, and more leisure and recreational activities.

As we age, the body experiences physiologic changes of normal aging. These changes include decreased proprioception and spatial orientation, and difficulty in balancing and righting oneself, which may lead to falls. Other changes include muscle atrophy and decreased muscle power. The ability to sustain maximal muscle contraction is also reduced. The intervertebral disc becomes more desiccated. Some aged become less active and deconditioned (356). The elderly who age normally are quite functional in the community setting.

In addition to normal aging changes, many elderly have comorbid diseases (i.e., congestive heart failure, stroke, diabetes, atherosclerotic heart disease (ASHD), peripheral neuropathy, and emphysema). This necessitates the use of many medications. They have psychosocial challenges as a result of loss of job, decreased function, diminished income, and death of family members and friends, all of which lead to isolation and loneliness (357).

OA is the most common arthritis seen in the elderly. Involvement of the hips and/or knees can be associated with significant impairments and diminished function. The elderly also may have chronic RA, which had its onset earlier in life. In addition, patients 60 years or older can present with elderly-onset rheumatoid arthritis (EORA). This tends to have more equal gender distribution, higher frequency of acute systemic onset with involvement of the shoulder, a higher disease activity, and in the later stages more radiographic damage and functional decline. Efficacy and tolerability of second-line drugs is similar in both age groups, but in the elderly caution is needed with NSAIDs and prednisone (358,359).

OA coexists with RA often in the elderly. With long-standing RA or RA in combination with OA, mechanically deranged joints, ligamentous laxity, joint effusions, bursitis, and tendinitis are common. They may also have other musculoskeletal and noninflammatory disorders such as fibromyalgia, and neck and back pain syndromes. The elderly can have other inflammatory arthritides (manifested by more pulmonary and neurologic involvement than in younger patients) other than RA, including gout, pseudogout, and Sjögren’s syndrome.

The elderly with arthritis tend to be less physically active. Rehabilitation treatment goals in the elderly with arthritis are to reduce impairment, functional problems, and disability by relieving pain, fatigue, and psychological stress and by improving joint motion, muscle strength, and aerobic capacity. Achieving these goals will enhance safe mobility and ADL, and help prevent disability.

Treatments, both pharmacologic and rehabilitative, need to be altered in the elderly to accommodate for physiologic aging changes in body mechanics and pharmacokinetics, the effects of comorbid diseases, and the increased presence of biomechanically deranged joints and psychosocial issues attendant to the group (360,361).

In general, exercises could include ROM, isometric, low-resistance isometric, and low-intensity aerobic activities. Isometrics are best if ligamentous laxity and effusion are present. Precautions need to be taken with joint replacements, which are common in the elderly. Comorbid cardiac and neurologic disease also limits exercise intensity and duration.
Several RCTs show that regular exercise does not increase joint pain or accelerate disease progression. These studies suggest that exercise training may increase physiologic reserve and decrease dependency in older adults with joint disease (362). A program of progressive interval training with bicycles and step climbing in frail elderly RA patients for 45 minutes twice a week increased work capacity 76% without increasing disease activity and has been shown to be effective (363). Heat/cold before exercise, unweighting painful lower-extremity joints, and cognitive behavioral strategies often diminish pain and improve function. Removing moderate to large joint effusions and carefully injecting any inflamed bursae or tendon sheaths also contribute to pain relief. Very painful biomechanically compromised joints that cause increased functional loss should be considered for joint replacement.

Correlates of fatigue in this population are pain severity, functional status, sleep quality, female gender, comorbid conditions, and disease duration (364). Treatment designed to decrease pain and improve sleep using short rest periods and energy conservation techniques may help reduce fatigue.

Foot problems are very common in the arthritic patient and even more common in the elderly. Proper footwear should provide a crepe sole to provide absorption of the ground reaction force, a wide base of support to promote balance, good fit and even more common in the elderly. Proper footwear should provide a crepe sole to provide absorption of the ground reaction force, a wide base of support to promote balance, good fit and softness to relieve metatarsalgia. Foot hygiene and prompt care of skin breakdown are essential to prevent infection in a patient with possible vascular compromise of the lower extremity.

REFERENCES

4. Fries JF, Williams GA, Morfeld D, et al. Reduction force, a wide base of support to promote balance, good fit and softness to relieve metatarsalgia. Foot hygiene and prompt care of skin breakdown are essential to prevent infection in a patient with possible vascular compromise of the lower extremity.


130. Rothwell AG, Cragg KJ, O’Neill B. Hand function following.


132. Gill DRJ, Morrey BE. The Conrad-Morrey total elbow arthroplast in.


Although it is not as large a part of the practice of physiatry as other areas, such as pain management or the rehabilitation of neurological disorders, cardiac rehabilitation is an area that should be a part of the scope of practice of all rehabilitation specialists. There are large numbers of patients with primary cardiac disability, and the principles of rehabilitation can be applied to their care, with energy conservation, adaptive devices, and lifestyle modifications all being part of the appropriate care of these patients. Additionally, many of the other patients who are seen in rehabilitation have underlying cardiac disabilities, and the appropriate application of cardiac rehabilitation techniques can enhance the outcomes of rehabilitation of those patients.

The goals of this chapter are to discuss the common types of cardiac disease and the principles of rehabilitation for those conditions.

EPIDEMIOLOGY OF HEART DISEASE

Prevalence and Incidence of Cardiac Disease

Because of the lack of activity and obesity, cardiovascular disease is a leading health care issue in the United States. Fortunately, exercise is an important and effective part of the solution to the problems we have seen above and can have profound public health benefits. As an illustration, hospital discharges for people over the age of 65 show an incidence of 767.9/10,000 for heart disease, 175.6/10,000 for cerebrovascular disease, while all cancers combined account for 172.2/10,000 population (1). The savings from even a small reduction in these numbers would be astounding, not to mention that the individual patients would be saved from illness and pain. Reviewing these numbers should provide a serious motivation to individuals to improve their own health status with a program of vigorous exercise as part of a cardiovascularly healthy lifestyle.

Despite the attention given to many other medical conditions, cardiovascular disease is still the most common cause of death and disability in the United States. Some startling statistics regarding cardiovascular health in the United States can be gleaned from the CDC reports on health behaviors of adults. The picture on leisure time physical activity is very alarming and is an area that most people can make a significant impact on their health. Despite this fact, little has been done on a national basis to engage this issue, and compliance with exercise regime remains a national concern. In the 2002 to 2004 time period, 38% of adults never engaged in any light, moderate, or vigorous physical activity, with only one in eight adults engaging in any vigorous activity at least five times a week (2). Additionally, women and older people performed less activity, and people with lower educational status and socioeconomic status were less likely to engage in physical activity. In the area of body weight, the nation has had a progressive increase in weight over the last 50 years, and in the most recent data, this trend continues (3). Overall, nearly six out of ten adults were overweight or obese in the 2002 to 2004 survey, and 23% of adults were obese. Overweight status was most common in adults between the ages of 45 and 74, being much lower for adults over 75 and under less than 45. Again, higher education was associated with a better health status, with less obesity being associated with greater education. With regard to smoking, just over a fifth of all American adults were currently smokers (4). Cigarette smoking also was associated with the onset of the disease at a younger age—four in five smokers started smoking before the age of 21. Fortunately, there is increased motivation on the part of many individuals to stop smoking, and this can help to improve the success rate.

As a rehabilitation specialist, the opportunity to intervene with your patients is especially strong, since most patients have just undergone a life-altering disability, or are seeking advice for exercise related to injuries, and may be more receptive to appropriate counseling.

Types of Heart Disease

It is essential for a practicing physiatrist to be familiar with the many types of heart disease that might be encountered. For practitioners of cardiac rehabilitation, post–myocardial infarction (MI) patients are most common; however, improved survival and increased availability of advanced treatments have increased the frequency of post–coronary artery bypass graft (CABG) surgery, posttransplant, heart failure, arrhythmia, and postvalvular surgery patients. The details of cardiac rehabilitation for these different populations will be discussed later in this chapter.

As noted above, the incidence of cardiac disease has been lowered due to the recognition of cardiac risk factors and interventions to prevent ischemic cardiac disease. Specifically, decreased cigarette smoking, lower red meat consumption, and increased exercise have all contributed to a decrease in coronary artery disease.
PART III | MAJOR CONDITIONS

AN OVERVIEW OF CARDIAC REHABILITATION

Currently, in the United States, only 15% to 20% of the one million survivors of acute MI go on to receive a cardiac rehabilitation program (5,6). This is an improvement over the rates of involvement of 10% to 15% a decade or so before but is still very low (7,8). Even with these low rates of participation, the cost of these cardiac rehabilitation programs was an estimated $160 to $240 million annually in 1990 (9), and the costs would be far higher with better participation. Cost can be reduced somewhat with the introduction of home-based programs, and there are attempts to make programs better suited for older patients. Additionally, a significant cost benefit would be realized as recent meta-analyses of the effect of the cardiac rehabilitation programs have shown that cost savings per life year gained was estimated between $2,193 and $28,193 and cost savings per quality adjusted life year gained was between $668 and $16,118 (10). These benefits are realized through decreased health care utilization and improved mortality. Cardiopulmonary conditioning and improved survival are outcomes that are well documented by numerous studies (7–10).

Simply stated, the goals of cardiac rehabilitation are to restore and improve cardiac function, reduce disability, identify and improve cardiac risk factors, and increase cardiac conditioning (11–13). These goals are achieved through the use of a prescribed exercise and education program performed under the supervision of a team composed of physicians and health professionals. The primary outcome for patients with cardiac disease is the ability to resume activities of normal life without significant cardiac symptomatology. Although the general outline of the cardiac rehabilitation program is similar for all patients with cardiac disease, specific cardiac conditions will require refinements of the exercise prescription.

PRIMARY PREVENTION

As a rehabilitation specialist, it is essential to address lifestyle modification and education as parts of a complete cardiac rehabilitation program. Lifestyle modification is needed to address reversible cardiac risk factors, and education includes teaching patients about all cardiac risk factors. The goal is to achieve a program of cardiac risk factor modification (14). Irreversible and reversible cardiac risk factors are shown in Table 41-1. Irreversible risk factors are those that cannot be altered and include male gender, past history of vascular disease, age, and family history. Most irreversible factors are found through a thorough patient history. Where significant irreversible cardiac risk exists, early and aggressive attention to reversible risk factors becomes essential and can help to appropriately target interventions. Reversible risk factors for cardiac disease have been known for several decades and include obesity, sedentary lifestyle, hyperlipidemia, cigarette smoking, and conditions such as diabetes mellitus and hypertension (14–27). Modification of all these risk factors is an essential part of a cardiac rehabilitation program. Both patient and family education are necessary and can introduce lifestyle modifications as a part of the general heart healthy routine care of all family members. The rehabilitation team also needs to work closely with the primary care physician and elicit a cooperative relationship that can reemphasize the necessary lifestyle modifications. With the disabled population, this is especially important, as a risk factor of relative immobility is often present, and thus more attention needs to be paid to the other modifiable risk factors.

REVIEW OF INDIVIDUAL RISK FACTORS

Diabetes

Diabetes is one of the most potent risk factors for the development or reoccurrence of ischemic cardiac disease. Close control of blood sugars has been shown to decrease the risk of cardiac disease through the slowing of the development of atherosclerosis and lowering the incidence of secondary conditions such as nephrogenic hypertension (28,29). In addition to oral hypoglycemic medications and the use of insulin, a combination of exercise training, weight loss, and dietary modification can assist in improving diabetic control (30). The appropriate selection of treatments for an individual patient can be helped with following the American Diabetes Association guidelines, and early intervention can be an essential component of the prevention of later cardiac disease. The exact benefits of exercise training in combination with good glucose control are still being elucidated, but they are present. Essentially, prevention of development of a combination of diabetes, hypertension, dyslipidemia, and obesity is essential as the combination, called the multiple metabolic syndrome, can lead to increased incidence of heart disease.

| TABLE 41.1 Risk Factors for Coronary Artery Disease |
|----------------------------------------|-----------------------------|
| Irreversible Risks                      | Reversible Risks            |
| Male gender                            | Cigarette smoking           |
| Family history of premature CAD (before age 55 in a parent or sibling) | Hypertension                |
| Past history of CAD                    | Low HDL cholesterol (<0.9 mmol/L [35 mg/dL]) |
| Past history of occlusive peripheral vascular disease | Hypercholesterolemia (>5.20 mmol/L [200 mg/dL]) |
| Past history of cerebrovascular disease | High lipoprotein A           |
| Age                                    | Abdominal obesity           |
|                                        | Hypertriglyceridemia (>2.8 mmol/L [250 mg/dL]) |
|                                        | Hyperinsulinemia             |
|                                        | Diabetes mellitus            |
|                                        | Sedentary lifestyle          |
|                                        | Metabolic syndrome           |

| TABLE 41-1 |
|-------------|-------------|
| Irreversible Risks | Reversible Risks |
| Male gender | Cigarette smoking |
| Family history of premature CAD (before age 55 in a parent or sibling) | Hypertension |
| Past history of CAD | Low HDL cholesterol (<0.9 mmol/L [35 mg/dL]) |
| Past history of occlusive peripheral vascular disease | Hypercholesterolemia (>5.20 mmol/L [200 mg/dL]) |
| Past history of cerebrovascular disease | High lipoprotein A |
| Age | Abdominal obesity |
| | Hypertriglyceridemia (>2.8 mmol/L [250 mg/dL]) |
| | Hyperinsulinemia |
| | Diabetes mellitus |
| | Sedentary lifestyle |
| | Metabolic syndrome |
Hypertension
Establishing adequate control of blood pressure is an important part of the management of individuals with cardiac disease. Although control of hypertension has been shown to be clearly beneficial in the prevention of stroke, the data for heart disease have been more mixed. Still, it is important to control hypertension in postinfarct patients and in patients with risk for cardiac disease. Historically, blood pressure control has been shown to be most useful for patients with normal electrocardiograms (31,32).

Lifestyle modification can provide two of the most important factors in the control of hypertension: (a) reduction of salt in the diet and (b) increasing exercise to improve conditioning. Although there are many classes of pharmacological agents available for the control of hypertension, there has not been a clear benefit shown with the use of one type of agent over another except in some special situations (32). The major groups of medications for the control of hypertension are divided into beta blockers, alpha blockers, diuretics, calcium channel blockers, and ACE inhibitors. The agents that used to be believed to be most beneficial of these were the beta blockers. This is still the case in individuals with CHF, arrhythmia, hypertrophic obstructive cardiomyopathy, and prior MI. In this group of patients, these agents provide cardiac protection by decreasing the maximum cardiac oxygen consumption and through decreasing inotropy and limiting heart rate (HR) response. However, in individuals with only hypertension and no other cardiac disease, the most recent evidence indicates that there is an increased risk of stroke and no clear benefit over other agents in the prevention of cardiac disease. The current recommendations are for beta blockers not to be used as monotherapy or as a first-line agent for uncomplicated hypertension (33,34). Diuretics have been shown in large trials to have beneficial effects on decreasing mortality, especially in isolated or uncomplicated hypertension (33). Consideration needs to be made for special populations in the treatment of hypertension. For example, hypertension is more prevalent, and severe, and occurs at a younger age in African Americans (35). The issue of which agent is most effective is often not a pertinent discussion, since combination therapy is usually required to adequately control blood pressure (36). The standards for the management of hypertension are evolving rapidly, and the latest European guidelines issued in 2007 by the European Society of Hypertension and the European Society of Cardiology (ESH-ESC) (37) are likely to be incorporated in the next revision of American guidelines that are issued periodically by the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure (36). The current guidelines are from 2003 and are due for revision. Since there is so much change in the management guidelines, it is recommended that rehabilitation physicians seek the advice of the treating cardiologist or internist for assistance in the optimal management of each individual patient.

Hypercholesterolemia
Elevated levels of serum cholesterol are a modifiable risk factor that has received a great deal of attention. In the popular press and among patients, this is the most commonly discussed risk factor. A combination of dietary modification, exercise, and medications can be very effective at controlling hypercholesterolemia. Decreased cholesterol levels and increased high density lipoprotein (HDL) have been known for a long time to be associated with decreased risk of cardiac disease (38–40). The initial approach to mild elevation of serum cholesterol is through a decrease in dietary intake of saturated fats and cholesterol, with the goal of lowering serum lipid levels, and thereby decrease cardiac risk. Even a moderate risk factor modification program can help.

Patients can decrease lipid levels by adhering to a low-cholesterol, low-fat diet along with weight reduction, even without the addition of exercise (41). Current recommendations are that the total amount of calories from fat in the diet should not exceed 30%. Control of cholesterol can be achieved through a three-step program, as outlined in the NCEP guidelines (35). Phase I is an adoption of nutritional guidelines, lifestyle changes, and general improvement in health habits. Phase II involves the addition of fiber supplements and possibly nicotinic acid. Phase III includes lipid-lowering drugs. Lipid-lowering programs have been shown to retard the progression of coronary artery disease. With the addition of physical activity, HDL cholesterol concentration can rise 5% to 16%, but the data on the lowering of low density lipoprotein (LDL) cholesterol are still controversial.

The Metabolic Syndrome
Any current discussions of risk factor modification must involve the metabolic syndrome, which is defined by a combination of a constellation of risk factors. These risks include hypertension, abdominal obesity, dyslipidemia, and insulin resistance. The definition of metabolic syndrome was formalized by the National Cholesterol Education Program in 2001 with the requirement of the presence of three or more risk factors out of five (41) (Table 41–2). The prevalence of the syndrome from the NHANES III survey was 23.7% in 47 million individuals (42). This incidence has likely only increased since

| Table 41.2 Identifying Patients with the Multiple Metabolic Syndrome (Diagnosis Based on Three or More Items Being Present) |
|-----------------|-----------------|
| **Risk Factors** | **Threshold Level** |
| Men and women   |                 |
| Fasting glucose | ≥110 mg/dL      |
| Blood pressure  | ≥130/85 mm Hg   |
| Triglycerides   | ≥150 mg/dL      |
| Women           |                 |
| Abdominal obesity (measured by waist circumference) | >88 cm (>35 in.) |
| HDL cholesterol | <50 mg/dL       |
| Men             |                 |
| Abdominal obesity (measured by waist circumference) | >102 cm (>40 in.) |
| HDL cholesterol | <40 mg/dL       |
the population has continued to have increased the incidence of obesity. Presence of the metabolic syndrome is associated with a fourfold increase in fatal CHD and a twofold increase in CVD and all-cause mortality, even after adjustment for age, LDL, smoking, and family history of CHD (43); in women, it is associated with a fivefold to ninefold increase in diabetes (44,45). The major issue in the prevention of the metabolic syndrome is the control of weight gain and requires changes in a variety of behaviors, including increasing activity, decreasing sedentary activities, and reducing caloric intake. Reduction of salt intake is also essential for these individuals (46).

**Obesity**

The importance of weight management has become clearer recently with the clarification of the role of central obesity in the metabolic syndrome. Weight management is an integral part of any cardiac rehabilitation program for individuals who are overweight, and there are many outside programs that can be used to help with weight management. Dietary counseling is also important, and specific counseling is based on lipoprotein levels, blood pressure, presence of diabetes or heart disease, and other risk factors. As little as a 5 lb weight loss can be associated with a 40% reduction in cardiovascular risk, according to the Framingham Heart Study (47). A 10% weight reduction can lead to a significant reduction in a number of cardiac risk factors (48). As a rehabilitation specialist, it is important to emphasize the importance of exercise and activity on the loss of weight and maintenance of target weight. The benefits of improved lipid profile and exercise have already been discussed above.

**Cigarette Smoking**

Cigarette smoking is one of the greatest single modifiable risk factors for cardiac disease (49,50). There are significant benefits to smoking cessation, even as secondary prevention. Ten year mortality in individuals with angiographically demonstrated coronary artery disease or MI who stopped smoking is decreased by over 30%. Part of the mechanism of smoking induced risk is through accelerated atherosclerosis, and as a contributor to hypertension. In evaluating techniques to help with smoking cessation, exercise alone does not contribute to decreased smoking (50), and smokers tend to be less compliant in cardiac rehabilitation programs (50). However, in a program of cardiac rehabilitation coupled with counseling for smoking cessation, with appropriate medication use, a decrease in smoking has been demonstrated (50). Smoking cessation is so critical for survival, it is essential to include cessation or enrollment in a cessation program counseling as part of a complete cardiac rehabilitation program.

**REVIEW OF CORONARY ANATOMY AND PHYSIOLOGY**

In order for the rehabilitation specialist to be able to better care for individuals with cardiac disease, it is essential that a basic understanding of cardiac physiology and anatomy be maintained. This is important for communication with cardiologists and cardiothoracic surgeons, as well as being an essential foundation that allows for better communication and education with your patients. This is not meant to be an exhaustive discussion; rather, it is a starting point from which the interested clinician can then increase his or her knowledge.

**Cardiac Anatomy**

Of particular importance is a familiarity with the normal distribution of the major arteries of the heart, cardiac valvular anatomy, and the structures at risk from ischemia or infarction in these distributions. Comfort with cardiac anatomy also facilitates an informed dialogue with the referring cardiac specialists and gives the rehabilitation specialist an ability to anticipate complications and problems associated with the specific design of an exercise program specific to the patient’s cardiac disease.

Overall, the heart consists of paired atria and ventricles, with deoxygenated venous blood entering the right atrium. Blood then flows into the right ventricle via the tricuspid valve and is then pumped out through the pulmonary valve into the pulmonary artery. After oxygenation in the lungs, the blood reenters the left atrium and passes through the mitral valve into the left ventricle. The left ventricle then ejects the blood through the aortic valve into the aorta and the systemic circulation. The heart valves ensure unidirectional flow of the blood, and the atria act in coordination with the ventricles to augment cardiac output (CO). Atrial contraction assists in diastolic filling of the ventricles, and can add up to 15% to 20% to the total CO. The benefits of atrial function are greater with increased HR and in conditions with decreased ventricular compliance (51). The loss of the contribution of atrial “kick” is especially important to consider in disease conditions where atrial dysfunction is seen, such as atrial fibrillation.

The cardiac conduction system is a specialized system of muscle cells (myocytes) adapted to facilitate the appropriate sequencing of the contraction of the atria and ventricles at the physiologically appropriate rate (52) (Fig. 41-1). It is a unique feature of all cardiac muscle cells that they have the intrinsic ability to contract. Additionally, myocytes are able to conduct electrical activity in order to synchronize their contractions with other cells. The usual pacemakers of the heart, with cells with the most rapid intrinsic conduction rate, are the cells of the sinoatrial (SA) node. The SA node is located in the right atrium. The electrical signal, or pulse, then travels through three atrial internodal pathways to the atrioventricular (AV) node. The AV node has a special quality of delayed conduction allowing for the sequential contraction of the atria, followed, after a short delay created by the AV node, by the contraction of the ventricles. The signal, after passing through the AV node, passes into the bundle of His, located in the intraventricular septum, which then divides further into left and right bundles. The left bundle has a final division into anterior and posterior fascicles. Terminal branches of both the right and left conduction systems carry the pulse signals that excite the myocytes, causing contraction. MI, aging, and other conditions can alter the conduction system, leading to a variety of conditions such...
as heart block and sick sinus syndrome. Congenital defects and accessory tracts can lead to life-threatening arrhythmias such as the Wolff-Parkinson-White syndrome (WPW). When prescribing or supervising a program of cardiac rehabilitation, it is important for the clinician to have a good understanding of the conduction system. This knowledge is helpful in evaluating arrhythmias and assessing the risks faced by patients prior to the initiation of cardiac conditioning programs.

Variation of Arteries

Another important part of cardiac anatomy that is important for the clinician practicing cardiac rehabilitation is the usual anatomy of the cardiac arteries. Normally, there are left and right coronary arteries arising from the base of the aorta in the left and right aortic sinuses. The left main coronary artery usually divides into the left anterior descending and the circumflex arteries, while the right coronary artery continues on as a single vessel. The standard distributions of the vessels and their rates of occurrence are seen in Table 41-3 (53). The most common anatomy of the coronary vessels is right dominant circulation, seen in 60% of individuals (see Table 41-3). When the posterior descending artery arises from the left circumflex, as seen in 10% to 15% of individuals, this is called left dominant circulation. For approximately 30% of individuals, the posterior descending arises from the left circumflex and right coronary arteries, in what is described as balanced circulation. The anatomy and the distributions of infarcts and associated cardiac syndromes usually seen are described in Table 41-4.

Cardiac Physiology

Cardiac myocytes are among the most metabolically active tissues in the body. In order to allow for this extremely high level of metabolic activity, oxygen extraction is nearly 65% at all levels of activity (compared to 36% for brain and 26% for the rest of the body) (54). The heart is most efficient at aerobic metabolism, but is able to perform both anaerobic and aerobic metabolism, using a variety of substrates. Carbohydrates are usually 40% of the metabolism, with fatty acids making up most of the remaining 60% metabolism (55). This high oxygen extraction and metabolism presents a relatively high risk for ischemic injury to cardiac myocytes, since coronary blood flow is only present during diastole. This is especially important for

![Figure 41-1. Schematic illustration of the conduction system of the heart with labeling of the key conduction pathways. SA node, sinoatrial node; AV node, atrioventricular node.](image)

<table>
<thead>
<tr>
<th>TABLE 41.3 Coronary Artery Anatomy</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Arteries</strong></td>
</tr>
<tr>
<td>Right coronary artery</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Left coronary artery</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
the endocardium, where increased wall tension and myocardial hypertrophy may place these myocytes at greater risk. Given the near maximum extraction of oxygen, there are only limited ways to increase oxygen supply in a situation of decreased cardiac perfusion. Under normal conditions, the coronary arteries can dilate to meet the demands of exercise. There are a number of substances secreted by the body that can increase the coronary blood flow, with nitric oxide as the final agent of many pathways (56). However, in disease vessels or situations of myocardial hypertrophy or excessive wall tension, this mechanism may not be able to allow for sufficient perfusion, and ischemia can result. Since it is so critical to restore or preserve myocardial perfusion, most medical and surgical therapies aim to restore the normal blood flow to the myocardium, either through vasodilatation or through bypass or endovascular procedures. It is also important to include exercise in the treatment regimen since regular exercise can increase cardiac collateral circulation and improve arteriolar vasodilation. These improvements are routinely seen as a result of the exercises that are a part of a cardiac rehabilitation program. Another consideration in maximizing cardiac function is maintaining adequate venous pressures to the right side of the heart, without overloading the ventricle. The CO is in part related to increase in venous return, which increases the length of the myocardial fibers in diastole prior to the initiation of cardiac contraction. Clinicians who work with patients with cardiac disease will often refer to maintaining this filling pressure as "preload." The benefit of myofibril stretch is to increase the overlap of the actin and myosin fibers in order to maximize the strength of contraction. However, excessive dilation of the ventricle, with further stretching in a weakened myocardium can cause the overlap of myosin and actin to decrease, yielding a decline in the strength of ventricular contraction. The relationship between the length of the fibers and the filling of the ventricle, which leads to increased contractility, is described by the Frank-Starling curve (Fig. 41-2). The clinical effects of this overlap of myosin and actin and the resultant decreases in cardiac function are seen in both constrictive heart disease (which limits the ability to move to the right on the Frank-Starling curve) and in patients with dilated cardiomyopathies, where there is a decrease in CO due to ventricular dilation moving too far out to the right on the curve. In cases of constriction, surgery can be done to allow greater dilation of the ventricle and restoration of CO, and in dilated heart failure, therapies are directed to decrease the size of the ventricles in order to increase CO (57).
In order for clinicians practicing cardiac rehabilitation to be able to discuss the basic principles of aerobic training and cardiac conditioning there is a need to have a basic understanding of the terminology and the principles of exercise physiology. These will be presented here.

**AEROBIC CAPACITY**

Aerobic capacity (VO$_{2\max}$) is the ability of the individual to perform exercise. It is a measure of work output and is analogous in some ways to a horsepower rating of an engine. Simply viewed, it is the work capacity of an individual. Aerobic capacity can be expressed simply as the oxygen consumed (liters of oxygen per minute, or more commonly, it is expressed in milliliters of oxygen per kilogram per minute corrected for weight). Oxygen consumption (VO$_2$) has a linear relationship with workload, increasing up to a plateau which occurs at the VO$_{2\max}$. VO$_2$ is measured through the analysis of expired gases, and for a given level of submaximal exercise, VO$_2$ reaches steady state after approximately 3 to 6 minutes of exercise. The slope of the line between VO$_2$ and workload represents the mechanical efficiency of the activity being performed. In conditions such as orthopedic limitations, deconditioning, or neurologic disorders, decreased efficiency is represented by an increase in the slope of VO$_2$ and work. A useful measure is to define submaximal effort as a percentage defined by VO$_2$ divided by VO$_{2\max}$. The use of percent VO$_{2\max}$ allows for normalization of data across individuals and for comparison of activities. VO$_{2\max}$ has been demonstrated to decrease with age in longitudinal studies such as the Baltimore Longitudinal Study of Aging (58) (Fig. 41-3).

**Heart Rate**

HR is a useful measure to guide exercise as it has a linear increase in relation to VO$_2$ or other measures of work. Maximum HR is determined by age and can be roughly estimated by subtracting the age of the individual in years from 220. The Karvonen equation is another equation to estimate peak HR and target HRs that takes resting HR into account. The slope of the line between HR and VO$_2$ is determined by physical conditioning and the maximum HR continues to decline with age even with ongoing exercise. The physiologic regulation of HR is mediated by the interaction of vagal and sympathetic tone and circulating catecholamines (Fig. 41-4).

**Stroke Volume**

Stroke volume (SV) is the quantity of blood pumped with each heartbeat. Since the heart is a muscle and strengthens with exercise, in a patient with normal myocardial function, SV can increase with exercise. During incremental exercise, SV increases the most during early exercise, with the major determinant of SV being diastolic filling time. SV is sensitive to postural changes, changing little in supine as it is near maximum at rest, while in erect position it increases in a curvilinear fashion until it reaches maximum at approximately 40% of VO$_{2\max}$ (Fig. 41-5). There is also a decreased response of SV seen with advancing age and in cardiac conditions which result in decreased compliance, such as left ventricular hypertrophy.

**Cardiac Output**

CO is the product of the HR and SV. CO increases linearly with work, and in early exercise the principal increase is via

---

**FIGURE 41-3.** Schematic illustration of the plateau of oxygen consumption while workload continues to rise at maximum exercise. VO$_2$ = work in volume of oxygen consumed.

**FIGURE 41-4.** Schematic comparison of heart rate response versus workload for normal, conditioned, and deconditioned individuals.
the Frank-Starling mechanism (SV increase), while in late exercise it is predominantly increased by HR. In general, the relationship between CO and VO$_2$ is linear with a break in the slope at the anaerobic threshold. The anaerobic threshold is the level of exercise at which the ability to deliver oxygen to the exercising muscles is below the demand for oxygen, marking the transition from aerobic to anaerobic metabolism. The maximum CO is the primary determinant of VO$_{2\max}$ and declines with age without any change in linearity or slope. The CO seen in submaximal work is parallel but lower in upright work compared to supine work, with VO$_{2\max}$ and maximal CO less in supine than erect positions (Fig. 41-6).

### Myocardial Oxygen Consumption

Myocardial oxygen consumption (MVO$_2$) is the oxygen consumption of the heart. MVO$_2$ rises in a linear fashion with workload. The anginal threshold is the point where MVO$_2$ exceeds the maximum coronary artery oxygen delivery. Although MVO$_2$ can be determined directly with cardiac catheterization, this is not practical. MVO$_2$ is usually estimated using the rate pressure product (RPP), calculated as the product of the HR and the systolic blood pressure (SBP) divided by 100. The increase in RPP with some activities versus others explains some of the seemingly paradoxical findings in patient symptoms. For example, activities with the upper extremities and exercises with isometric components to them have a higher MVO$_2$ for a given VO$_2$ due to higher SBP for a given level of work. Activities performed in supine also demonstrate a higher MVO$_2$ at low intensity and a lower MVO$_2$ at high intensity when compared to activities performed in the erect position. Finally, the MVO$_2$ increases for any activity when performed in the cold, after smoking, or after eating (Fig. 41-7).
AERobic training is the term for physical exercises which are performed in order to increase the cardiopulmonary capacity (VO2max). The basic principle to effectively perform aerobic training needs to take into account four areas in prescription: intensity, duration, frequency, and specificity.

Intensity of training is defined by either the physiologic response of the individual, or the intensity of the exercise performed. For example, a program of exercises may be aimed at a target HR or RPP or at a level of exercise intensity such as the speed and incline setting for a treadmill exercise. Usually, a target heart rate is the most simple for writing exercise prescriptions for an individual. Often target HR can be set at 80% to 85% of the maximum heart rate determined on a baseline exercise tolerance test (ETT). All exercises that evoke 60% or more of the maximal heart rate will have at least some training effect.

Duration of training is essential to establish the overall conditioning. A usual cardiac conditioning exercise program is 20 to 30 minutes long, with a 5 to 10 minute warm up period before exercise, and a 5 to 10 minute cooling down period after exercise. It is usually understood that exercise at lower intensity will require a longer duration to achieve a similar training effect to exercise at higher intensity.

Frequency of training is defined as how often exercise is performed over a fixed time period, and is usually expressed in sessions per week. At a minimum, training programs should be done three times per week. With low intensity programs, an increase to five times per week may be required to offset the decreased intensity of training.

Specificity of training refers to the performance of activities in training that are the same as those desired. It is essential to remember that training benefits are most applicable to the specific activities that are performed. For example, upper extremity ergometry will not as efficiently alter the cardiac response to walking as a treadmill training program. Specificity dictates that the design of a training program needs to consider the activities and muscle groups exercise based on the needs of the particular patient, based on known vocational and recreational activities. This is often called the law of specificity of conditioning, and is commonly referred to in cardiac conditioning programs (59).

Aerobic training causes benefits in a number of physiologic parameters as discussed below (52).

Aerobic capacity: The maximum aerobic capacity (VO2max) of a patient will increase with training. The resting VO2 does not change, and the VO2 at a given workload does not change. The changes are also specific to the muscle groups that are trained.

CO: The maximum CO increases with aerobic training. The resting CO does not change, but the HR at rest will decrease, and the SV at rest will increase, leading to a lower MVO2 at rest and submaximal exercise. The CO is directly related to VO2 at rest and at a given workload, up to the anaerobic threshold. The maximum CO increases with aerobic training. The direct relationship between VO2 and CO does not change during training.

Heart rate: The HR after aerobic training is lower at rest and at any given workload. The maximum HR is not changed, as the maximum HR is age determined.

Stroke volume: The SV is increased at rest and at all levels of exercise after aerobic training. It is the increase in SV that allows for maintenance of CO at a given workload with the decrease in heart rate described above.

Myocardial oxygen capacity: The MVO2 response to aerobic training is the most valuable part of training in cardiac rehabilitation. The maximum MVO2 does not usually change, since it is determined by the anginal threshold. However, at any given workload, the MVO2 decreased with training. This can allow individuals to markedly increase their exercise capacity and can lead to a marked improvement in function. After training, a patient will be able to perform more activities at an MVO2 below the anginal threshold. This will lead to less symptoms and increased safety to avoid myocardial injury than before training. Pharmacological interventions or revascularization procedures can also improve maximum MVO2.

Peripheral resistance: The peripheral vascular resistance (PR) decreases in response to exercise training. The PR is responsible for increases in systolic pressure and is a major contributor to myocardial wall tension, an important factor in limiting myocardial blood flow. The peripheral resistance is also often referred to by individuals involved in cardiac practice as “afterload.” The PR is decreased at rest and at all levels of exercise after a conditioning program. This response of the peripheral vasculature is due to the increased vasodilation in peripheral vascular beds. This results in a lower RPP and a lower MVO2 at a given workload and at rest.

It is important for rehabilitation specialists working with patients with cardiac disease to remember the basic physiology described above. It is even more important to be able to describe the benefits of training to the patients so that they understand the importance of exercise as a part of their treatment regimen. The benefits of cardiac conditioning are seen in two major areas: reduced cardiac risk and improved cardiac conditioning. The reduction of cardiac risk from cardiac rehabilitation has been demonstrated in numerous studies. As examples, from as long ago as 1989, pooled data from 22 randomized studies of exercise in 4,554 patients following acute MI demonstrated a 20% to 25% reduction in all-cause mortality, fatal MI, and cardiac mortality in a 3-year follow-up (60). The benefits of cardiac rehabilitation apply to the elderly (61–63), women (64), and postbypass patients (65). Cardiac rehabilitation programs also appear to be cost effective, implying that there can be significant cost savings based on comparison with other therapies, but further studies need to be undertaken (66–69).

An easier way to explain the benefits of a cardiac conditioning program can be through the use of a table of metabolic equivalents (METs). A metabolic equivalent, or MET is the amount of energy used by an individual at rest, corresponding to the basal metabolic rate which is approximately 1 kcal/min or 3.5 mL O2/kg/min. Remembering that at a given level of conditioning the anginal threshold is not changed but that the
workload that can be done is increased, it is possible to see that patients with cardiac disease will be able to do more work more comfortably as their VO2 for a given heart rate increases. This greater margin of safety between the MVO2 demanded at routine activity and the anginal threshold increases both vocational and recreational activities. A sample of METs is shown in Table 41-5.

### ABNORMAL PHYSIOLOGY

It is important for effective cardiac rehabilitation to be familiar with the alterations of normal cardiac physiology in disease. Cardiac disease states generally fall into two major areas, pump failure and ischemic disease, with a degree of overlap between the two. For example, MI will decrease the ejection fraction of the heart and thus reduce the SV and CO. Likewise, ischemic heart disease will cause the maximum heart rate to be lowered, with a resultant lowering of the MVO2 and VO2max that can be achieved. During ischemic episodes, the myocardium becomes less compliant, and also less contractile, leading to a decrease in SV due to ventricular stiffening. Valvular heart disease decreases maximum CO, either through flow limitation through tight stenotic valves (e.g., aortic or mitral stenosis) or through valvular regurgitation (e.g., aortic or mitral insufficiency). The end result of the valve disease is pump failure with a decrease in MVO2 and VO2max and increased VO2 at any level of submaximal exercise. Finally, congestive heart failure leads to a decreased

### TABLE 41.5 With Sample METS

<table>
<thead>
<tr>
<th>Energy Costs of Activities of Daily Living</th>
<th>METs</th>
<th></th>
<th>METs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sitting at rest</td>
<td>1</td>
<td></td>
<td>3.5–4</td>
</tr>
<tr>
<td>Dressing</td>
<td>2–3</td>
<td></td>
<td>5–6</td>
</tr>
<tr>
<td>Eating</td>
<td>1–2</td>
<td></td>
<td>4–7</td>
</tr>
<tr>
<td>Hygiene (sitting)</td>
<td>1–2</td>
<td></td>
<td>2–6</td>
</tr>
<tr>
<td>Hygiene (standing)</td>
<td>2–3</td>
<td></td>
<td>7–8</td>
</tr>
<tr>
<td>Sexual intercourse</td>
<td>3–5</td>
<td></td>
<td>6–7</td>
</tr>
<tr>
<td>Showering</td>
<td>4–5</td>
<td></td>
<td>3–4</td>
</tr>
<tr>
<td>Tub bathing</td>
<td>2–3</td>
<td></td>
<td>3–5</td>
</tr>
<tr>
<td>Walking 1 mph</td>
<td>1–2</td>
<td></td>
<td>2–4</td>
</tr>
<tr>
<td>Walking 2 mph</td>
<td>2–3</td>
<td></td>
<td>6–7</td>
</tr>
<tr>
<td>Walking 3 mph</td>
<td>3–3.5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Energy Costs of Avocational Activities</th>
<th>METs</th>
<th></th>
<th>METs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Backpacking (45 lb)</td>
<td>6–11</td>
<td></td>
<td>8–9</td>
</tr>
<tr>
<td>Baseball (competitive)</td>
<td>5–6</td>
<td></td>
<td>9–10</td>
</tr>
<tr>
<td>Baseball (noncompetitive)</td>
<td>4–5</td>
<td></td>
<td>10–11</td>
</tr>
<tr>
<td>Basketball (competitive)</td>
<td>7–12</td>
<td></td>
<td>6–7</td>
</tr>
<tr>
<td>Basketball (noncompetitive)</td>
<td>3–9</td>
<td></td>
<td>9–10</td>
</tr>
<tr>
<td>Card playing</td>
<td>1–2</td>
<td></td>
<td>5–9</td>
</tr>
<tr>
<td>Cycling 5 mph</td>
<td>2–3</td>
<td></td>
<td>5–7</td>
</tr>
<tr>
<td>Cycling 8 mph</td>
<td>4–5</td>
<td></td>
<td>7–8</td>
</tr>
<tr>
<td>Cycling 10 mph</td>
<td>5–6</td>
<td></td>
<td>8–9</td>
</tr>
<tr>
<td>Cycling 12 mph</td>
<td>7–8</td>
<td></td>
<td>9–10</td>
</tr>
<tr>
<td>Cycling 13 mph</td>
<td>8–9</td>
<td></td>
<td>1–2</td>
</tr>
<tr>
<td>Karate</td>
<td>8–12</td>
<td></td>
<td>4–9</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Energy Costs of Vocational Activities</th>
<th>METs</th>
<th></th>
<th>METs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assembly line work</td>
<td>3–5</td>
<td></td>
<td>4–5</td>
</tr>
<tr>
<td>Carpenter (light)</td>
<td>4–5</td>
<td></td>
<td>5–6</td>
</tr>
<tr>
<td>Carry 20–44 lb</td>
<td>4–5</td>
<td></td>
<td>6–8</td>
</tr>
<tr>
<td>Carry 45–64 lb</td>
<td>4–5</td>
<td></td>
<td>5–6</td>
</tr>
<tr>
<td>Carry 65–85 lb</td>
<td>5–6</td>
<td></td>
<td>3–4</td>
</tr>
<tr>
<td>Chopping wood</td>
<td>7–8</td>
<td></td>
<td>6–7</td>
</tr>
<tr>
<td>Desk work</td>
<td>1.5–2</td>
<td></td>
<td>7–9</td>
</tr>
<tr>
<td>Digging ditches</td>
<td>7–8</td>
<td></td>
<td>9–12</td>
</tr>
<tr>
<td>Handyman</td>
<td>5–6</td>
<td></td>
<td>5–6</td>
</tr>
<tr>
<td>Janitorial (light)</td>
<td>2–3</td>
<td></td>
<td>1.5–2</td>
</tr>
<tr>
<td>Lift 100 lb</td>
<td>7–10</td>
<td></td>
<td>6–7</td>
</tr>
</tbody>
</table>

CO with low SV, associated with a lower VO_{2max}, higher resting HRs, and often a greater MVO_{2} at a given VO_{2}.

When present, arrhythmias usually decrease the CO through decreased SV and increased heart rates. This may be due to a loss of atrial contribution (atrial "kick") with supraventricular arrhythmias (e.g., atrial fibrillation of supraventricular tachycardias) or through excessively high HRs without atrial coordination (e.g., ventricular tachycardias, ventricular bigeminy).

Usual surgical treatments of cardiac disease either aim to restore coronary circulation (e.g., bypass and intravascular procedures) or to restore normal anatomy (e.g., valve replacement). Medical therapy for ischemic disease aims to prevent or, if possible, improve coronary circulation, while treatment for heart failure aims to decrease afterload (reducing strain on the ventricle), reduce fluid overload (a fine balance of decreasing preload sufficiently to prevent heart failure while maintaining enough preload for ventricular function), and increase inotropy if possible. Surgical treatment of heart failure includes left ventricular assist devices (LVADs) (approved as destination devices). Medical treatment of arrhythmias with medications has been difficult, with some trials actually increasing the mortality. Implantable defibrillators and pacemakers have been the most efficient and efficacious treatment for these conditions. For intractable cases of all of these conditions, cardiac transplantation is the final possible treatment. However, rehabilitation of transplant patients also requires an understanding of the unique physiology seen in these patients. Even though the transplant may correct many of the pretransplant abnormalities such as congestive heart failure, intractable ischemia, or arrhythmia, posttransplant patients have a persistently high HR, and a limited ability to increase SV can limit exercise response. The details of a program of rehabilitation of all of these syndromes are discussed later in the chapter. The effects on the standard physiologic parameters in response to exercise, as compared to normal individuals, are seen in Table 41-6.

**CARDIAC REHABILITATION**

Cardiac rehabilitation programs are generally divided into primary prevention, which includes risk factor modification and education before a cardiac event, and secondary prevention, which is cardiac rehabilitation with exercise and risk factor modification after the establishment of cardiac disease, including MI and other conditions.

Primary prevention programs focus on the reduction of cardiac risk factors. Often this is education of patients at risk for cardiac disease and also for community-based cardiac disease prevention. The education of patients and individuals at risk for cardiac disease can have a profound effect on the rate of cardiac disease (70,71). An increase in physical activity has been shown to decrease obesity (72), lower SBP (73), and help modify lipid profiles. For behavior modification primary prevention, education about risk factor modification should begin in childhood in order to establish healthy behavior patterns to be maintained throughout life. Programs can be started in schools and parental involvement is also appropriate (74). Primary prevention can also include the use of medications for prevention of cardiovascular disease complications, and includes treatment of hypertension, lowering lipids, and antiplatelet agents. These are cost-effective approaches and large scale analyses in England and the United States that have demonstrated the cost benefits from decreased mortality and morbidity on a population-based scale, in addition to the individual benefits (75,76).

Secondary risk factor modification programs, which are the more common programs for individuals practicing cardiac rehabilitation, occur after an initial cardiac event. Secondary risk factor modification programs include all the features of the primary prevention programs with the addition of disease-specific education and formal exercise programs. The benefit in the prevention of a second cardiac event and in the reduction of mortality after MI with the initiation of risk factor modification programs has been known for over 20 years (77) and has been reemphasized in the latest guidelines (69). The lowering of blood cholesterol has been shown to be advantageous in numerous studies since the 1960s, including the Oslo Study (78), the Western Electric Study (79), the MRFFIT Study (80), Helsinki Heart Study (81), National Heart, Lung and Blood Institute Type II Study (82), and others. Cessation of cigarette smoking is important, as the risk of heart disease returns to that of nonsmokers after 2 years of nonsmoking (83,84).

### TABLE 41.6 Abnormal Physiology in Response to Exercise (as Compared to Normal Individuals)

<table>
<thead>
<tr>
<th>Condition</th>
<th>Aerobic Capacity (VO_{2max})</th>
<th>Cardiac Output</th>
<th>Heart Rate</th>
<th>Stroke Volume</th>
<th>Myocardial Oxygen Capacity (MVO_{2})</th>
<th>Peripheral Resistance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Myocardial infarction</td>
<td>↓</td>
<td>↓</td>
<td>↓</td>
<td>↓ or ×</td>
<td>↓ or ×</td>
<td>↑ or ×</td>
</tr>
<tr>
<td>Ischemic heart disease</td>
<td>↓</td>
<td>× or ↓</td>
<td>↑ or ×</td>
<td>↓ or ×</td>
<td>↓</td>
<td>↓ or ×</td>
</tr>
<tr>
<td>Valvular heart disease</td>
<td>↓</td>
<td>↓</td>
<td>↑ or ×</td>
<td>↓</td>
<td>↓</td>
<td>↑ or ×</td>
</tr>
<tr>
<td>Congestive heart failure</td>
<td>↓</td>
<td>↓</td>
<td>↑ or ×</td>
<td>↓</td>
<td>↓</td>
<td>↑</td>
</tr>
<tr>
<td>Arrhythmias</td>
<td>↓ or ×</td>
<td>↓</td>
<td>↑ or ×</td>
<td>↓ or ×</td>
<td>↓ or ×</td>
<td>↓</td>
</tr>
<tr>
<td>Cardiac transplant</td>
<td>↓</td>
<td>↓</td>
<td>↑ at submax</td>
<td>↓ at max</td>
<td>↓</td>
<td>↑ or ×</td>
</tr>
</tbody>
</table>

\[ x, \text{ unchanged; } \downarrow, \text{ decreased; } \uparrow, \text{ increased.} \]
The control of other risk factors, such as hypertension and diabetes can also be affected by the secondary program (69).

**CARDIAC REHABILITATION OF THE POST-MI PATIENT**

The actual program of cardiac rehabilitation following MI usually follows a modification of the classical model of cardiac rehabilitation as first described by Wenger et al. in 1971 (69,85). Cardiac rehabilitation is now viewed as being divided into three stages or phases. The first phase is the acute phase, which is the in hospital period immediately following the MI, leading up to discharge. This includes early mobilization. The second phase is the rehabilitation phase, which is initiated after healing is completed and is characterized by intense education and aerobic conditioning to achieve the desired results of exercise. Phase III is the final phase, which is devoted to the maintenance of the aerobic conditioning gains made in phase II, through a program of regular exercise. Risk factor modifications are taught and reemphasized throughout all phases.

**Acute Phase (Phase I)**

In the classical model of cardiac rehabilitation, early mobilization was a revolutionary approach to postinfarct patients in an era when 6 weeks of bed rest was considered to be the standard of care. The basics of the early mobilization program are outlined in Table 41-7. The original program was designed to get individuals from bed rest to climbing two flights of stairs in 14 days, and modern programs achieve this goal in 3 to 5 days (69). The rapid increase in mobilization now reflects both the current drive to decrease the length of stay and the recognition of the benefits of rapid mobilization. With respect to the steps of the old program, generally patients are encouraged to be sitting out of bed and in a chair as soon as medically stable, usually by day 1 to 2 (steps 1 to 5). By day 2 to 3, short distance ambulation can be initiated, and bathroom privileges are full (steps 6 to 9). Around day 3, the patient is introduced to the home exercise program, and stairs and increased duration of ambulation are encouraged (steps 10 to 13). After successful completion of a low level ETT for risk stratification on day 4 to 5, the patient completes learning the home program and is discharged (step 14). The educational program relating to risk factor modification should be introduced at this time, especially as many patients are ready to listen to advice in their acute hospitalization. A further modification of the classical program is that now most patients are rapidly evaluated for revascularization procedures, and often will start their rehabilitation after a catheterization or a bypass surgery. With or without revascularization, the acute mobilization should be done with cardiac monitoring (69) and under the supervision of a trained physical or occupational therapist or nurse (69). The post-MI HR rise with activity should be kept to within 20 bpm of baseline and the systolic BP rise within 20 mm Hg of baseline. Any decrease of systolic BP of 10 mm Hg or more should be considered worrisome and the exercise halted (86). The major goal of the phase I program is to condition the patient to perform activities up to four METs, which is within the range of most daily activities at home after discharge (86).

**Inpatient Rehabilitation Phase (Phase IB)**

In order to distinguish between patients who have a rapid recovery after their cardiac event and after revascularization, it has been proposed that admission to either an acute or a subacute rehabilitation facility prior to discharge home be considered phase IB rehabilitation (87). This is now more commonly seen as patients with cardiac disease often are of advanced age or are suffering with comorbidities that make mobilization more difficult. It is in this stage that many rehabilitation specialists will care for these patients. The guidelines for exercise are often the same as they are for the strict phase I patients, but the period of recovery is longer.

**Table 41.7: Wenger Protocol**

<table>
<thead>
<tr>
<th>Step</th>
<th>Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Passive range of motion (ROM), ankle pumps, introduction to the program, self-feeding</td>
</tr>
<tr>
<td>2</td>
<td>As above, also dangle at side of bed</td>
</tr>
<tr>
<td>3</td>
<td>Active assisted ROM, sitting upright in a chair, light recreation, and use of bedside commode</td>
</tr>
<tr>
<td>4</td>
<td>Increased sitting time, light activities with minimal resistance, patient education</td>
</tr>
<tr>
<td>5</td>
<td>Light activities with moderate resistance, unlimited sitting, seated ADL activities</td>
</tr>
<tr>
<td>6</td>
<td>Increased resistance, walking to bathroom, standing ADL, up to 1 h long group meetings</td>
</tr>
<tr>
<td>7</td>
<td>Walking up to 100 ft, standing, warm-up exercises</td>
</tr>
<tr>
<td>8</td>
<td>Increased walking, walk down stairs (not up), continued education</td>
</tr>
<tr>
<td>9</td>
<td>Increased exercise program, review energy conservation, and pacing techniques</td>
</tr>
<tr>
<td>10</td>
<td>Increase exercises with light weights and ambulation, begin education on home exercise program</td>
</tr>
<tr>
<td>11</td>
<td>Increased duration of activities</td>
</tr>
<tr>
<td>12</td>
<td>Walk down two flights of stairs, continue to increase resistance in exercises</td>
</tr>
<tr>
<td>13</td>
<td>Continue activities, education and home exercise program teaching</td>
</tr>
<tr>
<td>14</td>
<td>Walk up and down two flights of stairs, complete instruction in home exercise program and in energy conservation and pacing techniques</td>
</tr>
</tbody>
</table>
Exercise intensity is generally limited to a target heart rate which is known to be safe. The target HR is determined either during a low level ETT or from known limitations from the revascularization performed prior to discharge to the rehabilitation setting. The level of exercise is usually done at a level of approximately 70% maximum heart rate or a MET level of 5 (69,88). For a person more than 40 years old this generally represents a maximum heart rate of 130 bpm or five METs, and for an individual less than 40 years old, 140 bpm or seven METs (50,89,90). A Borg rating of perceived exertion scale of 7 (modified scale) or 15 (old scale) can also be used to determine the maximum tolerated exercise. The Borg scale and Modified Borg scale are shown in Table 41-8. For patients at high risk, more closely monitored programs can be designed. A guideline for determining high risk during cardiac rehabilitation and the need for monitoring are included in Table 41-9.

Training Phase (Phase II)
The training phase of the cardiac rehabilitation program is started after a symptom limited full level ETT or completion of a revascularization procedure and return to full activities. The maximum heart rate from the cardiac exercise test can be used to determine the maximum exertion during aerobic training. For patients who are in the low risk group, a program designed to achieve 85% of the maximum heart rate is generally regarded as safe. For individuals who are at greater risk, exercise programs at lower target HRs can be tailored to individual patients based on the results of the ETT and the reason for cessation of exercise. Generally, for patients with life-threatening arrhythmias or chest pain, lower target HRs are chosen. In patients with higher risk, a target HR of 65% to 75% of maximum can be safe and effective in a regular exercise program (50), and target rates as low as 60% can still yield a training benefit. For the patients at higher risk, it is appropriate to monitor individuals at each increase in activity. Some guidelines for activity classification and treatment are included in Table 41-10.

A classic cardiac training program is three sessions per week for approximately 8 weeks. Unfortunately, limitations of facilities, program availability, and cost often may limit access. Even more problematic is that the majority of eligible patients are not referred to cardiac rehabilitation. This causes a potential loss of treatment efficacy and a potential increase in morbidity and mortality for patients. In order to assist in increasing access to cardiac rehabilitation, creative programs have been developed, including at home programs for low-risk post-MI patients, both community-based (91) and home-based programs (92). A key to success in home-based programs is ensuring that patients are able to perform self-monitoring during their exercise program. Guidelines for self-monitoring are outlined in Table 41-11. Just as in the supervised programs, all exercise session should begin with a stretching session, followed by a warm up session, the training exercise, and end with a cool down period. It is important to remember that conditioning benefit is related to the specificity of training, and that the conditioning applies to the specific muscles exercised.

<table>
<thead>
<tr>
<th>Borg Scale</th>
<th>Perceived Exertion</th>
<th>Modified Borg Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>Nothing at all</td>
<td></td>
</tr>
<tr>
<td>0.5</td>
<td>Very, very weak</td>
<td></td>
</tr>
<tr>
<td>1.0</td>
<td>Very weak</td>
<td></td>
</tr>
<tr>
<td>1.5</td>
<td>Strong (light)</td>
<td></td>
</tr>
<tr>
<td>2.0</td>
<td>Weak (light)</td>
<td></td>
</tr>
<tr>
<td>2.5</td>
<td>Strong</td>
<td></td>
</tr>
<tr>
<td>3.0</td>
<td>Moderate</td>
<td></td>
</tr>
<tr>
<td>3.5</td>
<td>Very, very light</td>
<td></td>
</tr>
<tr>
<td>4.0</td>
<td>Somewhat strong</td>
<td></td>
</tr>
<tr>
<td>4.5</td>
<td>Strong</td>
<td></td>
</tr>
<tr>
<td>5.0</td>
<td>Strong (heavy effort)</td>
<td></td>
</tr>
<tr>
<td>5.5</td>
<td>Fairly light</td>
<td></td>
</tr>
<tr>
<td>6.0</td>
<td>Somewhat hard</td>
<td></td>
</tr>
<tr>
<td>6.5</td>
<td>Hard</td>
<td></td>
</tr>
<tr>
<td>7.0</td>
<td>Very strong</td>
<td></td>
</tr>
<tr>
<td>7.5</td>
<td>Strong</td>
<td></td>
</tr>
<tr>
<td>8.0</td>
<td>Very light</td>
<td></td>
</tr>
<tr>
<td>8.5</td>
<td>Somewhat hard</td>
<td></td>
</tr>
<tr>
<td>9.0</td>
<td>Hard</td>
<td></td>
</tr>
<tr>
<td>9.5</td>
<td>Very, very strong</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Maximal</td>
<td></td>
</tr>
<tr>
<td>&gt;10</td>
<td>Maximal</td>
<td></td>
</tr>
</tbody>
</table>

TABLE 41.8 Borg Scale

<table>
<thead>
<tr>
<th>Ischemic risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>Postoperative angina</td>
</tr>
<tr>
<td>LV ejection fraction &lt;35%</td>
</tr>
<tr>
<td>NYHA grade III or IV CHF</td>
</tr>
<tr>
<td>Ventricular tachycardia of fibrillation in the postoperative period</td>
</tr>
<tr>
<td>SBP drop of ten points or more with exercise</td>
</tr>
<tr>
<td>Excessive ventricular ectopy with exercise</td>
</tr>
<tr>
<td>Incapable of self-monitoring</td>
</tr>
<tr>
<td>Myocardial ischemia with exercise</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Arrhythmic risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acute infarction within 6 wk</td>
</tr>
<tr>
<td>Active ischemia by angina or exercise testing</td>
</tr>
<tr>
<td>Significant left ventricular dysfunction (LVEF &lt;30%)</td>
</tr>
<tr>
<td>History of sustained ventricular tachycardia</td>
</tr>
<tr>
<td>History of sustained life-threatening supraventricular arrhythmia</td>
</tr>
<tr>
<td>History of sudden death, not yet stabilized on medical therapy</td>
</tr>
<tr>
<td>Initial therapy of patients with automatic implantable cardioverter defibrillator (AICD)</td>
</tr>
<tr>
<td>Initial therapy of a patient with a rate adaptive cardiac pacemaker</td>
</tr>
</tbody>
</table>

TABLE 41.9 Patients at High Risk During Cardiac Rehabilitation
## TABLE 41.10 Activity Classification

<table>
<thead>
<tr>
<th>Activity Class</th>
<th>Clinical Characteristics</th>
<th>Activity Guidelines</th>
<th>ECG and BP Monitoring</th>
<th>Supervision Required</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Class A</strong></td>
<td>Healthy individuals</td>
<td>No restrictions</td>
<td>None needed</td>
<td>None</td>
</tr>
<tr>
<td></td>
<td>&lt;40 y old</td>
<td>No symptoms of heart disease</td>
<td>No clinical Heart failure</td>
<td>Individualized with exercise prescription by qualified personnel or restricted to walking</td>
</tr>
<tr>
<td></td>
<td>No known heart disease</td>
<td>No symptoms of heart disease</td>
<td>No ischemia or angina at rest or on ETT at ≤6 METs</td>
<td>Continuous during rehabilitation sessions until safety is established, usually 6–12 sessions</td>
</tr>
<tr>
<td><strong>Class B</strong></td>
<td>Known stable heart disease</td>
<td>Exercise capacity &gt;6 METs</td>
<td>No clinical Heart failure</td>
<td>Individualized with exercise prescription by qualified personnel</td>
</tr>
<tr>
<td></td>
<td>Low risk from cardiac exercise</td>
<td>NYHA class 1 or 2</td>
<td>No ischemia or angina at rest or on ETT at ≤6 METs</td>
<td>Individualized with exercise prescription by qualified personnel and supervised by nonmedical personnel trained in CPR</td>
</tr>
<tr>
<td></td>
<td>CAD</td>
<td>Exercise capacity &gt;6 METs</td>
<td>No high-grade ventricular ectopy</td>
<td>Individualized with exercise prescription by qualified personnel</td>
</tr>
<tr>
<td></td>
<td>Valvular heart disease</td>
<td>Appropriate rise in BP with exercise</td>
<td>Ability to self-monitor activity</td>
<td>Individualized with exercise prescription by qualified personnel</td>
</tr>
<tr>
<td></td>
<td>Congenital heart disease</td>
<td>No ischemia or angina at rest or on ETT at ≤6 METs</td>
<td>No high-grade ventricular ectopy</td>
<td>Individualized with exercise prescription by qualified personnel</td>
</tr>
<tr>
<td><strong>Class C</strong></td>
<td>Cardiomyopathy</td>
<td>No clinical Heart failure</td>
<td>Individualized with exercise prescription by qualified personnel and supervised by nonmedical personnel trained in CPR</td>
<td>Medical supervision during prescriptive sessions Nonmedical supervision for other exercise prescriptions</td>
</tr>
<tr>
<td></td>
<td>ETT abnormalities not directly related to ischemia</td>
<td>Appropriate rise in BP with exercise</td>
<td>Individualized with exercise prescription by qualified personnel</td>
<td>Medical supervision during prescriptive sessions Nonmedical supervision for other exercise prescriptions</td>
</tr>
<tr>
<td></td>
<td>NYHA class 1 or 2</td>
<td>Appropriate rise in BP with exercise</td>
<td>Individualized with exercise prescription by qualified personnel</td>
<td>Medical supervision during prescriptive sessions Nonmedical supervision for other exercise prescriptions</td>
</tr>
<tr>
<td></td>
<td>Previous V-Fib or sudden death not related to acute ischemia or cardiac procedure</td>
<td>Appropriate rise in BP with exercise</td>
<td>Individualized with exercise prescription by qualified personnel</td>
<td>Medical supervision during prescriptive sessions Nonmedical supervision for other exercise prescriptions</td>
</tr>
<tr>
<td></td>
<td>High grade ventricular arrhythmias that are uncontrolled at mild to moderate work intensities</td>
<td>Appropriate rise in BP with exercise</td>
<td>Individualized with exercise prescription by qualified personnel</td>
<td>Medical supervision during prescriptive sessions Nonmedical supervision for other exercise prescriptions</td>
</tr>
<tr>
<td></td>
<td>Left main disease or three vessel disease EF &lt;30%</td>
<td>Appropriate rise in BP with exercise</td>
<td>Individualized with exercise prescription by qualified personnel</td>
<td>Medical supervision during prescriptive sessions Nonmedical supervision for other exercise prescriptions</td>
</tr>
</tbody>
</table>
### TABLE 41.10  Activity Classification (Continued)

<table>
<thead>
<tr>
<th>Activity Class</th>
<th>Clinical Characteristics</th>
<th>Activity Guidelines</th>
<th>ECG and BP Monitoring</th>
<th>Supervision Required</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class E</td>
<td>Unstable disease which restricts activity</td>
<td>No activity is recommended for conditioning purposes</td>
<td>No conditioning program</td>
<td>No conditioning program</td>
</tr>
<tr>
<td></td>
<td>Unstable ischemia</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Uncontrolled heart failure</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Uncontrolled arrhythmia</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Severe symptomatic aortic stenosis</td>
<td>Daily activities as prescribed</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Other conditions that can be aggravated by exercise</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


### TABLE 41.11  Self-Monitoring Guidelines

<table>
<thead>
<tr>
<th>Patient Guidelines</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wear proper clothing</td>
<td>Good shoes, loose-fitting garments, appropriate to the ambient temperature</td>
</tr>
<tr>
<td>Follow pacing guidelines</td>
<td>Follow exertional guidelines established by rehabilitation team, perceived exertional guidelines</td>
</tr>
<tr>
<td>Follow exercise guidelines</td>
<td>5–10 min low intensity warm-up, 20–30 min exercise at full intensity, 5–10 min low intensity cool down</td>
</tr>
<tr>
<td>Stop exercising for adverse symptoms</td>
<td>Cardiac symptoms, chest pain, shortness of breath, lightheadedness, general symptoms, joint pain, faintness with exercise</td>
</tr>
<tr>
<td>No exercise while ill</td>
<td>Wait for 2 d after illness has passed, avoid extreme cold, wear warm clothing, use a face mask</td>
</tr>
<tr>
<td>No exercise in environmental extremes</td>
<td>Exercise indoors in winter, avoid extreme heat and humidity, lower pace, exercise in air-conditioned environment</td>
</tr>
<tr>
<td>No exercise after eating</td>
<td>Wait 2 h after meals</td>
</tr>
</tbody>
</table>

### Maintenance Phase (Phase III)

Despite usually receiving the least attention, the maintenance phase of a cardiac conditioning program is the most important part of the program. If the patient stops exercising, the benefits gained from phase II can be lost in a few weeks. From the beginning of the cardiac rehabilitation program the importance of an ongoing exercise program needs to be emphasized throughout. In order to facilitate compliance with maintenance exercise, the program needs to integrate the actual exercises into the patient’s lifestyle. There also needs to be an emphasis on secondary prevention measures and how to integrate these changes into the patient’s lifestyle. For moderate level exercises, patients should be told to perform their ongoing exercises at the target heart rate for at least 30 minutes three times a week. For low-level exercise, patients need to perform exercise five times a week. ECG monitoring is not necessary during the maintenance phase.

### CARDIAC REHABILITATION PROGRAMS IN SPECIAL CONDITIONS

Due to advances in the treatment of severe cardiac disease, many more patients are being referred to cardiac rehabilitation programs. Including patients with ischemic heart disease, revascularization, heart failure, valvular heart disease, and life-threatening arrhythmias, there are now patients with LVADs and heart transplantation entering rehabilitation programs. With ongoing advances in surgical techniques, it is likely that more patients with new treatments will be referred for cardiac transplantation. In this section, we will deal briefly with all of these separate groups of patients.
Cardiac rehabilitation is appropriate for patients with a stable anginal syndrome. The goal of rehabilitation in angina is to use the training effectively to improve the efficiency of exercise performance below the anginal threshold (Fig. 41-8). The benefit from exercise in angina is derived from the combination of the fact that the actual MVO$_2$ (and thus the maximum HR) at which angina occurs will not change with conditioning, while the work done to achieve the anginal heart rate will increase substantially. Additionally, with exercise there is a possible benefit of reduction of atherosclerotic lesions and increased cardiac collateralization that is cardioprotective and symptom reducing (93–95). Before the initiation of the program of cardiac conditioning, a full level ETT should be done in order to determine the target and maximum HRs and to rule out the potential of life-threatening events. Since anginal patients have suffered no actual infarction, the rehabilitation program can begin at phase II, and the primary and/or secondary prevention program is very important. The primary goal of rehabilitation in this group of patients is aimed at increasing the work capacity. The increase in work capacity achieved will often lead to an increase in functional capacity that may significantly decrease the disability caused by their recurrent chest pain.

**Cardiac Rehabilitation After Revascularization Procedures**

**Post–Coronary Artery Bypass Grafting**

There are a number of mechanisms to account for the benefits of cardiac rehabilitation in the postbypass patient which are listed in Table 41-12 (14,96). Many patients who undergo CABG or percutaneous revascularization procedures often have not had a recent MI and make excellent candidates for cardiac rehabilitation (14). Additionally, some patients have incomplete revascularization and can still be subject to ischemia, allowing them to have the same benefits from exercise seen in patients with angina. Among postrevascularization patients, patients with low ejection fractions and congestive heart failure need to be closely monitored. For these patients, the cardiac rehabilitation program has to be individualized to each patient’s needs. A symptom-limited cardiac stress test can be useful in determining the level of exercise that a patient will tolerate, especially if the revascularized patient has a non-diagnostic thallium or EKG stress test. For patients who have had CABG, the exercise test can be safely performed at 3 to 4 weeks after surgery (97–99). The purpose of the exercise test is to determine the maximal functional capacity, maximal HR, exercise blood pressure response, exercise-induced arrhythmias, and anginal threshold. The initiation of a cardiac rehabilitation program also allows for the initiation of the education program to help modify risk factors. As in post-MI programs, both supervised and unsupervised home programs have been shown to be of benefit in preventing recurrent cardiac disease.

Cardiac rehabilitation after CABG can be thought of as being similar to the post-MI rehabilitation program. Phase I is the immediate postoperative period, and this is followed by a phase II training program and then phase III maintenance. The in-hospital first stage is usually in the first week or so postoperatively, as patients are usually sent home within a week. The phase I program has three stages: (a) intensive mobilization in the immediate postoperative period; (b) progressive ambulation and daily exercises; (c) discharge planning and exercise prescription for the maintenance stage (96).

Unless the patient has an unstable postoperative course, or severe heart failure, intensive mobilization begins in the intensive care unit on postoperative day 1. The program starts with sitting upright, active leg exercises, and mobilization out of bed. Early intervention has several benefits including decreasing the deleterious effects of bed rest (DVT, PE, pulmonary complications and cardiac deconditioning). The program then rapidly progresses with supervised ambulation for distances of 150 to 200 ft, advancing to most
patients beginning independent ambulation by the third day. In the last few days prior to discharge, the patient and physicians develop a program that can be self-monitored at home and allows for gradual progression to previous levels of activity.

The phase II program for a post-CABG patient is usually conducted at home or as an outpatient. High risk patients or those who need other intensive interventions may require an inpatient phase Ib rehabilitation program. This is especially true for patients who may have had a neurological or other complication from their procedure. The home and outpatient rehabilitation patients are regularly supervised by their physicians. Patients can be stratified according to risk by the criteria shown in Table 41-9. Patients usually are stratified to low, moderate, or high intensity programs depending on their level of capacity. A low-intensity program is a progressive walking program with 2 to 4 MET energy expenditures and a target HR of 65% to 75% of maximum HR. Moderate-intensity programs are usually a progressive walk to walk-jog program from 3 to 6.5 METs, with target HRs of 70% to 80% of maximum HR. High-intensity programs progress from walk-jog state to jogging from 5 to 8.5 METs with a target HR of 75% to 85% maximum HR. For patients on β-blockade, the target HR is usually set at 20 beats per minute above the resting HR or can be an HR determined through an ETT or exercise performed at a target MET level (100). If available, a submaximal stress test before discharge is a useful way to evaluate the physiologic response to submaximal effort and assign precautions and target exercise levels for the exercise program. Usually, exercise at a rating of perceived exertion (RPE) of 13 is a level of training that can be safely prescribed in the outpatient setting (37,39). Phase Ib programs and outpatient programs for high risk patients need to be tailored to the specific needs of the patient and designed in cooperation with the patient’s cardiologist.

**Post-Percutaneous Transluminal Coronary Angioplasty**

The rehabilitation of patients after percutaneous transluminal coronary angioplasty (PTCA) is essentially the same as after CABG. PTCA patients tend to be younger and have disease limited to only one or two vessels (101,102). The exercise recommendations should be similar to those for the post-CABG patient population. Cardiac rehabilitation after PTCA is easier than post-CABG since there is no significant postoperative recovery. PTCA patients can begin a conditioning and maintenance program right after their procedure. While monitoring these patients, it is important to remember that some individuals may have incomplete revascularization or may develop restenosis. As in post-CABG rehabilitation, risk factor modification training should be undertaken, and both supervised institution-based and unsupervised home programs can be undertaken. Also, just as in the group of patients after CABG, high-risk patients require closer monitoring and closer physician supervision so that institution-based programs with cardiac monitoring are advised for these individuals. For low-risk post-PTCA patients, the recommendations in Table 41-10 can be followed.

**Cardiac Rehabilitation After Cardiac Transplant Surgery**

As the techniques of orthotopic cardiac transplantation have improved, the numbers of patients with cardiac transplantation have increased. Transplant half-life (the time at which 1/2 of all transplanted patients are alive) is currently 10 years (103), and since heart transplant began on 6/30/07, there have been 76,358 heart transplants and 3,262 heart and lung transplants (104). This means that there is an increased likelihood that physiatrists will see patients who have had cardiac transplants in their practice. Typically, cardiac transplant patients are middle aged, suffer from months of preoperative invalidism and general muscle weakness, and have depression and anxiety. The transplant itself usually resolves the cardiac disability, but a comprehensive approach to the patient is necessary. Due to the complexity of the procedure and the occurrence of vascular and neurological complications, some of these patients also require phase Ib programs and may come to acute or subacute rehabilitation settings.

Because of cardiac deinnervation and immunosuppressive medications, the physiology of the posttransplant patient is somewhat different than that of the normal cardiac patient. Transplantation causes loss of cardiac innervation, with loss of both sympathetic and vagal connections to central regulation. Loss of vagal inhibition to the SA node causes the resting HR of the denervated heart to usually be near 100 bpm (105). Because there is a small SV with the resting tachycardia, the CO can be increased via the Frank-Starling mechanism with a mild increase in SV as a response to increased ventricular filling in response to light exercise. With increased exercise, circulating catecholamines induce chronotropic and inotropic responses to increase the CO (106,107). Additionally, because of loss of direct sympathetic innervation of the new heart, there is a blunted heart rate response to an incremental exercise test, with peak HRs 20% to 25% lower than in matched controls. Resting hypertension, due to the renal effects of calcineurin inhibitors (e.g., cyclosporine, tacrolimus) and prednisone, is common (108). Diastolic dysfunction is another abnormality seen after cardiac transplantation and is due to increased myocardial stiffness. There are three possible causes of the diastolic dysfunction after transplant: (a) myocardial ischemia from accelerated atherosclerosis, (b) side effect of immunosuppressive drugs, and (c) prolonged ischemic time of the donor heart (109). Finally, because transplant recipients often have a 10% to 50% loss of lean body mass from the lack of activity and high-dose steroids in the perioperative period, their maximum work output and maximum oxygen uptake are reduced to about two thirds of the age-matched population (110). This results in higher than normal perceived exertion, minute ventilation, and ventilatory equivalent for oxygen at submaximal exercise levels, while oxygen uptake is preserved, and an earlier onset of anaerobic metabolism. At maximum effort, transplant patients demonstrate lower work capacity, CO, HR, SBP, and oxygen...
uptake, while resting HR and SBP are higher than in normal individuals. Finally, both resting and exertional diastolic blood pressures are higher after cardiac transplantation than in normal individuals.

Exercise testing can be performed after transplantation, but it is important to note that the donor denervated heart cannot demonstrate ischemia through anginal pain, and that close monitoring needs to be done for dyspnea, faintness, and ECG changes. In long-standing transplant recipients, accelerated atherosclerosis may develop and lead to cardiac ischemia.

The cardiac training regimen in transplant patients must address their overall conditioning and their cardiac function. Walking, jogging, cycling, and swimming are common exercises used in the program for transplant patients. In the initial postoperative period, sitting upright, lower extremity exercises, and mobilization from the bed are encouraged. The initial postoperative exercise program is similar to post-CABG patients. At the time of discharge, patients should have learned self-monitoring and are encouraged to increase ambulation as tolerated. The outpatient program then consists of progressively increasing distances for ambulation, with the pace designed to be at a level of 60% to 70% of peak effort for 30 to 60 minutes three to five times weekly (111). RPE targets, using the Borg scale, should be maintained at 13 to 14, with the level of activity increasing incrementally to stay at this level.

In addition to secondary prevention education for post-transplant patients, the rehabilitation program should also include addressing the complicated medical regimen and the vocational, debility, and psychological needs of the patient. The outcomes of rehabilitation in the cardiac transplant population have been generally favorable. Patients usually achieve increased work output and improved exercise tolerance (110), even resuming competitive athletics.

**Cardiomyopathy**

With improved cardiac care, the number of patients with left ventricular ejection fraction (LVEF) of less than 30% has increased, being one of the fastest growing subsets of the cardiac rehabilitation population (112). Due to poor LV function, these patients have increased complications compared to CABG or post-MI population, with a higher risk of sudden death, depression, and chronic cardiac disability (113,114). Due to decreased CO, heart failure patients demonstrate inconsistent responses to exercise (115). One of the earliest presenting findings and causes of the functional impairment seen with CHF is due to impaired exercise capacity. The normal physiological response to exercise is often absent, and there can be a decline in ejection fraction, a decrease in SV, with resultant exertional hypotension, and syncope. In the most severe cases, CO may not increase sufficiently to generate a dynamic exercise response at all. However, the hemodynamic alterations which are seen with exercise do not always correlate with the overall exercise capacity (116), and patients can have remarkably sustained performance with a very low LVEF. Still, low endurance and fatigue are frequently seen and prolonged postexercise fatigue is common, often lasting for hours to days after achieving a high aerobic workload (117). These limitations to exercise are often exacerbated in patients with CHF when atrial fibrillation, fluid overload, or medication noncompliance is also present. Despite these limitations, there is documented benefit from exercise in this patient population (118–120). A gradual program of increasing the HR above resting can be safe and increases the oxygen extraction efficiency. Exercise duration can increase by as much as 18% to 34% (121,122), and peak oxygen uptake can increase by 18% to 25% (121,123). The benefits of the cardiac rehabilitation programs in CHF are similar to those for angina, with lower HRs at rest and during submaximal exercise, improved anaerobic thresholds, and increased VO2max (124). Since many of these patients present initially with a very low exercise capacity, even a small improvement in VO2 can mean living independently for a patient with heart failure.

Since this population of patients is at higher risk than most cardiac rehabilitation patients, it is essential to have a graded ETT before starting. Other useful baseline values can be a measurement of LVEF by multiple-gated acquisition scanning (MUGA) or echocardiography during exercise. As with cardiac rehabilitation for ischemic disease, unstable angina, decompensated CHF, and unstable arrhythmias are contraindications to cardiac rehabilitation. Due to the poor adaptation to exercise seen in CHF patients, long warm-ups and cooldown periods are required with exercise being performed with a limited workload. It is important to remember that dynamic exercise is preferable to isometric exercise, with a target HR of 10 bpm below any significant endpoint found with cardiopulmonary exercise testing, such as exertional hypotension, marked dyspnea, or sustained arrhythmia (125). Because of possible increased diastolic pressure with subsequent increase in cardiac afterload, isometric exercise should be avoided where possible and limited to 2-minute intervals for those exercises performed. It is best for cardiac exercise to be done under supervision with cardiac monitoring initially. Patients with severe left ventricular dysfunction will need telemetry during warm-up, exercise, and cooldown. Once the patient is able to self-monitor and the clinicians are confident that there are no complications during the routine cardiac exercise, patients can start a self-monitored program. Other techniques to monitor CHF patients include closely following body weight (to observe for fluid accumulation) and blood pressure and HR responses to exercise.

For more severe heart failure, including patients who are awaiting cardiac transplant due to their heart failure, management of the patient often includes pharmacologic inotropic support, and if that fails, left ventricular mechanical support. Exercise can be done on intravenous (IV) inotropes with the same caveats as for normal CHF patients (120). More recently, LVADs have become more common and are now approved
for use as a bridge for cardiac transplant or as a destination device (126–128). Any of the several models of LVAD and IV inotropes sustain cardiac function and allow mobilization of these patients. Some in-hospital acute rehabilitation units can accommodate patients who have either inotropes or LVAD in a comprehensive inpatient CR program. This facilitates earlier successful transplantation through more rapid mobilization. Centers that wish to be able to treat LVAD patients with phase Ib cardiac rehabilitation require a trained staff, close cooperation with the LVAD team, and familiarity with the devices that are used locally out of the several models now available. The most important issue for the treating rehabilitation team is to have close contact with the cardiac surgical team and to include LVAD-specific training for the patient and family. CO is typically well sustained with the device, and the patients can have good exercise tolerance. The peak capacity for exercise is often limited by the peak flow of the device, and the treating team needs to keep the patient within parameters that do not exceed the peak flow capacity of the device. Finally, drive lines need to be protected, and precautions for anticoagulation (if needed) are all part of the standard precautions.

Valvular Heart Disease
For patients with valvular heart disease, the major issues are similar to patients with CHF or the pretransplant population. The management of the valvular heart disease patient with CHF is essentially the same as described for the CHF patient. After surgical correction of valvular heart disease, the patient improves in cardiac fitness as measured by improved oxygen consumption (129). Training can increase the physical work capacity by 60%, decrease RPE, and decrease the RPP by 15% (130). An issue that is present in many patients after valve replacement is the presence of anticoagulation postoperatively for patients with mechanical valves. When patients are on anticoagulant therapy, the exercise program has to avoid high-impact exercises to avoid hemarthrosis and bruising and include a component of education regarding injury avoidance (131). The overall training program is similar to that discussed for the post-CABG patient.

Cardiac Arrhythmias
The risk of death from cardiac arrhythmia during rehabilitation exercises is very low. In a recent study of 25,420 patients, with 743,471 hours of exercise training, there was a risk of one myocardial arrest event for every 1.3 million therapy hours and one cardiac event for every 49,565 patient hours of exercise (132). This is an improvement in outcomes and shows the relative safety of exercise in monitored settings. Table 41-9 identifies those patients at high risk of cardiac arrhythmias who should be exercised in the monitored setting rather than in a self-directed setting at home program. For those patients with life-threatening arrhythmias, the automatic implantable cardiac defibrillator (AICD) has been used increasingly and may improve safety. The modifications for a program of exercise with an AICD are limited and are mostly set to avoid the rate at which the device is set to fire. A prerehabilitation exercise stress test and cardiac precautions with target HR set well below the trigger threshold are sufficient modifications to the exercise program. Patients with AICD have been shown to have a good functional outcome from a rehabilitation program (133). Another important part of the rehabilitation of patients with AICD is support and reassurance about the safety of exercise, and there is a role for a cognitive-behavioral modification program for anxiety about recurrent arrhythmia (134,135).

Cardiac Rehabilitation in the Physically Disabled
As a last note, there are some special considerations for patients with both disability and cardiac disease. Most of the difficulties in cardiac rehabilitation in this population are due to limited mobility that presents difficulty in both testing and exercise training. Patients with disability may also be at higher risk of cardiac disease, and the presence of this disease must be remembered when engaging in a rehabilitation program. Patients with stroke or peripheral vascular disease are at particularly high risk, but any patient with disability may have a cardiovascular comorbidity, and in cases where cardiac disease is overt, cardiac rehabilitation should be provided for these individuals just as it would be for the able-bodied population (136,137). Overall, disabled patients have several issues that place them at greater risk for cardiac disease than their nondisabled counterparts. They are usually more sedentary, which yields a higher total cholesterol and a lower HDL (138–141). Obesity occurs more frequently with disability, and deconditioning is also often seen. Additionally, disabled individuals usually require much higher energy expenditures for mobility, with a resultant need for increased work capacity.

When prescribing cardiac exercise for disabled individuals, the exercise protocols need to be adapted to the individual patient. Patients with lower extremity impairment due to neurological or orthopedic conditions perform upper extremity ergometry and modification of lower extremity exercise equipment will allow them to exercise with their legs. Hemiplegic patients can use adapted bicycle ergometers or airdynes. Since exercise protocols for stroke and other conditions incorporate upper limb exercise, the high MVO₂ requirements for upper extremity exercise should be considered when designing a cardiac rehabilitation program for disabled patients. Table 41-13 shows some of the relative cardiovascular responses to arm and leg exercises. Exercise training for disabled individuals follows the same basic principles of cardiac rehabilitation for nondisabled individuals. The additional concern in disabled patients is that exercise should focus particularly on task-specific activities in order to improve aerobic conditioning and endurance, while seeking to lower MVO₂ needed with each task. The physiatrist is particularly suited to take a leadership role in the area of the design of cardiac rehabilitation programs.
for the disabled since most traditional cardiac rehabilitation programs have limited experience with the needs of physically disabled patients.

**REFERENCES**


<table>
<thead>
<tr>
<th>TABLE 41.13</th>
<th>Relative Cardiovascular Response to Upper (UE) and Lower (LE) Extremity Exercise</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameter</td>
<td>Submaximal Exercise</td>
</tr>
<tr>
<td>Work rate</td>
<td>UE = LE</td>
</tr>
<tr>
<td>Oxygen uptake</td>
<td>UE &gt; LE</td>
</tr>
<tr>
<td>Cardiac output</td>
<td>UE &gt; LE</td>
</tr>
<tr>
<td>Stroke volume</td>
<td>UE &lt; LE</td>
</tr>
<tr>
<td>Heart rate</td>
<td>UE &gt; LE</td>
</tr>
<tr>
<td>MVO₂</td>
<td>UE &gt; LE</td>
</tr>
<tr>
<td>Systolic BP</td>
<td>UE &gt; LE</td>
</tr>
<tr>
<td>Diastolic BP</td>
<td>UE &gt; LE</td>
</tr>
<tr>
<td>Total peripheral resistance</td>
<td>UE &gt; LE</td>
</tr>
</tbody>
</table>


American Heart Association scientific statement from the Council on Clinical Cardiology (Subcommittee on Exercise, Cardiac Rehabilitation, and Prevention) and the Council on Nutrition, Physical Activity, and Metabolism (Subcommittee on Physical Activity), in collaboration with the American association of Cardiovascular and Pulmonary Rehabilitation. Circulation. 2005; 111(3):369–376.


Rehabilitation of the Patient with Respiratory Dysfunction

The beginning of wisdom is to call things by their right names.

Ancient Chinese Proverb

The names that are attached to clinical phenomena determine in part how they are managed. For example, “ventilator associated pneumonia (VAP)” is responsible for over 60,000 excess hospital-related deaths annually. However, no ventilator, but rather the invasive interface, causes pneumonia. If this were acknowledged, there might be a greater emphasis on providing ventilator support by noninvasive interfaces. Likewise, symptomatic respiratory acidosis due to ventilatory pump failure is generally called “respiratory” not “ventilatory” failure. If the name were more appropriate, so might be the treatment, that is, assisted ventilation rather than supplemental oxygen. Many patients we treat must be “habilitated” rather than “rehabilitated.” This brings us to the definition of “pulmonary rehabilitation.”

Pulmonary rehabilitation has been defined as “a multidimensional continuum of services directed to persons with pulmonary disease and their families, usually by an interdisciplinary team of specialists, with the goal of achieving and maintaining the individual’s maximum level of independence and functioning in the community” (1). Interventions can include exercise, respiratory muscle rest and support, education, emotional support, oxygen, airway secretion clearance, promoting compliance with medical care, facilitating return to work, and a more active and emotionally satisfying life. These goals are appropriate for any patients with diminished respiratory reserve whether due to obstructive or intrinsic pulmonary diseases (oxygenation impairment) or neuromuscular weakness (ventilatory impairment). The former are normally eucapnic or hypocapnic, often despite severe hypoxia, and hypercapnia occurs only during episodes of acute respiratory failure (ARF) or with end-stage disease. For the latter, hypercapnia begins during sleep and precedes significant hypoxia or oxyhemoglobin desaturation. The ventilatory dysfunction causes decreased blood oxygenation. Many useful techniques for managing the latter were described, developed, or initially adapted by physiatrists in the United States (2–7), and elsewhere (8).

REHABILITATION OF PATIENTS WITH OBSTRUCTIVE LUNG DISEASE

Thirty-two sources of chronic obstructive pulmonary disease (COPD) prevalence estimation rates from 17 countries were surveyed. Prevalence ranged from 0.23% to 18.3%. In Europe and North America, rates were between 4% and 10% (9). COPD is the second most common noninfectious disease in the world, causing 2.75 million deaths annually with global mortality predicted to double by 2030. It is the fourth largest cause of major activity limitation (10). Thirty percent of COPD patients with forced expiratory volume in 1 second (FEV₁) less than 750 milliliters (mL) and 35% following an acute exacerbation die within 1 year (11) and 50% within 3 years (12).

Patient Evaluation

In addition to the elements of the patient evaluation noted in Table 42-1, any medical, physical, financial, or psychological factors that might interfere with a rehabilitation program need to be addressed (12). Commonly overlooked is the fact that 13% of COPD patients are anemic. This may indicate presence of erythropoietin resistance and be associated with increased serum inflammatory proteins (13).

Various dyspnea assessment surveys can be used to objectify the extent of dyspnea and the effects of rehabilitation (14–17). In addition, the presence of any dyspnea, coughing, wheezing, chest pains, neurologic or psychological disturbances, allergies, previous communicable diseases, injuries, and nutritional imbalance is explored.

Poor nutrition can be characterized by low protein values. Other tests that assess nutritional status include total iron binding capacity, cholesterol, and serum vitamin levels, especially of vitamins A, C, and E. This is especially important because COPD and neuromuscular disease (NMD) patients have a high incidence of vitamin deficiencies. Delayed cutaneous hypersensitivity is a function of cell-mediated immunity and is the immune response most sensitive to nutritional deprivation. Hypophosphatemia, hypomagnesemia, hypocalcemia, and hypokalemia may also cause respiratory muscle weakness, which is reversible following replacement (18). Social, educational, and vocational histories and any relevant environmental factors are explored.
Patient Evaluation

| Family history of pulmonary diseases |
| Symptom progression and impact on function |
| Exacerbation and hospitalization history |
| Appetite, nutritional status, and weight changes |
| Medications, substance abuse |
| Physical examination for hyper-resonant chest, poor breath, and cardiac sounds |
| Hemoglobin/hematocrit, sedimentation rate, C-reactive protein, white cell count |
| Radiographic assessment of (low, flattened) diaphragm, (long, narrow) heart shadow, retrosternal translucency, narrowing of peripheral pulmonary vessels |
| PaCO₂ and PaO₂ and diffusion capacities (decreased in emphysema, normal in bronchitis) |
| Pulmonary function tests for air and mucus trapping |
| Low-maximum midexpiratory flow rates and increased midexpiratory times, normal or increased lung compliance and increased flow work, increased residual volume and total lung capacity |
| Clinical exercise testing |
| A 3-, 6-, or 12-min walk |
| Assessment of the VAT and maximum exercise tolerance for a precise exercise regimen |

The pulmonary function studies of COPD patients demonstrate air trapping, low-maximum midexpiratory flow rates and increased midexpiratory times, normal or increased lung compliance, and increased flow work. Residual volume and total lung capacity are generally increased. Exertional dyspnea tends to occur when the forced expiratory volume in 1 second (FEV₁) is less than 1,500 mL. FEV₁ decreases by 45 to 75 mL/yr for COPD patients (19), a rate up to three times normal. Arterial oxygen tensions may be posturally related, significantly decreased with the patient supine (20), and desaturation may be episodically severe during sleep (21).

Since pulmonary function impairment does not predict the overall functional impairment, clinical exercise testing is done. This measures the functional reserve of all mechanisms taking part in oxygen and carbon dioxide transport and yields information regarding the capacity to perform exercise, the factors that limit exercise, the reasons for exercise-related symptoms, and the diagnosis (22). It permits the clinician to determine whether the primary disability is pulmonary, cardiac, or related to exercise-induced bronchospasm (23). The latter two diagnoses and even the presence of purely restrictive pulmonary syndromes are commonly mistaken for COPD and, therefore, may be mismanaged. Clinical exercise testing can also be useful for documenting patient progress with rehabilitation.

Clinical exercise testing, whether by using a treadmill, stationary bicycle, or upper-extremity ergometer, includes monitoring of: vital signs, electrocardiography, oxygen consumption, carbon dioxide production, the respiratory quotient, the ventilatory equivalent, minute ventilation, and metabolic rate. The respiratory quotient is the ratio of the carbon dioxide produced divided by the oxygen consumed. The ventilatory equivalent is equal to the volume of air breathed for 1 L of oxygen consumed. A metabolic equivalent (MET) is the resting metabolic rate per kilogram of body weight (i.e., 1 MET = 3.5 mL O₂/kg/min). Other useful measures for noninvasively assessing cardiac function include the oxygen pulse, a measure of the mL of oxygen consumed per heart beat (23). A clinical exercise test should advance until oxygen consumption fails to increase, maximum allowable heart rate for age is reached (usually 220—age in years), or electrocardiographic changes, chest pain, severe dyspnea, or intolerable fatigue occurs. A minute ventilation 37.5 times the patient’s FEV₁ can be the goal (24). Arterial blood gases may be normal at rest but are often abnormal during exercise. Oximetry is performed to determine need for supplemental oxygen therapy to maintain SpO₂ greater than 90% during reconditioning exercise or greater than 60 mm Hg long term (25). It has been suggested that supplemental oxygen benefits patients with COPD with moderate-to-severe airflow obstruction and mild hypoxemia at rest by improving exercise tolerance and reducing pulmonary hypertension during exercise (26).

When energy cost studies are not available, maximum exercise tolerance may be estimated from pulmonary function data (27). A 3-, 6-, or 12-minute walk test can also provide useful information. The patient is instructed to gradually increase walking speed and duration on subsequent walking tests. The test is simple and may be performed daily in the hospital or at home (28).

Any motivated COPD patient who has respiratory symptoms that limit activities of daily living (ADL) and who has adequate medical, neuromusculoskeletal, financial, and psychosocial status to permit active participation is a candidate for rehabilitation. Active patients who are still able to walk several blocks but who have noted yearly decreases in exercise tolerance or who have recently begun to require ongoing medical attention for pulmonary symptoms or complications are ideal candidates.

Organization of a Comprehensive Rehabilitation Program

Because, other than perhaps for smoking cessation (29,30), there is no evidence that inpatient programs are more effective than outpatient programs (29–32), the former should be reserved for severely debilitated patients (33), for tracheostomy tube removal or for optimizing the ventilatory aid regimen while initiating other aspects of comprehensive rehabilitation (34). Table 42-2 is a sample therapeutic prescription for an ambulatory, moderately affected COPD patient.

Therapeutic Interventions

Medications

The patient’s medical regimen is optimized prior to reconditioning exercise. Bronchodilators are delivered as aerosolized solutions but are not effective if the patient deposits the medications uselessly on the tongue. COPD patients require training in the use of “spacers” and nebulizers (35). Orally administered β agonists are used when aerosolized medications are ineffective or when metered-dose inhaler cannot be
One half to two thirds of 33 double-blind, randomized, placebo-controlled studies showed significant positive effects of anticholinergics and short-acting β-2 mimetics (especially salbutamol), respectively, on exercise capacity. Early medical attention is important during intercurrent respiratory tract infections (36). Antibiotics, glucocorticoids, and adjustment of bronchodilators and mucolytic agents may be indicated. N-acetylcysteine at 1,200 mg/d was demonstrated in a randomized, double-blind, placebo-controlled study to normalize C-reactive protein levels, lung function, and symptoms during acute exacerbations and possibly in preventing the exacerbation (37).

Counseling and General Medical Care
Dyspnea often causes fear and panic. This may worsen tachypnea while increasing dead-space ventilation, the work of breathing, hyperinflation, and air trapping. Relaxation exercises, such as Jacobson exercises and biofeedback (38,39), yoga, and diaphragmatic and pursed-lip breathing, can be used to decrease tension and anxiety.

COPD patients perceive impairment in quality of life. Depression has been reported in 50% of patients, and there is often severe reduction in social interaction (40). Integrating psychosocial support with multimodal pulmonary rehabilitation optimizes intervention (41). Loss of employment and physical independence may also need to be addressed.

COPD patients tend to overuse medications during periods of respiratory distress and underuse them otherwise. They are counseled on adhering to their medication regimens (42), and on avoiding atmospheric or vocational pollutants and other aggravating factors, such as pollen, aerosols, excessive humidity, stress, and respiratory tract pathogens. Yearly flu vaccinations are recommended; and pneumococcal vaccines are used one time or every 5 to 10 years for high-risk cases.

### TABLE 42.2  A Sample Therapeutic Prescription for a Patient with COPD

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>Chronic obstructive pulmonary disease</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prognosis</td>
<td>Favorable, patient on stable self-medication program</td>
</tr>
</tbody>
</table>
| Goals                 | • Ease sleep disordered breathing and rest inspiratory muscles  
                          • Improve endurance and efficiency  
                          • Optimize oxygen needs and control of secretions  
                          • Increase independence in ambulation and self-care activities  
                          • Reduce anxiety and improve self-esteem through enhanced body awareness |
| Precautions           | • Supplemental oxygen needed during exercise  
                          • Discontinue and notify physician if patient becomes severely dyspneic or develops chest pain with exercise |
| Respiratory therapy   | • Conduct ear oximetry at rest and during exercise to determine portable oxygen flow rate needed to maintain oxygen saturation higher than 90%  
                          • Instruct patient in diaphragmatic and pursed-lip breathing  
                          • Instruct patient and family in postural drainage techniques  
                          • Instruct patient and family in portable oxygen use  
                          • Instruct in use of metered-dose inhaler before exercise  
                          • Instruct in use of nocturnal BiPAP |
| Physical therapy      | • Assess baseline endurance, using 12-min walk test  
                          • Begin incremental exercise program to improve endurance through ambulation and stair climbing. Begin with 5-min sessions, followed by rest periods between sessions. When patient tolerates 20 min of total exercise per day, begin consolidating the sessions. Initial treatments on daily basis during weeks 1 and 2, taper to three times per week over weeks 3 and 4, and then taper to home program with self-monitoring in weeks 5 and 6.  
                          • Review proper body mechanics and coordinate with breathing patterns, using diaphragmatic and pursed-lip breathing when appropriate |
| Occupational therapy  | • Assess upper-extremity mobility, strength, and endurance  
                          • Evaluate basic and advanced self-care activities, and provide adaptive aids to improve independence with dressing, hygiene, bathing, cooking, and other chores  
                          • Train the patient in energy conservation and work simplification techniques  
                          • Evaluate home environment and make recommendations for workspace modifications and equipment to improve safety, efficiency, and independence  
                          • Provide relaxation exercise training with visual imagery techniques |
High-altitude travel may require an additional 0.5 L/min of supplemental oxygen administration for those already requiring supplemental oxygen; otherwise, oxygen therapy generally need not be used for short flights (43). Good hydration should be maintained with ample fluid intake.

**Nutrition**

Significant weight loss occurs in 19% to 71% of COPD patients (18). In a study of 255 stable COPD pulmonary rehabilitation patients, depletion of body weight, fat-free mass, and muscle mass was noted in 40% to 50% of patients with chronic hypoxemia and in normoxemic patients with FEV1 < 35% (44). In one study, 30 of 50 consecutive COPD patients presenting with ARF were significantly undernourished, and impaired nutritional status was more prevalent in those patients requiring mechanical ventilation (74% vs. 43%) (45).

Undernutrition has been associated with an increased susceptibility to infection that is due in part to impaired cell-mediated immunity, reduced secretory immunoglobulin A, depressed pulmonary alveolar macrophage function, and increased colonization and adherence of bacteria in the upper and lower airways. Patients with significant nutritional impairment are more frequently colonized by *Pseudomonas* species (46,47). In addition, malnutrition can adversely affect lung repair, surfactant synthesis, control of ventilation and response to hypoxia, respiratory muscle function and lung mechanics (48), and water homeostasis. It can lead to respiratory muscle atrophy and decreased exercise capacity, cor pulmonale, increased rate of hospitalization for pulmonary-related problems, hypercapnic respiratory failure, and difficulty in weaning from mechanical ventilation (45,49,50). Likewise, inappropriate nutrition, such as increasing carbohydrate intake, can exacerbate hypercapnia.

Short-term refeeding of malnourished patients can improve respiratory muscle endurance and increase respiratory muscle strength in the absence of demonstrable changes in skeletal muscle function (51). Because of bloating due to a low diaphragm, patients are advised to take small mouthfuls of food, eat slowly, and take smaller and more frequent meals. SpO2 can be evaluated while eating. If desaturation occurs, supplemental oxygen is used or increased. For those with hypercapnia, a dietary regimen high in calories derived from fat can decrease hypercapnia. Although short-term refeeding can be beneficial, refeeding programs lasting more than 2 weeks have not consistently resulted in increases in body weight. Growth hormone has not been shown to be useful. Possible beneficial effects of anabolic steroids as adjuncts to nutritional support and exercise have been reported to increase lean body mass and promote weight gain (52–54).

**Breathing Retraining**

Shallow, rapid breathing is commonly seen in anxious, dyspneic patients. This increases dead-space ventilation and airflow through narrowed airways, increasing the flow work of breathing. Patients with chronic airflow obstruction also have an altered pattern of ventilatory muscle recruitment in which the most effective ventilatory pressure is generated by the rib-cage inspiratory muscles rather than by the diaphragm, with significant contribution by primarily expiratory muscles (55). Diaphragmatic breathing and pursed-lip exhalation can help to reverse these tendencies. These techniques are usually initiated in the supine or 15% to 25% head-down position. Diaphragmatic breathing is guided by having the patient place one hand over the abdomen and the other on the thorax just below the clavicle. Breathing deeply through the nose, the abdomen is distended forward as appreciated by movement of the hand on the abdomen. Movement of the rib cage and, thus, the hand on the thorax should be kept to a minimum. Small weights can be placed on the abdomen to provide some resistance training and enhance the patient’s focus. During exhalation, the abdominal muscles and the hand on the abdomen compress the abdominal contents, and exhalation is via pursed lips (56). Classically, a lighted candle is put several feet in front of the patient, and the patient flickers the flame while exhaling. This equalizes pleural and bronchial pressures, preventing collapse of smaller bronchi and decreasing air trapping. Diaphragmatic and pursed-lip breathing decreases the respiratory rate, coordinates the breathing pattern, and can improve blood gases (57). It is used during routine ADL and exercise. It can improve exercise performance by relaxing accessory muscles and improving breathing efficiency.

Air-shifting techniques may be useful to decrease microatelectasis. Air shifting involves taking a deep inspiration that is held with the glottis closed for 5 seconds, during which time the air shifts to lesser-ventilated areas of the lung. Expiration is via pursed lips. This technique may be most beneficial when performed several times per hour.

**Airway Secretion Elimination**

Airway secretion clearance is crucial because exacerbations of COPD are caused by trapping of airway secretions in the peripheral airways. The patient’s cough may be weak or ineffective as a result of increased airway collapse in more central airways, and frequent bouts of coughing are fatiguing. The high expulsive pressures generated during coughing can exacerbate both air trapping and secretion retention. “Huffing,” or frequent short expulsive bursts following a deep breath, is often an effective and more comfortable alternative to coughing. Chest percussion and postural drainage can be useful for patients with chronic bronchitis or others with greater than 30 mL of sputum production per day (58), although caution must be taken to increase oxygen delivery as necessary during treatment. Autogenic drainage involves breathing with low tidal volumes between the functional residual capacity and the residual volume to mobilize secretions in small airways. This is followed by taking increasingly larger tidal volumes and forced expirations to transport mucus to the mouth (59).

Application of positive expiratory pressure (PEP) breathing is based on the theory that mucus in small airways is more effectively mobilized by coughing or forced expirations...
if alveolar pressure and volume behind mucous plugs are increased. PEP is applied by breathing through a face mask or a mouthpiece with an inspiratory tube containing a one-way valve and an expiratory tube having variable expiratory resistance. Expiratory pressures of 10 to 20 cm H2O are maintained throughout expiration. PEP increases functional residual capacity, reducing resistance to airflow in collateral and small airways (60,61). Studies on the clinical benefits of PEP breathing have been inconclusive for both cystic fibrosis (CF) and COPD (62–70).

Flutter breathing is a combination of PEP and oscillation applied at the mouth. The patient expires through a small pipe. A small stainless steel ball rests on the expiratory end of the pipe; it is pushed upward during expiration, producing PEP, and falls downward again, interrupting flow. The mucus-mobilizing effect is thought to be due to widening of the airways because of the increased expiratory pressure and airflow oscillations due to the oscillating ball (71). For this too, however, the results of clinical trials have been conflicting (72–76).

With currently available technology, mechanical vibration or oscillation can be mechanically applied to the thorax or directly to the airway to facilitate airway secretion elimination. Vibration is possible at frequencies up to 170 Hz applied under a soft plastic shell to the thorax and abdomen (the Hayek Oscillator, Breasy Medical Equipment Inc., Stamford, CT). Another device delivers rapid burst airflow at up to 25 Hz under a vest covering the chest and upper abdomen (THAIRapy System, American Biosystems, Inc., St. Paul, MN). The effects of mechanical chest percussion and vibration appear to be frequency dependent (77–79). In most animal studies, frequencies between 10 and 15 Hz appear to best facilitate mucous transport (77,79,80), especially the transport of a thicker mucous layer (81). Warwick and Hansen found long-term increases in forced vital capacity (FVC) and forced expiratory flows for CF patients treated with high-frequency chest-wall compression as compared with manual chest percussion alone (82). Others have reported improvement in pulmonary function and in gas exchange during high-frequency oscillation (83–87), and Sibuya et al. found that chest-wall vibration decreased dyspnea (86). Most studies on COPD and CF patients, however, have failed to demonstrate objective clinical benefits from percussion or vibration on mucous transport (87–90). Side effects of percussion and vibration can include increasing obstruction to airflow (91,92). In an animal model, the application of vibration and percussion was also associated with the development of atelectasis (93). Despite conflicting studies, the THAIRapy Vest has become popular for CF and familial dysautonomia patients, and studies have claimed decreases in hospitalization rates with its daily use (77). Patients with daily airways secretion encumbrance usually feel that its use is beneficial.

The Percussionator (Percussionaire Corp., Sandpoint, ID) can deliver aerosolized medications while providing high-flow percussive minibursts of air directly to the airways at rates of 2.5 to 5 Hz. This intrapulmonary percussive ventilation has been reported to be more effective than chest percussion and postural drainage in the treatment of postoperative atelectasis and secretion mobilization in COPD patients (94,95). The majority of such patients feel that it is helpful (96,97).

Patterson et al. found in a 10-year study that good CF patient compliance with airway secretion mobilization methods was associated with a slower rate of loss of pulmonary function (98). Patient compliance is usually poor, however (99–101). There is greater patient compliance for simple methods that can be used independently. Little has been documented concerning long-term safety and efficacy of any of these methods. Because expensive methods have not been shown to be more effective than simple handheld percussors costing about $400 (Jeannie Rub Percussor, Morfan Inc., Mishawaka, IN [Fig. 42-1] or NeoCussor General Physical Therapy, St. Louis, MO), the latter should be favored for routine clinical use.
Inspiratory Resistive Exercises

Inspiratory resistive exercises, including maximum sustained ventilation, inspiratory resistive loading, and inspiratory threshold loading, can improve the endurance of respiratory muscles (102–104). Typically, patients breathe through these devices for a total of 30 minutes daily for 8 to 10 weeks. The settings of the devices are adjusted to increase difficulty as patients improve and the program advances.

Levine et al. conducted an evaluation of the isocapnic hyperpnea method and determined that no more benefits could be derived from it than could be achieved using periodic intermittent positive pressure breathing treatments, a regimen that was considered equivalent to placebo (105). However, Ries and Moser randomly assigned 18 patients to either a home isocapnic hyperventilation training program or a walking program and found that the former led to improvements in ventilatory muscle endurance and exercise performance and significant improvements in the maximum rate of oxygen consumption (VO2max), whereas walking exercises improved lower limb exercise endurance but not ventilatory muscle endurance (106).

Twenty-one controlled studies of inspiratory resistive loading involving 259 COPD patients reported improvements in inspiratory muscle strength and endurance (107). The mean increase in maximum inspiratory pressure was 19%. However, the subjects using inspiratory resistance training reduce their inspiratory flow rates and lengthened their inspiratory time to reduce the severity of the imposed loads. Thus, “targeted” or threshold inspiratory muscle training is recommended over flow-resistive training to assure adequate intensity of inspiratory muscle activity.

With targeted training, the subject is provided feedback regarding the inspiratory flow rates through the resistor or the inspiratory pressure generated by flow through the resistor; with threshold training, the subject is unable to generate flow through the device until a predetermined pressure is achieved. Six of the nine controlled studies of the use of targeted or threshold resistor devices in COPD reported significantly greater improvements in inspiratory muscle function in the subjects than in the controls (107). In three of the six studies in which it was assessed, exercise tolerance was greater for trained subjects than for controls. In one controlled study comparing exercise reconditioning plus threshold inspiratory muscle training with exercise reconditioning alone, the former resulted in significantly greater increases in inspiratory muscle strength and endurance and in exercise tolerance (108). Exercise tolerance seemed to be improved particularly for those with electromyographic changes indicating inspiratory muscle fatigue with exercise (109). One controlled, well-designed, but small study of threshold inspiratory exercise for CF patients demonstrated significant improvements in inspiratory muscle strength, FVC, total lung capacity, and exercise tolerance in the experimental group (110). In another controlled study of patients taking corticosteroids, inspiratory muscle training appeared to prevent the weakness that would have otherwise resulted from the steroid use (111). The combination of inspiratory muscle training along with bronchodilator therapy and reconditioning exercise was demonstrated to very significantly reduce dyspnea by comparison to the use of bronchodilators and general exercise without inspiratory muscle training (112). A recent study showed benefit incorporating inspiratory muscle training into a pulmonary rehabilitation program (113). In general, improvements in inspiratory muscle function and in exercise tolerance were greater for the targeted and threshold studies than for the flow-resistor studies (114).

Respiratory Muscle Rest

Relatively minor changes in the pattern of breathing or respiratory muscle loading can trigger acute respiratory muscle fatigue and failure. Interspersing periods of exercise and muscle rest is a basic principle of rehabilitation. Hypercapnia is an indication of limited reserve before the appearance of overt fatigue and may indicate the need for periods of respiratory muscle assistance or rest before considering strengthening exercises (115). Diaphragm rest can be achieved by assisted ventilation using either body ventilators, mouth piece, or nasal noninvasive positive pressure ventilation (NIV).

Despite high ventilation rates in COPD, ventilatory response to both hypercapnia and hypoxia may be reduced. This is often exacerbated during sleep. The increase in pulmonary vascular resistance that occurs in the presence of pulmonary tissue hypoxia is exacerbated by acidosis. When this situation becomes severe, it may lead to right ventricular failure. The use of oxygen therapy alone may exacerbate CO2 retention and acidosis.

Two groups of patients may be suited to ventilatory assistance at home. The first and smaller group includes those who use ventilatory assistance around the clock, usually by tracheostomy, but who are medically and psychologically stable. These patients tend to require frequent hospital readmission, have a poorer prognosis than ventilator-assisted individuals with NMD, and can be candidates for decannulation to NIV. The second group may benefit from nocturnal assistance alone.

While both nocturnal negative pressure body ventilator (NPBV) use and NIV can normalize arterial blood gases, and have been reported to increase quality of life, 12-minute walking distance, respiratory muscle endurance, and decrease dyspnea (116), the former methods cause obstructive apneas during sleep (117,118). The NIV methods, provided by portable ventilators and bilevel positive airway pressure (BiPAP) machines, can rest inspiratory muscles, assist ventilation, and splint open the airway to prevent sleep apneas and airway collapse (119).

Belman et al. reported greater diaphragm relaxation by nasal ventilation than by use of NPBVs (120). Marino demonstrated reversal of nocturnal ventilatory insufficiency for COPD patients using nasal ventilation (121), and others have used assisted ventilation via oral-nasal interfaces as an alternative to intubation and tracheostomy for COPD patients in acute exacerbation (122). Nasal BiPAP has the additional benefit of the expiratory positive airway pressure countering auto-positive end-expiratory pressure (PEEP) in these patients.
who trap air. This decreases their work of breathing. A number of studies have reported long-term improvements in daytime blood gases with the use of nocturnal nasal ventilation for hypercapnic COPD patients (116,123,124), and it has been suggested that nocturnal ventilator use can also decrease dyspnea, improve quality of life (125), and improve survival (123). Nocturnal nasal BiPAP was also reported to improve sleep efficiency and total sleep time in hypercapnic COPD patients (126). In another study of 49 hypercapnic patients with COPD, while hospitalization rates were decreased by both long-term oxygen therapy (LTOT) and by LTOT with nocturnal nasal ventilation by BiPAP, only the latter group had a significant decrease in intensive-care admissions and a significant improvement in 6-minute walk distance (127). However, a systematic review (76) of nocturnal noninvasive positive pressure ventilation of at least 3 months duration in stable hypercapnic COPD patients did not find a consistent clinically or statistically significant effect on lung function, gas exchange, respiratory muscle strength, sleep efficiency, or exercise tolerance. However, these conclusions are limited by the small size of the underlying studies. Although it is widely considered appropriate to offer nocturnal BiPAP for hypercapnic COPD patients, little or no benefit has been reported with nocturnal use for nonhypercapnic patients for whom its role remains controversial (128). Use of proportional assist ventilation and pressure support during exercise has been reported to facilitate high-intensity exercise training in severe COPD (129–131). Patients benefiting from proportional assist ventilation were reported to achieve 15% higher exercise levels at 6 weeks than those exercising without it. Users had a significant reduction in plasma lactate concentration at equivalent workloads after training.

**Supplemental Oxygen Therapy**

Supplemental oxygen therapy is indicated for patients with $P_aO_2$ continuously less than 55 to 60 mm Hg (132). Home oxygen therapy can decrease pulmonary hypertension, polycythemia, and perception of effort during exercise, and it can prolong life (132,133). Patients with COPD have also been shown to have increased sympathetic modulation and reduced baroreflex sensitivity. Supplemental oxygen has been shown to significantly and favorably alter autonomic modulation and decrease blood pressure and pulse (134). In addition, cognitive function can be improved, or at least, better maintained, and hospital needs reduced.

An international consensus on the current status and indications for LTOT suggested that the prescription be based on (135):

1. An appropriately documented diagnosis
2. Concurrent optimal use of other rehabilitative approaches, such as pharmacotherapy, smoking abstinence, and exercise training
3. Properly documented chronic hypoxemia

Oxygen therapy should be given with caution to hypercapnic patients whether or not they are using NIV (136).

There can also be need for supplemental oxygen during exercise. Many patients exhibit exercise hypoxemia. Decreases in $SpO_2$ are noted at physical activity levels comparable to those necessary to perform ADL. Often, the decrease in $SpO_2$ occurs within the first minute, after which $SpO_2$ stabilizes; but occasionally there is a progressive decline in $SpO_2$ with exercise. In a study of 38 subjects in whom the mean resting $SpO_2$ was $93 \pm 3\%$, a decrease in $SpO_2$ of 4.7 \pm 3.6\% (range: 1\% to 18\%) was observed during submaximal exercise (137).

In a crossover study of 12 subjects with severe COPD (138), four patients more than doubled their duration of exercise while receiving 40\% oxygen, but in only two of these was desaturation observed in the absence of oxygen. Bradley et al. reported that in subjects with mild hypoxemia and exercise desaturation, supplemental oxygen by nasal prongs did not influence maximum work rate but did influence endurance (139). Davidson et al. noted that oxygen increased mean walking endurance time by 59\% and 6-minute walking distance by 17\%. Moreover, submaximal cycle time at a constant workload was increased by 51\% at a flow rate of 2 L/min and by 88\% at 4 L/min, suggesting a dose-response curve (140). The exercise response to oxygen could not be predicted from the degree of desaturation, resting pulmonary function tests, echocardiographic measurements of right ventricular systolic pressure, or other clinical parameters (138). In fact, in nonhypoxemic COPD patients performing moderate exercise, oxygen supplementation decreases ventilatory requirement by its direct effect on chemoreceptor inhibition (141). Thus, exercise tolerance can be increased by oxygen therapy without improving oxygen consumption or utilization. A recent study found that supplemental oxygen during exercise prevented exercise induced oxidative stress (142). Marcus et al. also reported significantly greater exercise tolerance in CF patients receiving oxygen (143). A recent study found improved exercise tolerance in CF patients in the Dead Sea basin compared with at sea level (144). This was thought to be due to the increased oxygen tension below sea level.

The most widely accepted guideline for prescribing oxygen use during exercise is that of exercise-induced $SpO_2$ below 90\%. However, it seems reasonable to recommend that measurements of dyspnea and exercise tolerance be undertaken with and without supplemental oxygen to determine which individuals are less short of breath or walk further (have greater exercise tolerance) when given supplemental oxygen (145). Certainly, exercise-induced decreases in $SpO_2$ below 90\% when combined with increased exercise tolerance with oxygen therapy warrant the prescription of oxygen therapy during exercise.

Patients with mild-to-moderate daytime hypoxemia often have marked nocturnal desaturation. Home overnight oximetry can be used to diagnose nocturnal oxyhemoglobin desaturation and assist in oxygen prescription although guidelines for sleep supplemental oxygen have not been established (146).

Inspiratory phase or pulsed oxygen therapy, especially when delivered transtracheal, avoids waste and decreases discomfort and drying of mucous membranes. Oxygen flow...
delivery is 0.25 to 0.4 L/min compared to 2 to 4 L/min when delivered via nasal cannula or face mask (147,148). Oxygen therapy should be used in combination with mechanical ventilation for patients with concomitant carbon dioxide retention.

Reconditioning Exercise

COPD patients are markedly inactive during and after hospitalizations for acute exacerbations. Those with low activity levels 1 month postdischarge were significantly more likely to be hospitalized during the next year. Patients who had hospitalizations for exacerbations in the previous year were even more deconditioned than patients with recent hospitalizations. Thus, efforts to enhance physical activity and exercise tolerance are especially important for patients experiencing acute exacerbations (149).

Because of decreased efficiency of gas exchange, there is an abnormally high ventilatory requirement and a rapid increase in breathing frequency by comparison to tidal volume during exercise. The COPD patient’s maximal exercise ventilation ($V_{E \text{max}}$) is close to or exceeds maximum voluntary ventilation. Cardiac output rises normally with exercise, but exercise can cease at relatively low heart rates because of the ventilatory limitation associated with dyspnea. Hypoxia—and in severely limited patients hypercapnia—may occur with exercise (150). Thus, not all patients can attain the 60% to 70% of predicted maximum heart rate and the minute oxygen consumption needed for cardiac or aerobic exercise training. Even when the ventilatory anaerobic threshold (VAT) cannot be achieved, however, maximum symptom-limited oxygen consumption does correlate with ability to perform ADL, and it can be significantly increased with exercise training (151).

Independent of LTOT, it has been suggested that re habilitation with exercise reconditioning may prolong survival for COPD patients (152). For 149 patients, 89% of whom had COPD and mean age of 69 years, age, sex, body mass index, and primary diagnosis were not related to survival after completion of a pulmonary rehabilitation program. However, a higher postrehabilitation functional activities score, a longer postrehabilitation 6-min walk distance, and being married were strongly associated with increased survival (153). Gerard et al. also demonstrated that postrehabilitation 12-min walk was a strong predictor of survival and much better than arterial partial pressure of oxygen (PaO₂), partial pressure of carbon dioxide (PaCO₂), FEV₁, and nutritional status (154). Thus, exercise tolerance is extremely important for predicting survival. A recent study suggests the importance of ongoing rehabilitation for COPD patients (155).

In normal male subjects, as well as in COPD patients, the intensity of training is the most important factor in improving maximum oxygen consumption ($VO_{2 \text{max}}$). Greater exercise tolerance results from high-intensity programs. Punzal et al. demonstrated that moderately to severely affected COPD patients could exercise at 95% of the baseline maximum treadmill workload. At training weeks 1, 4, and 8, they can train at 85%, 84%, and 86%, respectively, of initial baseline maximum. Maximum treadmill workload, $VO_{2 \text{max}}$, and endurance exercise time can be increased and symptoms decreased in this manner (156). The minority of patients who did not reach the VAT were able to train at a higher percentage of maximum exercise tolerance than patients who reached VAT, but the increase in exercise performance was similar for both groups. Indeed, moderate to severely affected COPD patients can perform exercise training successfully at intensity levels that represent higher percentages of maximum, than typically recommended in normal individuals or other patients (156). Although both men and women benefit from short-term (3-month) exercise therapy with improvements in dyspnea, fatigue, and emotional function, men, but not women, continued to benefit from long-term exercise therapy (157).

The VAT is defined as the highest rate of oxygen consumption during exercise above which sustained lactic acidosis occurs. The VAT can be consistently achievable for patients with less than very advanced disease, including patients with 0.5 L of FEV₁. The parameters of a 45-minute, maximally intense, training prescription can be derived from bilevel exercise in cycles of 4 minutes of exercise at the VAT (lower level) and 1 minute at the $VO_{2 \text{max}}$ (upper level) (158). For the minority of cases for whom the VAT cannot be achieved, bilevel exercise can be provided by using the maximum $VO_{2}$ attained for the upper (1-minute) level and 40% of this figure for the lower (4-minute) level. These intensities can be maintained for 45 minutes in a maximally intense endurance exercise program using a cycle ergometer, such as a Square Wave Endurance Exercise Trainer (158).

Imprecisely prescribed exercise regimens—for example, when the exercise is too intense to be tolerated for a 45-minute session—will result in patients failing to complete the prescribed sessions and deriving suboptimal benefits on endurance. Others, for whom the prescription is submaximal, will be able to complete the 45-minute sessions but at exercise intensities inadequate for optimal reconditioning. Submaximal exercise programs have not been demonstrated to increase $VO_{2 \text{max}}$ or reduce lactate production. Increases in $VO_{2 \text{max}}$ can result in both lower ventilation requirement at given exercise levels and increased exercise tolerance.

Some advocate guiding initial exercise intensity by parameters of clinical exercise testing, for example, heart rate at ventilation levels of 37.5 times the FEV₁ (24). Such patients can increase exercise ventilation and sustain it at a high percentage of their maximum voluntary ventilation. With training, they can exceed the levels attained during initial exercise testing (106). For example, 34 patients in one study could walk at a work level of 86% of their baseline maximum for a mean duration of 22 minutes (106). Carter et al. also trained patients at near their ventilatory limits and reported mean peak exercise ventilation of 94% to 100% of measured maximum voluntary ventilation (159). Training above the VAT leads to a reduced ventilatory requirement during exercise and, therefore, improved maximum exercise tolerance (159).

Others prescribe maximum sustainable exercise by encouraged 6-minute walking (160). This can be a high-intensity
submaximal exercise protocol in which oxygen uptake plateaus after the third minute. Relatively high levels of oxygen consumption to 1.6 ± 0.3 L/m were achieved in eight subjects with moderate COPD and exercise tolerance increased (161). The pulse should return to baseline 5 to 10 minutes after exercise.

Walking, stair climbing, calisthenics, bicycling, and pool activities may be used for reconditioning. The patient is made responsible for a progressive program. As a minimum, we recommend the purchase of a stationary bicycle for the patient’s home and prescribe its daily use. A daily 12-minute walk is recommended, as well as daily 15-minute sessions of inspiratory muscle training. A daily log of time and distance walked, bicycled, and the inspiratory resistance tolerated during the 15-minute sessions, provides feedback to both the patient and the physician. A typical program consists of weekly re-evaluations for 10 to 12 weeks, during which the patient logs are reviewed and exercise parameters modified. Educational and peer groups sessions reinforce the activities. Exercises for skeletal musculature that include bench presses, lateral pull down, leg extension, and presses may improve the 6-minute walk (160). Such programs are inexpensive, minimally intrusive, and optimize the chances for continued independent adherence to the protocol following the 10- to 12-week supervision period.

Upper Extremity Exercise
Upper extremity reconditioning is part of any comprehensive program (162–167). Many arm and shoulder muscles are also accessory muscles of respiration and, as such, are very active for patients with COPD. The overlap in function explains why COPD patients are particularly short of breath when performing upper extremity ADL.

Unsupported upper extremity activities range from typing, lifting, reaching, and carrying, to athletic activities and personal daily care (eating, grooming, cleaning). Unsupported arm exercise shifts work to the diaphragm, leading to earlier fatigue (164). In a randomized controlled trial comparing supported arm exercise with unsupported arm exercise within a general rehabilitation program for COPD patients, the group performing the unsupported upper extremity exercise demonstrated significantly greater improvements and reduced oxygen consumption during upper extremity exercise than the supported upper extremity exercise group (165). Other studies have substantiated the greater benefits to be derived from unsupported rather than supported upper extremity exercise (164–168).

Physical Aids
Wheelchairs and other assistive equipment are covered elsewhere in this text (see Chapters 50 and 74). However, certain aids like motorized scooters and Rollators can greatly improve function and quality of life for COPD patients. Rollators (essentially rolling walkers that permit the user to sit) were reported to be effective in improving functional exercise capacity by reducing dyspnea and rest duration among stable individuals with severe COPD (169,170). Those who walked less than 300 m and those who required a rest during an unaided 6-minute walk benefited the most by reduced dyspnea, reduced rest time, and improved distance walked. Hospital beds with an overhead trapeze, reachers, elevated toilet seats, and strategically placed hand rails at home can also be very useful.

The Results of Pulmonary Rehabilitation
In a review of 48 pulmonary rehabilitation studies that included exercise reconditioning, exercise tolerance improved significantly in all 48, including 14 controlled studies (29). It improved significantly (proportionally) in patients with mild or advanced (hypercapnic) disease (171). Consistent improvements included decreases in the ventilatory equivalent or the ventilation/oxygen consumption ratio, increases in work efficiency (external work per unit of oxygen consumed) and thus in exercise tolerance, ambulation capacity, general well-being, and dyspnea tolerance. The patients developed better performance strategies and greater confidence in performing the tests. In many programs, decreases in blood lactate levels were also observed in combination with higher VO2max, implying a physiologic training effect as well as improved motivation and effort. Maximum tolerated intensity exercise regimens yielded better results than low-intensity exercise for proportionally longer periods (158,172,173). Twenty session out-patient programs result in better outcomes than ten session programs (174), and minimally supervised home exercise programs can also result in improved health status (175).

In general, peak performance appeared to be reached in 26 to 51 weeks and lasted for as long as 5 years (172,173,176–178). Quality-of-life measures (29,179–181), hospitalization rates, postoperative pulmonary complications (182), and physical functioning were likewise reported to be significantly improved in the majority of both the controlled and the repetitive measure studies (29). Outcomes were equivalent for inpatient and outpatient programs, although many of the outpatient programs were predominantly home based. Pulmonary function parameters, such as FEV1, did not significantly improve in 31 of 35 studies (29).

Thus, virtually all studies indicate that pulmonary rehabilitation including exercise training results in significant increases in ambulation capacity and exercise endurance for COPD patients, as well as for many patients with other intrinsic lung pathology (183). The often-reported decreased resting oxygen consumption and carbon dioxide production may, at least in part, account for the significant decrease in perception of dyspnea, the general increase in functional performance, and in the often-found improved sense of well-being. Indeed, a recent Cochrane meta-analysis of pulmonary rehabilitation in COPD (184) concluded that “Rehabilitation relieves dyspnea and fatigue, improves emotional function, and enhances patients’ control over their condition. These improvements are moderately large and clinically significant. Rehabilitation forms an important component of the management of COPD.”

One study of 120 advanced hypercapnic COPD patients, 117 of whom with FEV1 less than 1 L, demonstrated significant
increases in PaO₂, vital capacity (VC), FEV₁, maximum inspiratory pressures, and ambulation tolerance with pulmonary rehabilitation, including exercise reconditioning. Maximum expiratory pressures did not change significantly. The higher the initial PaCO₂, the more PaCO₂ fell and PaO₂ rose during the rehabilitation program. The patients’ improvement in performance of ADL correlated with increased walking distance. Thus, hypercapnia is not a contraindication to intensive rehabilitation, and the use of rigorous reconditioning exercise does not precipitate respiratory muscle fatigue in this population (171).

In a study of 61 of 80 bedridden ventilator-dependent COPD patients in a respiratory care unit, 60 underwent a stepwise rehabilitation program consisting of skeletal and respiratory muscle training and early ambulation. Twenty percent of both the rehabilitation and nonrehabilitation patients died while in the unit. However, all but nine of the rehabilitation patients were weaned from invasive mechanical ventilation after a mean stay of 38 ± 14 days (34). Six-minute walk distances and mean inspiratory pressure increased significantly more in the rehabilitation group.

Following the acute rehabilitation period, continued surveillance and attention to abstinence from smoking, bronchial hygiene, breathing retraining, physical reconditioning, oxygen therapy, and airway secretion mobilization have been shown to reduce hospital admissions, the length of hospital stays, and cost (185). As noted, the benefits of pulmonary rehabilitation on exercise performance and quality of life persisted for up to 5 years (172,173,176,177,186–189). Pulmonary function parameters, dyspnea, exercise tolerance, and quality of life are not further improved long term by yearly repeated 2-month outpatient pulmonary rehabilitation interventions, but rates of hospitalizations and acute exacerbations are further decreased by repeated interventions (190). The principles of rehabilitation for COPD are being increasingly applied to patients with asthma with similar outcomes (191).

**PHYSICAL MEDICINE AND HABILITATION OF PATIENTS WITH PARALYTIC AND NEUROMUSCULAR IMPAIRMENTS**

Physiatrists have been instrumental in developing the habilitation techniques, orthotics, and assistive equipment that have been used to optimize quality of life for patients with neuromuscular disability (192). Since the early 1950s, they have also reported most of the physical medicine aids that have been used to prevent respiratory failure for patients with dysfunctional respiratory muscles (193). Despite this history, few physiatrists today are aware of these methods.

**Pathophysiology**

Ventilatory insufficiency is defined by hypercapnia in the presence of a normal arterial-alveolar A-a gradient. It is hypercapnia not caused by intrinsic lung disease, or irreversible upper or lower airway obstruction, as in COPD. These patients can have airway obstruction from bronchial mucous plugging that causes an elevated A-a gradient. However, the mucous plugging is reversible by using expiratory (coughing) aids. Symptomatic hypercapnic patients benefit from the use of noninvasive ventilation for at least part of the day and, more often overnight. With progressive inspiratory muscle weakness, ventilator-free breathing ability is eventually lost.

Ventilatory muscle failure is defined by the inability of the inspiratory and expiratory muscles to sustain one’s respiration without resorting to ventilator use. Ventilatory insufficiency leading to failure can be nocturnal only resulting from diaphragm dysfunction with the patient unable to breathe when supine, or can result from a lack of central ventilatory drive, or can result from severe generalized respiratory muscle dysfunction. Many patients with ventilatory insufficiency survive for years without ventilator use with the cost of orthopnea and increasing hypercapnia, associated symptoms and dangers, and a compensatory metabolic alkalosis that depresses central ventilatory drive. The alkalosis allows the brain to accommodate to hypercapnia without overt symptoms of acute ventilatory failure. Hypercapnic patients not using NIV, and especially those receiving supplemental oxygen, develop increasingly severe hypercapnia that eventually results in coma from carbon dioxide narcosis and ventilatory arrest. When symptomatic, hypercapnic patients are treated with NIV, blood gases normalize, and the alkalosis resolves as the kidneys excrete excess bicarbonate ions. Because of the need to take bigger breaths to maintain normal PaCO₂ and blood pH levels, a nocturnal-only ventilator user can become dyspneic when discontinuing ventilator use in the morning and require gradually increasing periods of daytime NIV and eventually become continuously NIV dependent. Such patients eventually require NIV for increasing periods during the day until continuously NIV dependent. Some patients with ventilatory muscle failure and no measurable VC with their respiratory muscles use only nocturnal aid and rely on glossopharyngeal breathing (GPB) to ventilate their lungs during daytime hours.

For patients with primarily ventilatory impairment, respiratory morbidity and mortality are a direct result of respiratory muscle dysfunction and can be avoided by assisting respiratory muscles, as long as bulbar muscle dysfunction is not so severe that the SpO₂ remains below 95% because of continuous aspiration of saliva. Such patients develop essentially irreversible upper airway obstruction and require tracheostomy tubes to protect the airway. In general, this scenario occurs for bulbar amyotrophic lateral sclerosis (ALS) patients who have lost the ability to speak and for few other patients with NMD.

**Epidemiology**

There are 500,000 people in the United States, or 0.15% of the population, with pediatric NMD (194), and many more with thoracic wall restrictive lung disease. One out of 3,500 boys is born with DMD, one in 5,000 children have spinal muscular atrophy (SMA), and 1 in 1,800 people develop ALS. It has been estimated that of the patients who had acute
poliomyelitis, 1.63 million with a median age of 57 were still alive in 1987 (194). Many of these patients have ventilatory insufficiency and require ventilator use currently. Respiratory muscle dysfunction amenable to treatment by respiratory muscle aids occurs in many people with the diagnoses listed in Table 42-3.

Surveys in the United States, Western Europe, and Japan indicate that the use of home mechanical ventilation is increasing rapidly (194). Most people who appear to require more than nocturnal ventilator use, however, and those who have difficulty weaning when intubated during chest infections generally undergo tracheotomy. This usually occurs because of lack of awareness of how to introduce physical medicine aids. Likewise, 90% of episodes of ARF for young patients with NMD occur as a result of inability to generate effective cough peak flows (CPF) during otherwise benign upper respiratory infections (195). These episodes are usually avoidable (196).

**What Are Physical Medicine Respiratory Muscle Aids?**

Inspiratory and expiratory muscle aids are devices and techniques that involve the manual or mechanical application of forces to the body, or intermittent pressure changes to the airway, to assist inspiratory or expiratory muscle function. The devices that act on the body include body ventilators that create pressure changes around the thorax and abdomen. Negative pressure applied to the airway during expiration assists coughing, just as positive pressure applied to the airway during inhalation (NIV) assists the inspiratory muscles. Continuous positive airway pressure (CPAP) does not assist ventilation and is not useful for patients with primarily ventilatory impairment.

**TABLE 42.3 Conditions with Chronic Alveolar Hypoventilation Manageable with Respiratory Muscle Aids**

| Myopathies: Congenital, metabolic, inflammatory, and mitochondrial myopathies, myopathies of systemic disease such as carcinomatous myopathy, cachexia/anorexia nervosa, medication and ICU-associated, muscular dystrophies such as Duchenne and Becker, limb girdle, Emery-Dreifuss, facioscapulohumeral, congenital, and myotonic |
| Endocrine related as with hypothyroidism, acromegaly |
| Mixed connective tissue disease and arthrogryposis |
| Anterior horn cell disorders: Spinal muscular atrophies, motor neuron diseases, poliomyelitis |
| Neuropathies: Hereditary sensory motor neuropathies |
| Familial hypertrophic interstitial polyneuropathy |
| Phrenic neuropathies |
| Guillain-Barré syndrome |
| CNS disorders: Multiple sclerosis |
| Disorders of supraspinal tone such as Friedreich’s ataxia |
| Central hypoventilation syndromes, hypoventilation associated with diabetic microangiopathy, familial dysautonomia, Prader-Willi syndrome |
| Encephalopathies: Syringomyelia, myelomingingocele |
| Myelopathies: Traumatic and spondylitic medical including rheumatoid, infectious, vascular, botulism/pancuronium bromide, or idiopathic |
| Restrictive lung disease/increased work of breathing: |
| Obesity-hypoventilation, Kyphoscoliosis |
| Lung resection/tuberculosis, Milroy’s disease, congenital diaphragmatic hernia, vocal cord paralysis/postlaryngotracheal reconstruction |
| Mixed ventilatory-respiratory impairment like COPD, CF |

**Patient Evaluation**

Patients with diminished ventilatory reserve who are able to walk commonly complain of exertional dyspnea. Eventually, morning headaches, fatigue, sleep disturbances, and hypersomnolence develop (197). For wheelchair users, symptoms may be minimal, except during intercurrent respiratory infections when anxiety, inability to fall asleep, and dyspnea become problems.

The patient is observed for increased respiratory rate, decreased depth, or irregularity of breathing. Paradoxical breathing, or asymmetric movement of the abdomen or thorax is often present. Hypophonia, nasal alae flaring, use of auxiliary respiratory musculature, peribuccal or generalized cyanosis, flushing or pallor, hypertension, difficulty controlling airway secretions, dysphagia, regurgitation of fluids through the nose, nasality of speech, cor pulmonale, confusion, and fluid retention may all be signs of ventilatory insufficiency.

Maximum inspiratory and expiratory pressures generated at the mouth correlate best with inspiratory and expiratory muscle strength. Maximum voluntary ventilation gauges respiratory muscle endurance. The VC gives an indication of both these parameters and is simple, easy to measure, objective, and very reproducible. Because hyperventilation is often worse during sleep, the supine rather than sitting position VC is the most important indicator of ventilatory dysfunction. Spirometry is also useful for monitoring progress with GPB and air stacking. A patient’s maximum insufflation capacity (MIC) is determined by giving the patient the largest volume of air that can be held with a closed glottis from a manual resuscitator or a portable ventilator that is volume cycled. The patient then expels the air into the spirometer. Patients who learn GPB can often air stack consecutive GPB gulps to or beyond the MIC (198). A nasal interface or lip seal can be
used for air stacking when the lips are too weak for effective air stacking via the mouth (Fig. 42-2).

Unassisted and assisted CPF are measured using a peak flow meter (Access Peak Flow Meter, Healthscan Products Inc., Cedar Grove, NJ). CPF of 160 L/m are the minimum needed to cough effectively (199), and this is the best indicator for tracheostomy tube removal irrespective of remaining pulmonary function (200) (Fig. 42-3). Indeed, almost 40% of patients with ALS can survive despite continuous ventilator dependence using strictly noninvasive aids (201). Patients with VC s less than 1,500 mL have assisted CPF measured from a maximally stacked volume of air and with an abdominal thrust delivered simultaneously with glottic opening.

For the stable patient without intrinsic pulmonary disease, arterial blood gas sampling is unnecessary. Besides the discomfort, 25% of patients hyperventilate as a result of anxiety or pain during the procedure (99). Noninvasive continuous blood gas monitoring, including capnography and oximetry, yield more useful information, particularly during sleep.

Nocturnal noninvasive blood gas monitoring can be performed for patients with diminished supine VC, especially for those with rapidly evolving conditions and symptoms suggestive of hypoventilation. The oximeter and the capnograph, which measures end-tidal $pCO_2$, must be capable of summarizing and printing out the data (197). These studies are most conveniently performed in the home. When symptoms are obvious, a trial of nocturnal NIV rather than nocturnal end-tidal $CO_2$ and oximetry monitoring is more appropriate.

Any symptomatic patient with decreased VC, multiple nocturnal oxyhemoglobin desaturations below 95%, and elevated nocturnal PaCO$_2$ certainly requires treatment for nocturnal hypoventilation. For symptomatic patients with normal VC, an unclear pattern of oxyhemoglobin desaturation, and no apparent carbon dioxide retention, sleep disordered breathing is suspected. This is particularly true when loud high-pitched snoring, interrupted breathing, and hypersomnolence dominate the picture (202). These patients undergo polysomnography and are considered for CPAP therapy. Obesity-hypoventilation patients are treated with nocturnal ventilatory support, as are NMD patients. When concurrent COPD or interstitial lung disease is documented, capnography is correlated to PaCO$_2$.

**The Intervention Objectives**

The intervention goals are to maintain lung and chest-wall compliance and to promote normal lung and chest-wall growth for children, to maintain normal alveolar ventilation

---

**FIGURE 42-2.** Fifteen-year-old patient with acute high-level spinal cord injury, continuously ventilator dependent from 12 h postinjury. Never intubated, he uses lip seal intermittent positive pressure ventilation (IPPV) while sleeping (seen here) and a simple mouth piece for IPPV while awake, without ventilator-free breathing ability, for 10 days before weaning from ventilator use. Because there is no neck movement, the lip seal is used with a single Velcro strap to the posterior poster rather than the customary two straps.

**FIGURE 42-3.** Cough peak flows increase from ineffective levels (below 160 L/m) to effective levels by maximally insufflating the patient and then providing an abdominal thrust. (With permission and appreciation to Dr. Yuka Ishikawa, Dept. of Pediatrics Yakumo Byoin National Sanatorium, Yakumo-cho Hikkarclo.)
around the clock, and to maximize CPF. The long-term goals are to avert episodes of ARF during intercurrent chest infections, avoid hospitalizations, and prolong survival without resorting to tracheotomy. All goals can be attained by evaluating, training, and equipping patients in the outpatient setting and at home.

**Goal One: Maintain Pulmonary Compliance, Lung Growth, and Chest-Wall Mobility**

Pulmonary compliance is lost because the ability to expand the lungs to the predicted inspiratory capacity is lost as the VC decreases. As the VC decreases, the largest breath that one can take can only expand a small portion of the lungs. Like limb articulations and other soft tissues, regular range of motion (ROM) is required to prevent chest-wall contractures and lung restriction. This can only be achieved by providing deep inspirations, air stacking, or nocturnal NIV (203). The extent to which the MIC is greater than the VC predicts the capacity of the patient to be maintained by noninvasive rather than tracheostomy ventilatory support (199). This is because the MIC:VC difference, like assisted CPF, is a function of bulbar muscle integrity. Patients who cannot close the glottis and, therefore, cannot air stack, must be passively insufflated using a手动 resuscitator. In general, because of the importance of ventilatory support. Such interfaces deliver air via mouthpieces, nasal, and oral-nasal interfaces for nocturnal ventilatory support. Mouthpiece and nasal IPPV are open systems that require the user to rely on central nervous system reflexes to prevent excessive insufflation leakage during sleep (197,210).

The NPBVs available today include the iron lung, the Porta-Lung, chest-shell ventilator, and various wrap-style ventilators. They are only practical for use during sleep. Although NPBVs have been used by patients with little or no VC for decades, they become ineffective with aging and decreasing pulmonary compliance (117). Such patients are switched to noninvasive IPPV (207,208). Although NPBVs continue to be used in a few centers as a “bridge” to noninvasive IPPV while extubating unweanable patients (209), it is debatable whether their long-term use is ever warranted today due to the positive pressure devices available.

**Goal Two: Continuously Maintain Normal Alveolar Ventilation by Assisting Inspiratory Muscles as Needed**

I. The Nocturnal Inspiratory Muscle Aids

A. Negative pressure body ventilators

The NPBVs available today include the iron lung, the Porta-Lung, chest-shell ventilator, and various wrap-style ventilators. They are only practical for use during sleep. Although NPBVs have been used by patients with little or no VC for decades, they become ineffective with aging and decreasing pulmonary compliance (117). Such patients are switched to noninvasive IPPV (207,208). Although NPBVs continue to be used in a few centers as a “bridge” to noninvasive IPPV while extubating unweanable patients (209), it is debatable whether their long-term use is ever warranted today due to the positive pressure devices available.

B. Noninvasive intermittent positive pressure ventilation

Although for acutely ill patients, introduction of noninvasive ventilation needs to be done in the hospital setting, the great majority of noninvasive IPPV users are introduced to it in the clinic or home setting.

IPPV can be noninvasively delivered via mouthpieces, nasal, and oral-nasal interfaces for nocturnal ventilatory support. Mouthpiece and nasal IPPV are open systems that require the user to rely on central nervous system reflexes to prevent excessive insufflation leakage during sleep (197,210).

There are numerous commercially available nasal interfaces (CPAP masks). Each interface design applies pressure differently to the paranasal area. One cannot predict which model will be most effective and preferred by any particular patient. Nasal bridge pressure and insufflation leakage into the eyes are common symptoms with several of these generic models. Such difficulties can be avoided by using nasal prong systems or custom designs (6,197,211). No patient should be offered and expected to use only one nasal interface. Alternating IPPV interfaces nightly alternates skin pressure sites, minimizes discomfort, and is to be encouraged.

Excessive insufflation leakage can be avoided by switching to the use of lip seal–nasal prong systems that provide an essentially closed system of noninvasive ventilatory support. Such interfaces deliver air via mouth and nose during sleep with minimal strap pressure. This optimizes skin comfort and minimizes air (insufflation) leakage. Excessive leakage can also be prevented.
by sustaining ventilatory drive by maintaining normal daytime CO$_2$ and avoiding supplemental O$_2$ and sedatives.

II. The daytime inspiratory muscle aids
C. Body ventilators

The intermittent abdominal pressure ventilator (IAPV) involves the intermittent inflation of an elastic air sac that is contained in a corset or belt worn beneath the patient's outer clothing (Fig. 42-4) (Exsufflation Belt, Respironics Inc., Murrysville, PA). The sac is cyclically inflated by a positive pressure ventilator. Bladder inflation moves the diaphragm upward to assist in expiration. During bladder deflation, gravity causes the abdominal contents and diaphragm to return to the resting position, and inspiration occurs passively. A trunk angle of 30 degrees or more from the horizontal is necessary for it to be effective. If the patient has any inspiratory capacity or is capable of GPB, he or she can add volumes of air autonomously taken in to that taken in mechanically. The IAPV generally augments tidal volumes by about 300 mL, but volumes as high as 1,200 mL have been reported (212). Patients with less than 1 hour of breathing tolerance usually prefer to use the IAPV rather than use noninvasive IPPV during daytime hours.

D. Mouthpiece intermittent positive pressure ventilation

Mouthpiece IPPV is the most important method of daytime ventilatory support. Some patients keep the 15-mm angled mouthpiece between their teeth all day. Most patients prefer to have the mouthpiece held near the mouth. A metal clamp attached to a wheelchair can be used for this purpose, or the mouthpiece can be fixed onto motorized wheelchair controls—most often, sip and puff, chin, or tongue controls (Fig. 42-5). The ventilator is set for large tidal volumes, often 1,000 to 2,000 mL. The patient grabs the mouthpiece with his mouth and supplements or substitutes for inadequate autonomous breath volumes. The patient varies the volume of air taken from ventilator cycle to ventilator cycle and breath to breath to vary speech volume and cough flows as well as to air stack to fully expand the lungs.

To use mouthpiece IPPV effectively and conveniently, adequate neck rotation and oral motor function

FIGURE 42-4. A postpolio survivor with no measurable VC since 1952 using an IAPV (Exsufflation Belt, Respironics International Inc., Murrysville, PA) during daytime hours and lip seal IPPV nightly since 1956. The air bladder inside the girdle is connected to the ventilator circuit (seen here), then the girdle is placed under the clothes and over the patient's abdomen.
are necessary to grab the mouthpiece and receive IPPV without insufflation leakage. To prevent the latter, the soft palate must move posterior and caudally to seal off the nasopharynx. In addition, the patient must open the glottis and vocal cords, dilate the hypopharynx, and maintain airway patency to receive the air. These normally reflex movements may require a few minutes to relearn for patients who have been receiving IPPV via invasive tubes.

E. Nasal intermittent positive pressure ventilation

Because patients prefer to use mouthpiece IPPV or the IAPV for daytime use (163,183), nasal IPPV is most practical for nocturnal use. Daytime nasal IPPV is indicated for infants and for those who cannot grab or retain a mouthpiece because of oral muscle weakness, inadequate jaw opening, or insufficient neck movement. Nevertheless, 24-hour nasal IPPV can be a viable and desirable alternative to tracheostomy, even for some patients with severe lip and oropharyngeal muscle weakness (197). Nasal IPPV users learn to close their mouths or seal off the oropharynx with their soft palates and tongues to prevent oral insufflation leakage.

Complications of Noninvasive Intermittent Positive Pressure Ventilation

Besides orthodontic deformities and skin pressure from the interface, other potential complications include very infrequent allergy to the interface, dry mouth (65%), eye irritation from air leakage, nasal congestion (25%) and dripping (35%), sinusitis (8%), nose bleeding (4% to 19%), gum discomfort and reeding from nasal interface or bite-plate pressure (20%), maxillary flattening in children, aerophagia (213), and, as for invasive ventilation, barotrauma. In addition, occasional patients experience claustrophobia. Proper interface selection eliminates or minimizes these difficulties.

Pressure drop-off through the narrow air passages of the nose is normally between 2 and 3 cm H2O. Suboptimal humidification dries out and irritates nasal mucous membranes, causes sore throat, and results in vasodilatation and nasal congestion. Increased airflow resistance to 8-cm H2O can be caused by the loss of humidity that is due to unidirectional airflow with expiration via the mouth during nasal CPAP or IPPV (214). This problem cannot be ameliorated by using a cold pass-over humidifier, but the increase in airway resistance can be reduced by 50% by warming the inspired air to body temperature and humidifying it with the use of a hot-water bath humidifier (214). Decongestants can also relieve sinus irritation and nasal congestion. Switching to lip seal–only interface can relieve most if not all difficulties associated with nasal IPPV. There are no absolute contraindications to the long-term use of noninvasive inspiratory muscle aids (215). Relative contraindications are listed in Table 42-4.

Abdominal distention tends to occur sporadically in noninvasive IPPV users. The air usually passes as flatus once the patient gets up or is placed into a wheelchair in the morning. When severe, it can present as intestinal pseudo-obstruction with diminished bowel sounds and increased ventilator dependence. A rectal tube can usually decompress the colon; a gastrostomy tube, when present, can be burped; or a nasogastric tube can be passed to relieve the problem.

Barotrauma results from rupture of the boundary between the alveoli and the bronchovascular sheath. Although its incidence has been cited as 4% to 15% for intensive care unit invasive ventilation users with primarily respiratory impairment (216), in 139 patients it has been reported to occur in 60% of those with acute respiratory distress syndrome but was absent for those with congestive heart failure or neurologic

<table>
<thead>
<tr>
<th>TABLE 42.4</th>
<th>Relative Contraindications for Long-Term Noninvasive Intermittent Positive Pressure Ventilation (IPPV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Lack of cooperation or use of heavy sedation or narcotics</td>
<td></td>
</tr>
<tr>
<td>2. Need for high levels of supplemental oxygen therapy</td>
<td></td>
</tr>
<tr>
<td>3. Oxyhemoglobin saturation (SpO2) cannot be maintained above 94% despite noninvasive IPPV and optimal use of assisted coughing techniques when needed</td>
<td></td>
</tr>
<tr>
<td>4. Substance abuse or uncontrollable seizures</td>
<td></td>
</tr>
<tr>
<td>5. Bulbar-innervated muscle impairment with inability to close the glottis</td>
<td></td>
</tr>
<tr>
<td>6. Conditions that interfere with the use of IPPV interfaces, i.e., facial fractures, inadequate bite for mouthpiece entry, presence of facial hair that hamper airtight interface seal</td>
<td></td>
</tr>
<tr>
<td>7. Inadequate caregiver support</td>
<td></td>
</tr>
</tbody>
</table>

FIGURE 42.5. Forty-one-year-old man with Duchenne muscular dystrophy who has used 24-h mouthpiece IPPV for 29 years, now with less than 1 minute of breathing tolerance. The mouthpiece is fixed adjacent to the chin/tongue controls of his motorized wheelchair.
disease (217). We have had one case of pneumothorax in over 1000 NIV users.

Inspissated secretions can be life threatening (218,219). Secretion encumbrance for patients with primarily ventilatory impairment results from failure to use assisted coughing. Chronic aspiration of saliva to the extent of lowering baseline SpO₂ can overwhelm normally sterile airways and lead to pneumonia (220), tracheitis, bronchitis (221), and chronic lung disease (220). The indication for tracheostomy is SpO₂ baseline less than 95% as a result of chronic airway secretion aspiration. This is rare other than for averbal, severe bulbar impaired ALS patients.

**Goal Three: Provide Functional Coughs by Assisting Expiratory Muscles**

I. Why are expiratory muscle aids needed?

Bulbar, inspiratory, and expiratory muscles are needed for effective coughing. The latter are predominantly the abdominal and intercostal muscles. Clearing airway secretions and airway mucus can be a continual problem for patients who cannot swallow saliva or food without aspiration. For patients with respiratory muscle dysfunction and functional bulbar musculature, it becomes a problem during chest infections, following general anesthesia, and during any other periods of bronchial hypersecretion.

II. Manually assisted coughing

Assisted CPF can be greatly increased in patients receiving maximal insufflations followed by manual thrusts for assisted coughing (222). In 364 evaluations of our NMD patients able to air stack, the mean VC in the sitting position was 996.9 mL, the mean MIC was 1647.6 mL, and although CPFs were 2.3 L/s (<2.7 L/s or the minimum needed to eliminate airway secretions) mean assisted CPF were 3.9 L/s.

Techniques of manually assisted coughing involve different hand and arm placements for expiratory cycle thrusts (Fig. 42-6). An epigastric thrust with one hand, while applying counterpressure across the chest to avoid paradoxical chest expansion with the other arm, further increases assisted CPF for 20% of patients (223). Manually assisted coughing is usually ineffective in the presence of severe scoliosis because of a combination of restricted lung capacity and the inability to effect diaphragm movement by abdominal thrusting because of severe rib-cage and diaphragm deformity. Abdominal compressions should not be used for 1 to 1.5 hours following a meal. When inadequate, and especially when inadequacy is due to difficulty air stacking, the most effective alternative for generating optimal CPF and clearing airway secretions is the use of mechanical insufflation-exsufflation (MI-E).

The inability to generate more than 2.7 L/s or 160 L/m of assisted CPF despite having a VC or MIC greater than 1 L usually indicates fixed upper-airway obstruction or severe bulbar muscle weakness and hypopharyngeal collapse during coughing attempts. Vocal cord adhesions or paralysis may have resulted from a previous translaryngeal intubation or tracheostomy (224). Because some lesions, especially the presence of obstructing granulation tissue, can be corrected surgically, laryngoscopic examination is warranted especially before decanulation.

III. Mechanical insufflation-exsufflation

A. Introduction of MI-E

MI-Es (Cough Assist TM, Respironics Internation Inc., Murrysville, Pa) deliver deep insufflations followed immediately by deep exsufflations. The insufflation and exsufflation pressures and delivery times are independently adjustable. Insufflation to exsufflation pressures of +40 to –40 cm H₂O are usually the most effective and preferred by most patients. Onset of insufflation generates an insufflation flow peak and a lung insufflation of more than 2 L. Mechanical exsufflation generates two exsufflation flow notches.
One occurs when the insufflation pressure stops and is due to the elastic recoil of the lung. The second one, a bit greater, is caused by the exsufflation pressure itself. Except after a meal, an abdominal thrust is applied in conjunction with the exsufflation (MAC, or mechanically assisted coughing). MI-E can be provided via an oral-nasal mask, a simple mouthpiece, or via a translaryngeal or tracheostomy tube. When delivered via the latter, the cuff, when present, should be inflated. The Cough Assist can be manually or automatically cycled. Manual cycling facilitates caregiver-patient coordination of inspiration and expiration with insufflation and exsufflation, but it requires hands to deliver an abdominal thrust, to hold the mask on the patient, and to cycle the machine.

One treatment consists of about five cycles of MI-E or MAC followed by a short period of normal breathing or ventilator use to avoid hyperventilation. Insufflation and exsufflation pressures are almost always from +35 to +60 cm H₂O to −35 to −60 cm H₂O. Most patients use 35 to 45 cm H₂O pressures for insufflations and exsufflations. In experimental models, +40 to −40 cm H₂O pressures have been shown to provide maximum forced deflation VCs and flows (225). Insufflation and exsufflation times are adjusted to provide maximum chest expansion and rapid lung emptying. In general, 2 to 4 seconds are required. Multiple treatments are given in one sitting until no further secretions are expelled and any secretion or mucus-induced oxyhemoglobin desaturations are reversed. Use can be required as frequently as every 30 minutes around the clock during chest infections.

The use of MI-E via the upper airway can be effective for children as young as 11 months of age. Patients this young can become accustomed to MI-E and permit its effective use by not crying or closing their glottises. Between 2.5 and 5 years of age, most children become able to cooperate and cough on cue with MI-E. Exsufflation-timed abdominal thrusts are also used for infants.

Whether via the upper airway or via indwelling airway tubes, routine airway suctioning misses the left main stem bronchus about 90% of the time (226). MI-E, on the other hand, provides the same exsufflation flows in both left and right airways without the discomfort or airway trauma of tracheal suctioning. Patients prefer MI-E to suctioning for comfort and effectiveness, and they find it less tiring (227). Deep suctioning, whether via airway tube or via the upper airway, can be discontinued for most patients.

B. Efficacy of MI-E

The efficacy of MI-E has been demonstrated both clinically and on animal models (228). Flow generation is adequate in both proximal and distal airways to eliminate respiratory tract debris (229,230). VC, pulmonary flow rates, and SpO₂ when abnormal improve immediately with clearing of airway secretions and mucus by MI-E (231,232). An increase in VC of 15% to 42% was noted immediately following treatment in 67 patients with “obstructive dyspnea,” and a 55% increase in VC was noted following MI-E in patients with neuromuscular conditions (233). We have observed 15% to 400% (200 to 800 mL) improvements in VC and normalization of SpO₂ as MI-E eliminates airway mucus for ventilator-assisted NMD patients with chest infections (223).

C. Indications for MI-E

Of the three muscle groups required for effective coughing, MI-E can only take the place of the inspiratory and expiratory muscles. Thus, it cannot be used to avert tracheostomy very long if bulbar function is inadequate to prevent airway collapse, as is often the case in advanced bulbar ALS. On the other hand, patients with completely intact bulbar muscle function, such as most ventilator users with traumatic tetraplegia, can usually air stack to volumes of 3 L or more, and, unless very scoliotic or obese, a properly delivered abdominal thrust can often result in assisted CPF of 6 to 9 L/s. These flows should be more than adequate to clear the airways and prevent pneumonia and respiratory failure without need for MI-E. Thus, the patients who need MI-E the most are those whose bulbar muscle function can maintain adequate airway patency but is insufficient to permit optimal air stacking for assisted CPF more than 250 to 300 L/m. This is typical of most nonbulbar ALS/NMD patients. The most typical example of patients who can consistently avoid hospitalization and respiratory failure by using MI-E during intercurrent chest infections is DMD patients (232). Patients with respiratory muscle weakness complicated by scoliosis and inability to capture the asymmetric diaphragm by abdominal thrusting also greatly benefit from MI-E.

**Glossopharyngeal Breathing**

Both inspiratory and, indirectly, expiratory muscle function can be assisted by GPB (198). GPB can provide an individual with weak inspiratory muscles and no VC or breathing tolerance with normal alveolar ventilation when not using a ventilator or in the event of sudden ventilator failure day or night (198,234). The technique involves the use of the glottis to add to an inspiratory effort by projecting (gulping) boluses of air into the lungs. The glottis closes with each “gulp.” One breath usually consists of 6 to 9 gulps of 40 to 200 mL each (Fig, 42-7). During the training period, the efficiency of GPB can be monitored by spirometrically measuring the milliliters of air per gulp, gulps per breath, and breaths per minute. A training manual (235), and numerous videos are available (236), the best of which was produced in 1999 (237).

Although severe oropharyngeal muscle weakness can limit the usefulness of GPB, we have managed 13 DMD ventilator users who had no breathing tolerance other than by
GPB (238). Approximately 60% of ventilator users with no autonomous ability to breathe and good bulbar muscle function can use GPB and discontinue ventilator use for minutes to up to all day (234,239). GPB is also rarely useful in the presence of an indwelling tracheostomy tube. It cannot be used when the tube is uncapped as it is during tracheostomy IPPV, and even when capped, the gulped air tends to leak around the outer walls of the tube and out the stoma as airway volumes and pressures increase during the GPB air-stacking process. The safety and versatility afforded by GPB are additional reasons to eliminate tracheostomy in favor of noninvasive aids.

Because of their generally intact bulbar musculature, spinal cord injury (SCI) patients are ideal candidates to master GPB. High-level SCI patients typically use it for ventilator-free breathing. In some centers, these patients are decannulated to free them from the fear of ventilator failure or accidental ventilator disconnection (234,239). Some patients with no measurable VC have awoken at night, breathing glossopharyngeally, to find that their ventilators had failed, a scenario not possible for tracheostomy IPPV users. When combined with nocturnal noninvasive IPPV, SCI patients can be decannulated, permitting mastery of GPB (Table 42-5) and the security and benefits that accompany decannulation.

**TABLE 42.5 Management of Patients with Spinal Cord Injury**

<table>
<thead>
<tr>
<th>Level</th>
<th>VC, mL</th>
<th>Neck/bulbar muscles</th>
<th>Daytime</th>
<th>Nocturnal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Above C1</td>
<td>0</td>
<td>Inadequate/ inadequate</td>
<td>TIPPV</td>
<td>TIPPV</td>
</tr>
<tr>
<td>C2-C3</td>
<td>&lt;200</td>
<td>Adequate/ inadequate</td>
<td>EPR</td>
<td>N/MIPPV</td>
</tr>
<tr>
<td>Below C2</td>
<td>&gt;200</td>
<td>Adequate/ adequate</td>
<td>MIPPV/IAPV</td>
<td>N/MIPPV</td>
</tr>
</tbody>
</table>

*Motor levels.*

*Adequate neck function involves sufficient oral and neck muscular control to rotate, flex, and extend the neck to grab and use a mouthpiece for IPPV; adequate bulbar function to prevent aspiration of saliva to the degree that the SpO2 baseline decreases below 95%.*

EPR, electrophrenic pacing; IAPV, intermittent abdominal pressure ventilator; MIPPV, mouthpiece intermittent positive pressure ventilation; NIPPV, nasal intermittent positive pressure ventilation; TIPPV, tracheostomy intermittent positive pressure ventilation.

**FIGURE 42-7.** Normal minute ventilation (60 to 90 mL per gulp, 6 to 8 gulps per breath, 12 breaths per minute) throughout daytime hours by GPB for an individual with no measurable VC otherwise. Maximum glossopharyngeal single breath capacities can exceed 3,000 mL for such individuals.

**Maintenance of Respiratory Muscle Strength and Endurance**

In the few studies of respiratory muscle exercise performed on NMD patients, short daily sessions of inspiratory resistive exercise alone were reported to have no effect on spirometry or maximum inspiratory or expiratory pressures (240,241) but did improve respiratory muscle endurance (240,242). However, the degree of improvement in endurance correlated significantly with the level of VC and maximum inspiratory pressure at the outset of training, and no patient with less than 30% of predicted VC improved (242). This is the level of VC at which point patients often require nocturnal ventilatory aid and have considerable difficulty during respiratory tract infections (243). There is also no evidence that beginning an exercise program earlier would preserve more muscle function for the time when the patient requires aid, or that the improvement in endurance for relatively strong patients (mean VC: 54% to 59% in the patients studied) delays the occurrence of pulmonary complications as suggested (240). Indeed, there is some evidence to the contrary. Mildly affected ALS patients were reported to respond to a respiratory muscle–resistive exercise program with a decrease in VC and inspiratory pressures (242). There may also be a greater subsequent rate of loss of muscle strength in any temporarily strengthened muscles. The training itself may be hazardous for advanced patients. Thus, for those most likely to have respiratory complications, the use of resistance exercise is likely to be of little or no value.
Oximetry Monitoring and Feedback Protocol

For a patient with chronic alveolar hypoventilation who has not been using ventilatory support, or the patient being weaned from tracheostomy IPPV, introduction to and use of mouthpiece or nasal IPPV is facilitated by oximetry feedback. An SpO₂ alarm may be set from 93% to 94%. The patient sees that by taking slightly deeper breaths, SpO₂ will exceed 95% within seconds. He or she is instructed to maintain SpO₂ above 94% all day (232). This can be achieved by unassisted breathing for a period of time, and once tiring, by mouthpiece or nasal IPPV. With time, the patient requires increasing periods of IPPV to maintain adequate ventilation (SpO₂ >94%). In this manner, an oximeter may also help to reset central ventilatory drive.

Oximetry feedback is especially important during the management of respiratory tract infections. The cough of infants and small children who can never sit is inadequate to prevent chest cold–triggered pneumonia and respiratory failure. Children who can sit are usually protected from this until after 2 years of age. Older children and adults whose assisted CPF decrease below 300 L/min are also at high risk for chest cold–triggered ARF. Such patients require continuous SpO₂ monitoring and are taught that any dip in SpO₂ below 95% is due either to underventilation, or bronchial mucous plugging, and if these two causes are not quickly addressed, may lead to atelectasis or pneumonia. They are instructed to use noninvasive IPPV to maintain normal ventilation and manually or mechanically assisted coughing to reverse mucous plug associated oxyhemoglobin desaturations. In this way, most episodes that would otherwise cause ARF are successfully managed at home.

For adults with infrequent chest colds, rapid access to this equipment may be all that is necessary. Likewise, with intact bulbar muscle function, manually assisted coughing may be sufficient to generate effective cough flows, whereas with complete bulbar muscle paralysis, even MAC may not spare patients from respiratory failure.

Treatment of Sleep-Disordered Breathing

Significant weight reduction can improve the central and obstructive apneas and hypopneas of sleep-disordered breathing in the morbidly obese (244). CPAP can be effective for patients with primarily obstructive events, but it is not adequate for those who have restricted pulmonary volumes and hypercapnia. Mask discomfort and air leakage into the eyes can make the use of CPAP via CPAP masks intolerable for about 35% of patients (245).

Independently varying inspiratory positive airway pressures (IPAP) and expiratory positive airway pressures (EPAP) with BiPAP machines can be effective for hypercapnic patients. The greater the pulse pressure difference—that is, the difference between the IPAP and EPAP—the greater the inspiratory muscle assistance. Often, 20 to 30 cm H₂O IPAP and minimum EPAP are most effective. Portable volume ventilators are used instead of BiPAP to deliver nasal IPPV for hypercapnic morbidly obese patients who require greater inspiratory muscle assist at higher peak ventilator pressures than can be provided by BiPAP.

Another convenient long-term solution, effective for many obstructive sleep apnea syndrome patients, is the use of an orthodontic splint that brings the mandible and tongue forward, thus, helping to splint open the hypopharynx (246). Uvulopalatopharyngoplasty and mandibular advancement procedures have also been used but are usually ineffective (247–249).

Invasive Ventilatory Support

The use of noninvasive aids can be contraindicated by the presence of any of the following: saliva aspiration causing SpO₂ to remain below 95%, depressed cognitive function, orthopedic conditions interfering with the application of noninvasive IPPV interfaces and exsufflation techniques, pulmonary disease necessitating high FiO₂, or uncontrolled seizures, or substance abuse (245,250). Also, the presence of a nasogastric tube can hamper the fitting of a nasal interface and the use of mouthpiece or nasal IPPV by interfering with both soft palate closure of the pharynx and seal at the nose. Although tracheostomy IPPV can extend survival for patients with neuromuscular ventilatory failure (251), morbidity and mortality outcomes are not as favorable as by noninvasive approaches (195). Tracheostomy is indicated for severe bulbar ALS patients (199), rarely if ever for DMD patients (252), and rarely for SMA patients. It has been found that even SMA type 1 children can have less long-term morbidity when managed without tracheostomy tubes (205). Patients with DMD, even those who are continuously ventilator dependent on noninvasive IPPV, can avoid hospitalizations and pulmonary morbidity and mortality for decades, and tracheotomy indefinitely when properly managed by using respiratory muscle aids (252).

Although widely thought to be fatal before age 2, management of SMA1 according to a recently described noninvasive respiratory aid protocol has thus far allowed many of these patients to survive age 15 without a tracheostomy. Of 47 patients, 32 used high-span BiPAP only during sleep; 6 required it over 16 h/d and become hypercapnic and dyspneic when not using it; and 9 required it continuously with little or no breathing tolerance (253).

Considering the eight deceased NIV patients, three died suddenly at home from bradycardias. One died from intracranial hemorrhage. Two died from mucus plugs during upper respiratory infections when intubation was unsuccessful; one from septic shock; and one from pulmonary embolism. Of the 32 patients who used high-span BiPAP only during sleep, 10 were able to speak clearly, 16 had severe dysarthria, and 6 were aeral. Of the six 16 to 20 h/d users, 5 had severe dysarthria and 1 was aeral. Of the 9 who require continuous support, 1 could speak clearly, 2 had severe dysarthria, and 6 were aeral. Thus, 34 of the 47 NIV users could communicate verbally in contrast to 6 of the 27 with tracheostomy tubes.
When tracheostomy IPPV is used despite the fact that oropharyngeal muscles are sufficient for swallowing, speaking, and permitting decannulation to NIV, either cuffless tubes or tracheostomy cuff deflation should be used up to 24 hours a day (101). Delivered air volumes are increased to compensate leak and support speech and one-way valves used to further facilitate verbal communication (98). Tracheostomy buttons are useful to optimize air passage through the upper airway for autonomous breathing as well as during transition from tracheostomy to noninvasive IPPV (239).

Quality of Life

Misconceptions about the undesirability of "going on a respirator" have far-reaching negative effects for persons now happily being supported on a respirator, and mitigate the positive effects it could have for some types of chronically impaired persons whose quality of life also could be enhanced by the use of a ventilator (254). Although it would seem that intelligent, self-directed individuals should be fully informed about therapeutic options, in the frenzy of seeking a less expensive health care delivery system, some physicians have suggested eliminating the patient from the decision-making process. As recently as 1989, it has been recommended that a physician's assessment of patients' quality of life be done "independent of the patient's feelings" to guide the clinician in whether to institute mechanical ventilation (255).

Poor quality of life is usually given as the reason for withholding ventilator use (256). However, no quality-of-life criteria can be appropriately applied to all individuals. Life satisfaction depends, rather, on personal preferences (257) and on subjective satisfaction in physical, mental, and social situations, even though these may be deficient in some manner. Thus, not quality of life but potential satisfaction with life should be considered. It is particularly appropriate that the life satisfaction of individuals who are living the consequences of having chosen to use ventilators be considered when deciding about such ethically and financially complex matters as ventilator use for others. Interestingly, data indicate that severely disabled, long-term postpoliomyelitis, DMD, and SCI ventilator users (256) generally have very positive views of their lives and life satisfaction. These individuals find quality of life in interpersonal activities, and they are very significantly more satisfied with life than health care professional estimates suggest (257). Crucial for this is often the availability of personal attendant-care services. Thus, in the face of calls to limit entitlement spending, it should be noted that a society willing to provide free room, board, health care, legal and educational services, vocational training, and cable television for felons at exorbitant cost has the ethical responsibility to provide attendant-care services to those in need, some of whom are crime victims themselves. Our specialty needs to do more than instruct the patient and his/her family; it must advocate for them to assure that they have access to appropriate community resources.


CHAPTER 42  |  REHABILITATION OF THE PATIENT WITH RESPIRATORY DYSFUNCTION  


Burn Rehabilitation

INTRODUCTION

Burn injuries pose complex physical and psychological rehabilitation challenges. The incidence of burns has decreased dramatically in the past 50 years as a result of public education and home and work safety efforts. Additionally, survival after burn injury has increased significantly in the same time period (1). Advances in the field that have contributed to survival include the formation of specialized burn centers, early excision and grafting, improved resuscitation and intensive care, and the development of topical and systemic antibiotics (1,2). With dramatic improvements in survival, the focus of burn care is increasingly shifting toward rehabilitation. Burn survivors have complicated rehabilitation needs including scarring, contractures, pain, amputations, neurologic injuries, psychological problems, and community integration issues. The period of rehabilitation may last from months to many years after injury. The physiatrist is an integral member of the burn care team from the time of injury to long-term follow-up.

EPIDEMIOLOGY

It is estimated that 1.25 million people experience burn injuries each year. Of those, approximately 500,000 receive some form of medical treatment and 40,000 are hospitalized (3,4). Burns predominantly affect young men (mode age: 20 to 40; male: 70%). Two thirds of burn injuries affect adults and one-third affect children. Most burns occur by fire/flame (43%) or scald injuries (36%) (Fig. 43-1). Other etiologies that comprise the minority of burns include electrical, contact, chemical, tar, radiation, and grease injuries as well as skin diseases. Approximately one third of burn injuries are associated with concomitant alcohol or drug use. A large majority of burn survivors have less than or equal to a high-school education (82%). Most injuries (65%) are result of an accident that is not work related. A minority of burn injuries (17%) occur at work.

Approximately 5% of burn injuries are the result of child abuse or adult assault or abuse. Among children less than 2 years old, burn injuries represent the most common cause of accidental death; most of these deaths are a result of abuse. Overall, the survival rate is approximately 95%. The risk of death is increased for those at the extremes of age, with inhalation injury and with larger burns (5–7).

PATHOPHYSIOLOGY

Normal Skin

The skin is the largest organ of the body. It serves multiple functions; it acts as a protective barrier to the external environment, regulates temperature and fluid homeostasis, plays a key role in sensation, and contributes to our sense of identity and communication. Skin is a complex organ composed of two layers, the epidermis and dermis (Fig. 43-2).

The outermost epidermis consists of stratified squamous epithelium. This layer contains no blood vessels and the cells are nourished by diffusion from capillaries in the upper layers of the dermis. Cells are formed through mitosis at the basal layer (stratum basale). In addition to forming the bottom layer of the epidermis, the stratum basale lines the hair follicles and sweat glands. (Hair follicles and sweat glands are a source of epidermal cells that enable partial thickness burn injuries to heal spontaneously.) The mitotic daughter cells migrate toward the surface changing shape and composition as they die due to isolation from their blood source. The cytoplasm is released and the protein keratin is inserted. They eventually reach the outer layer and desquamate. This process is called keratinization and takes approximately 30 days. This keratinized layer of skin (stratum corneum) is responsible for keeping water in the body and keeping harmful chemicals and pathogens outside the body, making skin a natural protective barrier.

The dermis resides beneath the epidermis and consists of a vascular connective tissue that supports and provides nutrition to the epidermis and skin appendages. The dermis contains nerve endings, capillaries, lymphatic vessels as well as appendages that include hair follicles, sweat glands, and sebaceous glands. The eccrine sweat glands release heat from the body’s surface through sweat, thereby contributing to thermoregulation. The sebaceous glands secrete an oily substance called sebum that protects the skin and hair and moisturizes the skin.

Classification of Burn Injury

One of the most common classification systems uses depth of injury to categorize the severity of the burn. Superficial injuries, previously termed first-degree burns, solely affect the epidermis. The category of second-degree burns is now divided into superficial and deep partial thickness burns. The former interrupt the epidermis and superficial dermis and present with blistering, moist and painful skin that blanches with
Burn injuries are also classified by size. Lund and Browder (8) diagrams provide a systematic method for calculating total body surface area (TBSA) burned for both adults and children. In contrast to an adult, a child’s head represents proportionally more and their legs represent proportionally less of their body surface area (Fig. 43-3). The rule of nines is used clinically as a quick estimate of TBSA (Fig. 43-4).

Many burn injuries are best treated in specialized burn centers. The American Burn Association in consultation with the American College of Surgeons has developed clinical criteria for referral to a burn center (9) (Table 43-2).

<table>
<thead>
<tr>
<th>Old Classification</th>
<th>New Classification</th>
<th>Appearance/Symptoms</th>
<th>Course/Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>First degree (epidermis)</td>
<td>Superficial thickness</td>
<td>Erythematous, dry, mildly swollen, blanches with pressure, painful</td>
<td>Exfoliation, heals spontaneously in 1 wk, no scarring</td>
</tr>
<tr>
<td>Second degree (dermis)</td>
<td>Superficial partial thickness</td>
<td>Blistering, moist, weeping, blanches with pressure, painful</td>
<td>Reepithelialization in 7–20 d</td>
</tr>
<tr>
<td></td>
<td>Deep partial thickness</td>
<td>No blisters, wet or waxy dry, variable color, less painful, at risk for conversion to full thickness because of marginal blood supply</td>
<td>Reepithelialization weeks to months. Skin grafting may speed recovery. Associated with scarring</td>
</tr>
<tr>
<td>Third degree (all of dermis and epidermis)</td>
<td>Full thickness</td>
<td>White waxy to leathery gray to charred black, insensate to pain, does not blanch to pressure</td>
<td>Reepithelialization does not occur, requires skin grafting, associated with scarring</td>
</tr>
<tr>
<td>Fourth degree (extends to muscle bone, tendon)</td>
<td>N/A</td>
<td>Black (eschar), exposed bones, ligaments, tendons</td>
<td>May require amputation or extensive deep debridement</td>
</tr>
</tbody>
</table>
Effects of Thermal Injury

In thermal injury, the extent of tissue damage is related to the location, duration, and intensity (temperature) of heat exposure. Also of significance, those at the extremes of age have fewer protective layers of epithelium; therefore the same location, duration, and intensity of heat will produce a more serious burn injury in children and elderly than in other adults.

After burn injury a cascade of physiologic processes affect the thermal injuries ultimate impact. There is a complex interplay of local and circulating mediators, including histamine, prostaglandins, thromboxane, kinins, serotonin, catecholamines, oxygen free radicals, platelet aggregation factors, angiotensin II, and vasopression. Initially, there is vasoconstriction at the site of injury mediated by release of norepinephrine and serotonin. A few hours after injury vasoconstriction turns to vasodilation, increased capillary permeability, and leakage of plasma into the extravascular space. Histamine is released. Damaged cells swell. Fluid shifts result in increased extravascular edema and intravascular hypovolemia. Platelets and leukocytes aggregate, leading to thrombotic ischemia (10, 11). In severe burn injuries, inflammatory mediators are released and compromise cardiovascular function. Burn shock ensues, resulting in decreased intravascular volume, increased systemic vascular resistance, decreased cardiac output, end-organ ischemia, and metabolic acidosis. Resuscitation treatment helps reverse this potentially deadly cascade.

Damaged skin results in impairment in most major functions of the integumentary system. In areas of burn injury, skin loses its ability to act as a protective barrier and homeostatic regulator. This may lead to significant losses of body fluid, impaired thermoregulation, and increased susceptibility to infection. In large burns, loss of fluid by evaporation contributes to the development of hypovolemia and shock.

<table>
<thead>
<tr>
<th>Table 43.2 American Burn Association Criteria for Referral to a Burn Center</th>
</tr>
</thead>
<tbody>
<tr>
<td>Partial-thickness burn greater than 10% TBSA</td>
</tr>
<tr>
<td>Burns that involve the face, hands, feet, genitalia, perineum, or major joints</td>
</tr>
<tr>
<td>Any full-thickness burn</td>
</tr>
<tr>
<td>Electrical burns, including lightning injury</td>
</tr>
<tr>
<td>Chemical burns</td>
</tr>
<tr>
<td>Inhalation injury</td>
</tr>
<tr>
<td>Burn injury in patients with preexisting medical disorders that could complicate management, prolong recovery, or affect mortality</td>
</tr>
<tr>
<td>Any patient with burn injury and concomitant trauma in which the burn injury poses the greatest risk of morbidity and mortality</td>
</tr>
<tr>
<td>Burned children in hospital without qualified personnel or equipment for the care of children</td>
</tr>
<tr>
<td>Burn injury in patients who will require special social, emotional, or rehabilitative intervention</td>
</tr>
</tbody>
</table>
Release of catecholemines plays a key role in the development of a catabolic state after burn injury. Tachycardia, increased nutritional demand and weight loss typically ensue. Gastric dilation and gastrointestinal ileus are common in the first few days after burn injury. Also, immune function is impaired. The arachidonic acid and cytokine cascades alter the function of lymphocytes, macrophages, and neutrophils. As a result, patients are at increased risk of infection. Inhalation injury is commonly associated with fire injuries. Noxious gaseous components of smoke directly damage the respiratory tract. Patients are at risk for carbon monoxide intoxication, upper airway edema and obstruction, pneumonia, and dependence on mechanical ventilation (11,12).

Skin Regeneration and Scarring
Spontaneous reepithelialization is impossible with a full-thickness burn injury because of destruction of the dermal appendages. Full-thickness burns result in hair loss, sensory impairment, loss of normal skin lubrication, and heat intolerance because of destruction of sweat glands.

Healing and regeneration of skin in partial-thickness burns arise from the epithelial linings of the hair follicles and sweat glands (stratum basale). Depending on the depth, healing is completed within 14 to 21 days. The new skin again functions as a temperature regulator and a protective barrier against bacteria. After epithelialization there is continued healing with regeneration of the peripheral nerves, sometimes associated with symptoms of neuropathic pain and itching. Although epithelium covers the wound, dermal scarring occurs in the burn wound on a continuous basis for several months after injury. The healing process is ongoing from 6 months to 2 years until the skin is mature. By that point the vascularity of the wound has returned to near normal, and there is no further collagen deposition in the wound.

ACUTE MEDICAL MANAGEMENT

Acute Care of Burn Wounds
While the long-term goals of burn wound care are to restore skin integrity, function, and appearance, the immediate goals post resuscitation are to prevent infection, decrease pain, prepare wounds for grafting, prevent contracture and scarring, and maintain strength and function.

Debridement, the removal of eschar and necrotic tissue, prepares a viable base for wound healing and grafting. Eschar is a composite of coagulum and other tissue debris. Like necrotic tissue, eschar provides an excellent environment for bacterial growth. Since eschar has no microcirculation, bacterial invasion cannot be resolved with systemic antibiotics.

Debridement is performed by several methods. Water immersion, water spray, and wet to dry dressing are examples of mechanical debridement. Commercially available topical enzymes are available for debridement and include substances such as sultains that induce proteolysis, fibrinolysis, and collagenolysis.

There are different types of surgical debridement. Sequential, also called tangential, debridement is the process of removing thin slices of necrotic tissue. Tissue is removed until a viable tissue bed is reached. Fascial debridement surgically removes tissue down to fascia. In this type, a viable wound bed is assured but a significant soft-tissue defect results. Circumferential fascial debridement places patients at high risk for chronic edema.

Deep skin burns are inelastic and the injured skin does not accommodate to the massive edema associated with acute burn injury. In a circumferential burn, the inelastic tissue acts as a tourniquet. The tourniquet effect can lead to a compartment syndrome, defined as compartmental pressure of at least 40 mm Hg. If this occurs, escharotomy is indicated. Escharotomy, surgical decompression of the compartment, is urgently performed to avoid necrosis of the underlying tissues that result from sustained elevated pressure. Escharotomy incisions are performed along the medial and lateral aspects of the extremity. This procedure is critical in situations involving full-thickness circumferential injuries of the chest, arms, or legs. If escharotomy does not successfully reduce the elevated pressure, a fasciotomy is indicated.

Grafting
The introduction of early excision and grafting in the last 40 years has contributed significantly to decreased length of hospitalization as well as improved survival rates, cosmesis, and functional outcome.

Homografts, tissue taken from one’s own species, include cadaveric tissue and human fetal membranes. Heterografts, also known as xenografts, are tissues taken from nonhuman species that are used as human grafts. These biological tissues, most commonly porcine tissues, provide wound closure, modulate metabolic needs, and reduce evaporative fluid. These temporary grafts also act as a mechanical barrier to infection and aid pain relief. Temporary grafts are useful as “test grafts” to determine if the wound bed will accept an autograft. Typically, homografts and heterografts are removed or replaced after several days because the patient’s immune system rejects them. It is not routine to use immunosuppressant therapy in conjunction with grafting.

Synthetic wound dressings are available, including polyvinylchloride, polyurethanes, and other plastic membranes. They are vapor and gas permeable. Such grafts are employed until autografting is feasible or the wound heals. Bilaminate analogs composed of thin sheets of silastic as well as epidermal and dermal components are available. Biobrane and Integra are the two most common biosynthetic dressings.

In the world of skin substitutes and biologic dressings, autografts are the gold standard. The bioengineered substitutes can be used as a temporary covering prior to autografting, or as a neodermal base on which the autograft is placed.

Pham, in a recent evidence-based literature review of bioengineered skin substitutes, concluded that Biobrane, TransCyte, Dermagraft, and allogeneic cultured skin are at least as safe and effective as other wound dressings or allografts for partial thickness wounds. The authors concluded that Integra is best used for smaller burns; evidence suggests that burns of 45% or greater experienced higher infection rates. TransCyte provides good coverage in areas of high contour such as the face (13).
Autografts are harvested from the patient’s own skin. In this process, skin is surgically removed from one’s own body and is relocated to another site (grafted) (Figs. 43-5 and 43-6). Autografts are placed on a wound bed clean of any nonvital tissue or debris and without evidence of infection.

Split thickness grafts are applied in sheets or may be meshed prior to application. Cutting small regularly staggered parallel slits in the sheet of harvested skin creates a meshed skin graft. This expands the size of the graft to several times its original surface area. Meshed grafts are less cosmetic than full-thickness grafts. Meshed grafts heal quickly, and the epithelialization that occurs in the interstices creates a meshlike appearance to the healed skin (Fig. 43-7).

Full-thickness skin grafts are not meshed and result in a more cosmetic appearance. They are typically applied to cosmetically critical areas such as the face, neck, and hands. In addition, full-thickness grafts are used almost exclusively in reconstructive surgery.

**Dysphagia**

Dysphagia is a common problem with large burn injuries and complicates the delivery of adequate nutrition for healing and recovery. Burn induced hypermetabolism increases the caloric needs of the patient. Weakness, inhalation injury, tracheostomy, medication, oral motor dysfunction, and multiple other factors contribute to the development of dysphagia. Tracheostomy is associated with aspiration, pneumonia, and the development of tracheal-esophageal fistulae. Vocal cord paresis is associated with inhalation injury and endotracheal intubation.

Vigilant monitoring is required to recognize dysphagia early and prevent aspiration and associated morbidity. Edelman et al. (14) demonstrated the importance of evaluating patients swallowing function. In this study, dysphagia was initially assessed by bedside swallowing evaluation, and, if abnormal, followed by modified barium swallow examination. They found that oral phase dysphagia is usually due to impaired range of motion (ROM), weak mastication, and impaired oral seal from burn wounds or scar formation. Esophageal dysfunction is a secondary complication of intubation or tracheostomy. Pharyngeal phase dysphagia is usually a result of inhalation injury, complications of tracheostomy, intubation, or burn scar. Pharyngeal dysphagia is associated with the highest risk of aspiration.

Interventions include evaluating endotracheal tube position, size, and location. Selection of appropriate food consistency and positioning, including head and neck positioning, are successful measures to reduce the risk of aspiration.

**REHABILITATION**

Restoration of independent function is the ultimate goal of rehabilitation. Functional restoration includes all aspects of the human life such as strength, ROM, mobility and self-care, reintegration into family and community, adaptive psychosocial responses, and self-determination.

Rehabilitation after a severe burn injury is a multistage process that may take years. Acute rehabilitation goals include interventions to facilitate wound healing, achieve pain
control, prevent joint contracture and weakness, and promote independent mobility and activities of daily living (ADL). Details of the injury, age, premorbid functional level, and health are determinants of an individual’s rehabilitation plan. Therapy is individualized according to burn location, depth, and size as well as other associated injuries or complications. Successful rehabilitation involves multiple disciplines working collaboratively with the patient to achieve the highest level of functioning possible.

**Positioning**

Contractures are a common complication of deep partial and full-thickness burn injuries. Proper positioning is a basic tenet of contracture prevention. Proper positioning also helps prevent other complications such as pressure ulcers and compression neuropathies (Fig. 43-8; Table 43-3).

Contracture prevention is based on the principle of tissue elongation. Patients often prefer to position injured tissue in a shortened, nonstretched state for comfort. Typically, this is a position of flexion and adduction. Such positions of comfort lead to contracture. Positions of extension and abduction are usually indicated to counter the position of comfort. One must prescribe positioning in accord with the location of the injury and direction of the contracture. Joints with overlying deep burns are placed in a position of tissue elongation. For example, a deep burn of the elbow is kept in extension to avoid a flexion contracture that would impede reaching.

Contractures are not limited to joints. Other areas, such as the soft tissue of the lips and mouth, require stretching, exercise, and therapeutic devices to maintain tissue length and function.

**Splinting**

Splints are commonly employed for burn injuries. These devices provide multiple functions including facilitating proper positioning, preventing joint contractures, protecting skin grafts or fragile wounds, or assisting desired motion.

Splints are fabricated from many materials. Low-temperature thermoplastic orthotics are most common. Low-temperature plastics have several advantages including the ability to be warmed at the bedside in heated water and fitted to the patient immediately. These plastics are conformable at low temperatures and therefore can be readily remodeled and adjusted as needed in the clinic or at bedside.

Custom splints can be designed for virtually all parts of the body. Hand injuries commonly require custom splints. Custom splints are advantageous for difficult to fit areas and for sites that require unique design or built-in features, such as facilitating motion at a single joint. They are costly, require experienced staff for construction and maintenance, and necessitate that materials are available on site for fabrication.

Commercially available prefabricated splints may be cost-effective; however, these splints often require modification to fit properly or to achieve the intended purpose. Some clinicians believe that the commercially available prefabricated splints are best used for positioning the knee and ankle. Splints that are simple and straightforward in design and function are “user friendly.” Such splints are more likely to be applied correctly and compliantly. An incorrectly applied splint can lead to further injury, including nerve damage, loss of skin graft, and other skin trauma. A properly designed and fitted splint avoids pressure over bony prominences and is compatible with wound dressings and topical medications. It is often desirable to fabricate the splint with moldable materials that can be modified as a patient’s needs change.

The basic rule of splinting is to splint the body part in a position opposite of the expected deformity. Factors to consider when prescribing a splint include burn size, burn location, burn type, functional goals, and patient activity level.

The wearing schedule for splints is individualized. In the case of the comatose patient, splints should be worn for 2 to 4 hours and then removed for a similar length of time and then reapplied. The wearing schedule can be modified as the patient’s level of participation is increased. Splints may be worn to maintain the gains made in therapy.

If normal ROM of a joint is preserved, a splint is not indicated unless a joint or tendon is exposed or the patient is noncompliant with positioning. Common splints include the knee extension splint to prevent knee flexion contracture and posterior foot drop splint to maintain neutral ankle positioning.

The upper extremity is the most common site for contractures. For axillary burns, an “airplane” splint is used to prevent shoulder adduction contracture. An “airplane” splint holds the upper extremity in approximately 15 degrees of horizontal abduction.

and 90 degrees of abduction. This splint prevents shortening of the anterior and posterior axillary folds. Modifications or inserts are used to maintain the contours of the axillary apex and increase ROM as tolerated (Fig. 43-9). In injuries of other upper-extremity joints, splints are fabricated to meet the specific positioning demands of the elbow, forearm, and wrist.

To correct a contracture various static progressive splints or orthoses are designed to provide a slow progressive sustained stretch. The literature documents success using serial casting to achieve contracture correction as well (15). Staley and Serghiou (16) summarized serial casting use in burn injury:

- Long duration of stretch with minimal force.
- Protection of exposed tendons
- Mechanical forces to remodel scar
- Cost-effective
- Treatment useful in children and noncompliant patients
- Treatment option when an open wound is present.

### TABLE 43.3  Proper Positioning for Preservation of Function in Acute Burns

<table>
<thead>
<tr>
<th>Joint/Body Region</th>
<th>Preferred Position</th>
<th>Positioning Devices</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Upper Extremity and Trunk</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neck</td>
<td>Slight extension</td>
<td>Neck collar</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Splint that conforms to the neck</td>
</tr>
<tr>
<td></td>
<td></td>
<td>No pillows under the head</td>
</tr>
<tr>
<td>Axillae</td>
<td>Horizontal adduction 15 degrees</td>
<td>Airplane splint</td>
</tr>
<tr>
<td></td>
<td>Abduction 80 degrees</td>
<td>Wedge to positioning to abduction</td>
</tr>
<tr>
<td></td>
<td></td>
<td>If rest of upper extremity involved</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Can support UEs:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bedside table</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Side boards/bedside extensions</td>
</tr>
<tr>
<td>Elbow</td>
<td>Extension 5 degrees</td>
<td>Arm trough splint</td>
</tr>
<tr>
<td>Forearm</td>
<td>Supination</td>
<td>Elbow extension splint</td>
</tr>
<tr>
<td>Wrist</td>
<td>Neutral or slight extension</td>
<td>Arm trough</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Wrist cock-up splint</td>
</tr>
<tr>
<td>Hand</td>
<td>I, P joints: full extension</td>
<td>Part of resting hand splint</td>
</tr>
<tr>
<td></td>
<td>MCP: 70–80 degrees flexion</td>
<td>Resting hand splint</td>
</tr>
<tr>
<td></td>
<td>Abducted from palm</td>
<td>Soft web spacers</td>
</tr>
<tr>
<td></td>
<td>Thumb opposition</td>
<td>Intrinsic plus hand splint</td>
</tr>
<tr>
<td></td>
<td></td>
<td>C bar for thumb</td>
</tr>
<tr>
<td>Chest and trunk</td>
<td>Neutral with level hips</td>
<td>Figure of eight device to reduce protraction</td>
</tr>
<tr>
<td><strong>Lower Extremity and Trunk</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hip</td>
<td>Neutral extension</td>
<td>Wide soft straps to avoid frog leg position especially in children</td>
</tr>
<tr>
<td></td>
<td>Abduction 20 degrees</td>
<td></td>
</tr>
<tr>
<td>Knee</td>
<td>Extension</td>
<td>Knee extension splint; immobilizer</td>
</tr>
<tr>
<td>Ankle</td>
<td>90 degrees</td>
<td>Posterior shell with ankle in neutral</td>
</tr>
<tr>
<td></td>
<td>That is neutral</td>
<td>L/Nard; PRAFO-like devices</td>
</tr>
<tr>
<td></td>
<td>Dorsiflexion, plantar flexion</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Inversion/eversion</td>
<td></td>
</tr>
<tr>
<td>Foot</td>
<td>Neutral forefoot</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Supination/pronation; toes extended</td>
<td></td>
</tr>
</tbody>
</table>

**FIGURE 43-9.** Airplane splint fabricated to prevent contracture development of the shoulder.
Splints are also valuable in the postacute period to prevent contractures until the tissue length is stabilized, after surgical release of a contracture, or after skin grafting.

**Hands**

Hands require special attention. When evaluating the burned hand, individual joint motion should be assessed. Limitations in ROM results from decreased tendon sliding, decreased muscle strength, tendon shortening, skin/muscle/ligament/tendon tightness, joint restriction, or a combination of these processes.

During the acute phase, the hands are positioned and splinted to prevent shortening of the joint capsules, collateral ligaments, and muscle tendons. Edema can complicate care as it accentuates metacarpophalangeal (MCP) joint extension and interphalangeal (IP) joint flexion. The resulting combination of MCP hyperextension and IP joint flexion produces the intrinsic minus deformity of the hand, also called claw hand.

The hand is splinted with the wrist in slight extension, the MCP joints in 70 to 80 degrees flexion, IP joints in extension, and the thumb abducted from the palm. Fingers are positioned in abduction (Fig. 43-10). If the burns are limited to the palmar aspect, then the MCP and IP joints are positioned in extension, fingers abducted, and the thumb abducted from the palm. Palmar splints may need modifications to maintain the palmar arch. Elastic wraps or straps are employed to secure the splint. Elastic wraps are applied in a figure of eight formation to avoid circumferential constriction. Straps should be soft and applied in a crisscross orientation.

Circumferential hand burns often damage the web spaces of the hand. Preservation of the web spaces is important for hand rehabilitation. For example, loss or shortening of the first web impairs thumb opposition and abduction and interferes with grasp (17). Early intervention is critical. In addition to abduction, exercise to all the digits, web space dressings, and soft inserts or straps are placed in the web spaces. Also, thermoplastic inserts may be customized and fitted in these areas. Compression gloves may be used in conjunction with the splints. In mild cases, compression gloves may be sufficient to preserve web spaces and function.

Exposed tendons require splinting in a slack position. The tendon should be kept moist to avoid desiccation and denaturation. With time, the tendon may revascularize and become functional. If the exposed extensor hood of the fingers is not kept slack, the central slip can fail and lead to a boutonniere deformity. In cases of extensor hood rupture, the finger is positioned in extension. In approximately 6 weeks, scar tissue may form to bridge the extensor surface and act as a functional substitute. After that time, active ROM can be initiated (18).

For exposed joints, a splint is prescribed to provide protection. Gentle ROM is performed to maintain motion. If the joint capsule is violated, the joint is at risk for ankylosis. In such cases, the joint is best placed in a function position.

Edema can cause a claw hand. After 12 to 18 hours, edema changes consistency to a gelatinous substance secondary to lymphatic vessel occlusion and impaired fluid clearance. This gel-like quality makes the edema more difficult to resolve. In the acute stage after injury, the upper extremity and hand are typically suspended above the patient’s head to aid fluid return and reduce edema. Care is taken not to stretch the brachial plexus or damage the ulnar nerve.

The resting hand splint, with the inclusion of wrist extension, contributes to edema control of the hand (see Fig. 43-10). This splint stretches the normally redundant dorsal skin, thereby limiting space for edema collection. Edema formation coupled with the flexor tendon pull increases flexion at the IP joints resulting in a claw hand deformity.

Once skin closure is achieved, a compression glove is prescribed to assist with edema control (Fig. 43-11). Passive exercise and scar suppression are also employed to assure restoration and maintenance of normal ROM.

**Exercise**

In burn rehabilitation, one of the earliest goals of exercise is to maintain or achieve normal ROM. For the obtunded or comatose patient, passive ROM exercises emphasizing the end ROM is appropriate. Alert and cooperative patients can participate in active and active-assisted exercise. With children, developmentally appropriate exercise and play activities are indicated to achieve the therapy goals.

Surgical anesthesia can provide an opportunity to perform ROM exercises and determine objective measurements of

---

**FIGURE 43-10.** Resting hand splint.

**FIGURE 43-11.** Compression gloves for hypertrophic scar suppression.
range. In this setting, one can perform ROM without concern for pain. The opportunity to assess ROM under anesthesia is valuable in the case of a child, an uncooperative patient, or if pain is inhibiting ROM exercise.

Stretching exercises are prescribed when ROM is abnormal. The biomechanics of skin and muscle are different; therefore, the two tissues are stretched differently. Stretching of injured skin or scar tissue requires a slow sustained mechanical stretch to enhance elongation of collagen and underlying fibers. When a prolonged stretch is performed, the stretch is maintained until the tissue blanches. Blanching indicates that the tissue is near its yield point, the point at which the skin is at risk of tearing (19,20).

After burn injury, muscular weakness, fatigue, and deconditioning are serious problems. These sequelae interfere with function, such as ambulation, ADL skills, and endurance. Loss of muscle mass and, in children, the additional loss of bone mass, interfere with the restoration of function, return to work or school (21,22).

Literature indicates that a structured exercise program composed of aerobic and resistance training leads to increased function as measured by increased muscle mass, strength, and cardiovascular endurance (23). Suman and Herndon (24) demonstrated the efficacy of a supervised exercise program for children, ages 7 to 18 years, that included resistance and aerobic exercise. His team trained children in eight basic exercises, including bench press, leg press, and triceps curls under the supervision of a physical therapist in an outpatient setting. The control group was instructed in a home exercise program. The participants of the supervised resistance and aerobic program required significantly fewer surgical releases up to 2 years after the intervention (25). Suman et al. (26) also demonstrated that children had increased lean body mass and muscle strength with structured exercise and the concurrent administration of the anabolic steroid, oxandrolone, for 1 year after injury. Anabolic agents, such as oxandrolone and human growth hormone, have been demonstrated to reduce the effects of hypermetabolism while also increasing muscle mass and strength and decreasing resting energy expenditure in children and adults (27–31).

St-Pierre et al. (32) demonstrated that individuals with small burns did not differ from those without injury in muscle strength. However, for those with greater than 30% TBSA burn produced less torque, work, and power in their quadriceps when compared to matched controls. de Lateur et al. (33) reported that a structured aerobic exercise program, consisting of three times a week aerobic treadmill training for 12 weeks combined with a standard burn therapy program, achieved improved aerobic capacity. This literature suggests that regular exercise after burn injury, like in other adults, results in improved flexibility, endurance, balance, and strength. Such gains are important for returning to full independence and function. Other likely benefits are reduced anxiety and an improved sense of well-being (34).

**Gait**

Independent walking may be the single most important factor in determining discharge disposition after severe burn injury. Farrell et al. (35) found that independent ambulation predicted discharge from the acute burn unit to home.

As soon as the patient’s condition permits, ambulation should begin. In addition to reducing the risk for contractures, deconditioning and deep venous thromboses, early gait training maintains balance, lower extremity ROM, strength, and endurance. Physical therapists use a tilt table to perform graduated upright positioning. Literature indicates that muscles of the lower extremities, particularly the gastrocnemius and quadriceps, are the first muscles to lose strength with bed rest. These same muscles experience a greater proportional loss of muscular torque when compared to other skeletal muscles (36,37).

The physical therapist can initiate gait training and determine appropriate assistive devices; however, physical therapy sessions should not be the only time the patient walks. Nurses, family, and others can learn how to assist and provide needed guarding during walking. Walking reduces the effects of bed rest and improves aerobic conditioning.

Many consider new autografts to the lower extremities a contraindication to ambulation. Typically, ambulation is not initiated until the surgeon is confident that the graft will tolerate a dependent position. At 5 to 7 days after grafting, patients are instructed to begin lower extremity–dependent positioning. Dangling the lower extremities is a preambulation exercise that helps determine if the graft tolerates the dependent position. The application of elastic wraps or other elastic devices are used to minimize venous pooling and decrease the risk of graft loss. The typical protocol for dangling begins for 5 minutes two to three times per day. If inspection of the graft after dangling shows no signs of intolerance, the duration is progressively increased. Once the graft shows tolerance for dependency, ambulation is initiated. Supportive dressings such as ace wraps are often employed. Like dangling, the time spent walking is methodically increased as the graft tolerates.

Gait deviations are common after burn injury. Some are transient while others persist. Early correction of abnormal stance and gait reduces the risk that the deviation will be long-standing. Deviations result from pain, location of injury, deconditioning, weakness, contractures, and sensory and central nervous system dysfunction. Gait devices are commonly used to protect injured areas, reduce pain, or assist with weight bearing. Devices are also employed to correct or prevent poor posture or gait deviations secondary to injury to the trunk or other areas.

During times that the patient is unable to ambulate, wheelchairs provide mobility and are easily adapted to the patient’s specific needs. Adaptations include attaching splints or wedges to elevate the arm or adding lower-extremity positioning devices.

**Surgical Reconstruction**

The surgical goal is to minimize surgery and maximize results. If multiple areas are involved, an overall plan and timetable are
PART III | MAJOR CONDITIONS

developed that prioritize the fewest number of surgeries with the maximal functional benefit to the patient. The timing of surgical release of a contracture varies. Some surgeons believe that surgery should wait 6 to 12 months after injury. However, there is limited evidence to support this recommendation. Other literature supports early intervention (38–40). There is evidence that early release does not worsen outcome, and surgery is indicated with the development of secondary deformities (41). Tendon and ligament injury impacts the completeness of the release that can be performed. Age, severity of the deformity, and time since injury impact surgical outcome (39,42).

Many reconstructive procedures are utilized in the burn-injured patient. Listed in the order of complexity are Z-plasty, skin graft, local skin flaps, local muscle flaps, fasciocutaneous flaps, free flaps, and cross limb flaps (38).

Simple excision is indicated for scars that are limited in size and location. Z-plasties are employed to limit or correct joint contractures caused by hypertrophic scarring. Contractures of soft tissue that disrupts appearance or tissue flexibility are also treated with Z-plasty. For example, this technique is employed to correct hypertrophic scarring over the chin that prevents adequate mouth closure (43). Release and skin grafting typically involves a fish-mouth incision and graft placement. Scar pressure therapy and silicone application are often combined with surgery to maximize outcome.

Severe axillary contractures are difficult to treat (Fig. 43-12). For tight bands formed near unburned skin, five flap releases are used (44). Local flaps are often used if anterior or posterior axillary fold contractures are present (39). Contractures that are broader and involve the entire axilla require fasciocutaneous or similar flap procedures (45).

Contractures at the elbow are often complicated by heterotopic ossification (HO) and this is taken into consideration during the surgical plan (46). A Z-plasty or five flap release is used to treat a thin band of scarring (45). Surgical intervention for hand contractures requires expertise and experience. Full-thickness grafts are commonly used. Correction of flexion contracture of the MCP joint has a higher success rate than an extension deformity. K wires are placed for 2 to 4 weeks after surgery for joint stabilization (47). Like the MCP joint, extension contractures of thumb exhibit poor outcomes. Contractures associated with dislocated or subluxed joints have the worst results (45,48,49).

In recent years, face transplantation has garnered significant interest in the lay press. The first reported human face transplantation was performed in 2005 (50). To date, there are only a few case reports in the literature (50–52). These investigators utilize a cadaveric facial allograft that requires lifelong immunosuppressive treatment. There are unresolved questions regarding ethical, immunological, and psychological issues. Does the improvement in quality of life justify the long-term risks of immunosuppression, including cancer, infection, and renal toxicity? Will the recipient experience graft rejection? What are the psychological consequences of a transplanted face? These and other issues are only beginning to be addressed in the literature (53–55).

COMPLICATIONS

Burn survivors experience a wide range of complications that include neurologic, orthopedic, dermatologic, metabolic, pain, and psychosocial problems. These complications may develop in few months to years after injury. The physiatrist plays an instrumental role in managing these problems, both on the wards and in the clinic.

NEUROLOGIC INJURIES

Localized Neuropathies

Peripheral mononeuropathies and plexopathies are common in severely burned patients. However, this complication is underreported in the literature, as the diagnosis is often delayed or missed entirely. The neurologic assessment is complicated by the complexity of medical problems and impaired consciousness of the critically ill patient. The reported incidence of peripheral neuropathy ranges from 11% to 30%. Kowalske et al. (56) examined 572 burn survivors and found that electrical injury, history of alcohol abuse, and length of intensive care unit stay were significant risk factors for the development of mononeuropathies. It is also felt that elderly and diabetic patients are predisposed to peripheral nerve compromise (57,58). Compression and stretch of peripheral nerves places them at risk for injury. Bulky dressings can cause compression to superficial peripheral nerves. Improper and prolonged positioning is also a risk factor. Clinical pearls of specific mononeuropathies and brachial plexopathy are reviewed below (Table 43-4).

Several bed and intraoperative positions may put the brachial plexus at risk for injury. Brachial plexus stretch injury likely occurs in patients positioned supine with shoulders abducted at least 90 degrees and externally rotated (59). This position may be used in the surgical suite for grafting of the axilla or lateral chest wall. This position may also be implemented to decrease arm edema or prevent an axillary
contracture. To prevent compression or stretch injury to the brachial plexus, it is recommended to position the patient supine with 30 degrees of shoulder horizontal adduction (58).

In the upper extremity, mononeuropathies include the ulnar, median, and radial nerves. The primary area of involvement of the ulnar nerve in the burn patient is at the elbow in the cubital tunnel. The position of elbow flexion and forearm pronation stretches the nerve and places it at risk for compression at this point. Most injuries to the median nerve occur at the wrist level and may be caused by edema and prolonged or repeated hyperextension of the wrist that compresses the nerve at the carpal tunnel. Sustained stretch to the wrist in a hyperextended position, either with splints or in exercise programs, should be performed with caution. The radial nerve is most commonly compressed at the spiral groove of the humerus. Injuries at this level often result from arm resting on the siderails or hanging over the edge of the operating table. The superficial cutaneous branch of the radial nerve may be compressed by restraints at the wrist. This pure sensory neuropathy is a less frequent injury.

In the lower extremity, mononeuropathies may include the peroneal and femoral nerves. Peroneal nerve injuries are relatively common. Stretch injuries are associated with the frog leg position, defined as an externally rotated and flexed hip, flexed knee, and inverted foot. This position is often the result of a short bed, or tender medial thigh, or perineum burns. Compression injuries of the peroneal nerve at the fibular head are associated with the lateral decubitus position, metal stirrups, leg straps, and heavy bulky dressings. Windowing of dressings over the fibular head helps relieve pressure. Femoral nerve injury is uncommon. Injury may occur in the femoral triangle by compression from a hematoma caused by venous or arterial blood draw. Retroperitoneal hemorrhage should be considered in patients with a new femoral nerve injury who are on anticoagulation or have had recent abdominal surgery (60).

The pneumatic tourniquet used to establish a bloodless field in the operating room may cause nerve injury. Improper tourniquet inflation pressure can cause a direct pressure injury of the nerve at the cuff edge (61). The radial nerve is most vulnerable, but ulnar and median nerves are also at risk.

### Peripheral Polyneuropathy

Generalized peripheral polyneuropathy is a similarly common neurologic disorder in burn injury. The incidence ranges from 15% to 30% (57,62,63). Kowalske et al. (56) found that age and length of intensive care unit stay are risk factors for developing polyneuropathy. Polyneuropathy is more commonly seen in those with greater than 20% TBSA burns and electrical injuries (64–66). The etiology of peripheral neuropathies is uncertain; however, metabolic complications and neurotoxic drugs have been implicated. Electrophysiologic evidence of polyneuropathy is seen within 1 week of severe burn injury (67). The patient may have symptoms of paresthesia and signs of mild to moderate weakness in the muscles of the distal extremities. On manual muscle testing, most patients eventually recover their strength, although they may complain of easy fatigability for years after the burn (57,62,63). Critical illness neuropathy is not explicitly documented in the burn literature. Clinically, severely burned patients who commonly experience prolonged intensive care unit stays, sepsis, and multiple organ failure are at risk for critical illness neuropathy.

### Mononeuritis Multiplex

Mononeuritis multiplex is an asymmetric sensory and motor peripheral neuropathy that involves two or more isolated peripheral nerves. The pathophysiology is not well understood, but is thought to result from a combination of circulating neurotoxins, metabolic factors, and mechanical compression. Multiple mononeuropathy was documented in 7 of 121 subjects with greater than 40% burns in one study (68). In a separate study, mononeuritis multiplex was the most common diagnosis in burn patients with a neuropathy (65). At 1 year after injury, lower-extremity nerve lesions demonstrated better functional recovery than upper-extremity nerve lesions (68).

### Pruritis

Itch is a significant complaint for many patients. The prevalence of pruritis is as high as 87% at 3 months and 70% at 1 year after injury (69,70). The mechanism of pruritis is not well understood. Some investigators believe that it is related to axonal sprouting in the dermis. Predictors of pruritis include deep dermal injury, extent of burn, and early posttraumatic stress symptoms (70,71). Various treatment regimens have demonstrated a decrease in reported itch symptoms; however, such interventions lack strong empirical evidence. Nonetheless, there exist multiple clinical treatment options. Topical moisturizers (72) and scar massage (73) are used in clinical practice on healed burn wounds. Topical medications include antihistamines and prudoxin, a tricyclic antidepressant. Oral medication options also include antihistamines (74) and prudoxin (DPT Laboratories, San Antonio, TX).
There are reports of the use of transcutaneous electrical nerve stimulation (TENS) and topical anesthetics in the treatment of pruritis (75,76). For those with severe itching, often a combination of interventions is needed to control symptoms.

**BONE AND JOINT CHANGES**

**Contractures**

Contractures are defined as an inability to perform full ROM of a joint. They result from a combination of possible factors—limb positioning, duration of immobilization and muscle, soft tissue, and bony pathology. Individuals with burn injuries are at risk for developing contractures. Burn patients are often immobilized, both globally, as a result of critical illness in the severely burned, and focally, as a result of the burn itself due to pain, splinting, and positioning. Burns, by definition, damage the skin and may also involve damage to the underlying soft tissue, muscle, and bone. All of these factors contribute to contracture formation in burn injury.

Contractures place patients at risk for additional medical problems and functional deficits. Contractures interfere with skin and graft healing. Functionally, contractures of the lower extremities interfere with transfers, seating, and ambulation. Contractures of the upper extremities may affect ADL, such as grooming, dressing, eating, and bathing, as well as fine motor tasks.

Approximately 40% of hospitalized burn patients develop contractures. The shoulder, elbow, and hand are the most commonly involved joints (Fig. 43-13). Those with more extensive burns, amputations, and inhalation injuries are more likely to develop severe contractures (77). Positioning and ROM exercises are the mainstays of contracture prevention (18,78). Positions of comfort often include joint flexion and adduction and lead to contracture formation. For the bedbound burn patient, the ideal position to prevent contractures involves neck extension, shoulder abduction and external rotation, elbow extension and supination, hip abduction, and ankle dorsiflexion (see Fig. 43-8). Such positioning is coupled with regular ROM exercises. Once a patient has developed a contracture, treatment usually begins with conservative measures, including splinting (79–86) and serial casting (15,16,87). Richard et al. studied 52 patients with burn contractures, comparing treatment with a multimodal approach (massage, exercise, pressure) to a progressive approach (splinting and serial casting). Contractures corrected in less than half the time in the progressive compared to the multimodal treatment group (83). In another study, Bennett et al. performed a case series of 15 subjects, with 35 contractures treated with serial casting for a mean of 13 days. They found a mean increase in ROM from 56% of normal before casting to 86% of normal after casting (15).

Some investigators have reported success using ultrasound (88) and silicone gel (89,90) in treating contractures. Surgical correction of contractures (41,91) is reserved for contractures that significantly impact one’s function and are not improved by conservative measures.
Bone Growth
Significant growth delays occur in children after a severe burn injury (92). Growth disturbances in pediatric burn survivors may result from premature fusion of the epiphyseal plate of affected long bones. Bone growth issues should be considered in growing children with burn scars that cross a joint and with joint contractures. Partial epiphyseal plate fusion may cause bone deviation and deformity (93,94). In addition, case reports document that pressure garment for the treatment of facial burn injury in children alters facial bone growth. It is recommended to monitor closely facial development during and after pressure garment use in children for the development of normal dental and facial proportions (95,96).

Children with burn greater than 15% TBSA exhibit decreased bone mineral density early (at 8 weeks) after injury and the loss is sustained (5 years after injury) (22). The mechanism for loss of bone mass is still being worked out; however, recent research demonstrates causal roles for multiple factors including increase in endogenous glucocorticoids, resorptive cytokines from the systemic inflammatory response and vitamin D deficiency and disruption of calcium metabolism. Reduced bone density places children at risk for long bone fractures (97–99). Mayes et al. examined 104 burned children with greater than 40% TBSA and found a 5.8% incidence of fracture (100). Investigators have studied the use of recombinant human growth hormone without proven effect on bone formation (101). Recent studies have demonstrated improved bone mineral density with bisphosphonate therapy. Klein et al. performed a randomized controlled trial of 43 children with greater than 40% TBSA and examined the effects of acute administration (within 10 days of injury) of intravenous pamidronate. Subjects receiving pamidronate demonstrated higher whole body and lumbar spine bone mineral content at discharge, 6 months and 2 years compared to controls (102,103).

Osteophytes
Evans and Smith report that osteophytes are the most frequently observed skeletal alteration in adult burn patients. They are most often seen at the elbow and occur along the articular margins of the olecranon or coronoid process (104,105).

Heterotopic Ossification
HO is the abnormal formation of bone in soft tissue. The incidence of HO is estimated at 1% to 2% of hospitalized burn patients (106–108). Only those with symptomatic joints, impaired ROM, joint pain, or other symptoms, underwent radiographic examination. Therefore, reports in the literature reflect the incidence of clinically significant HO, not the true incidence. The etiology of this process is unknown. The elbow is the most frequent joint affected, comprising greater than 90% of cases in a 21-year review (106). Risk factors associated with the development of HO include size of burn, ventilator support, intensive care unit stay, prolonged wound closure, wound infection, and graft loss (106,109).

HO may occur as early as 5 weeks but usually develops at around 3 months after injury. One of the earliest signs of HO is loss of joint ROM. Other clinical findings may include swelling, erythema, pain, and peripheral nerve injury. Symptoms may precede radiologic findings. A bone scan is the most sensitive diagnostic imaging test and may demonstrate positive findings up to 3 weeks prior to positive radiographic findings.

Treatment of HO begins with conservative measures, including positioning and ROM to prevent worsening of joint motion. There is no evidence in the literature to support HO prophylaxis for burn patients. Timely surgical intervention is indicated when HO results in nerve entrapment. Surgical treatment is also indicated for HO that causes significant functional deficits, including impairments in upper- and lower-extremity function, impaired mobility and ADL. In these cases, it is common practice for surgeons to wait until the bone is mature to operate, which can take 12 to 18 months. One can follow the HO with serial radiographs every few months to monitor for bone stabilization. Surgical excision of HO at the elbow results in improvement in ROM (110,111). Tsionos et al. performed HO in 28 subjects and 35 elbows at a mean of 12 months after injury. At a mean follow-up of 21 months, the flexion/extension improved from 22 degrees preoperatively to 123 degrees postoperatively (110). In a separate study of eight children with elbow HO, at 17 months after surgery all subjects demonstrated improved ROM and all were able to reach their face and perineum for functional tasks of feeding and toileting (111).

Bony Changes in Electrical Burns
New bone formation is found at residual limb long bones in electrical burn survivors with amputation. Bony changes in the electrical burn include bone splitting, bony necrosis, bone swelling, lucent holes in bone, and periosteal new bone formation (112). Helm et al. reviewed 61 amputation sites in 43 burn survivors with electrical injuries. Twenty-three of twenty-eight patients with long-bone amputations developed new bone formation at the amputation site. No new bone growth was evidenced in the 15 subjects with small bone amputations and disarticulations. The average time from amputation to diagnosis of new bone was 38 weeks. Of those with new bone formation, 12% required surgical revision of the stump and 7% required replacement of their prosthesis secondary to new bone formation. The etiology of new bone formation in the electrical amputee is unknown (113).

Scoliosis and Kyphosis
Asymmetric burns of the trunk, hips, and shoulder girdle can cause the patient to favor the affected side. In the growing child, the contracture of burn scars and resultant postural change can result in structural scoliosis. Investigators have reported a case series of four children scalded on the back as infants that developed adolescent scoliosis. The deformities of all four cases were corrected surgically with good results (114). Similarly, childhood burns on the anterior neck, shoulders, and chest wall may produce a rounding of the shoulders and
sunken chest. Likewise, burn scar shortening and protective posturing can result in kyphosis. Both scoliosis and kyphosis are amenable to bracing and surgical interventions. An orthopedic surgeon should follow such survivors.

**Septic Arthritis**

Septic arthritis is challenging to diagnose in the severely burned patient. The characteristic signs and symptoms are often absent or masked by the overlying burn wound. Joint pain, swelling, color change, and tenderness are common symptoms at the site of burn injury or grafting and therefore are difficult to distinguish from septic arthritis.

The two major causes of a septic joint are penetrating burns into a joint and hematogenous seeding in cases of bacteremia. Burn patients are at risk for infection because of their impaired immune system and concurrent illness. Septic arthritis may cause gross dislocation because of capsular laxity or cartilage and bone destruction (115), or it may result in severe restriction of movement or ankylosis. It occurs most frequently in the joints of the hands, hips, knees, and wrists.

**Subluxations and Dislocations**

Joint subluxation of the hands and feet are common after burn injury. Burns of the dorsal surface may contract resulting in joint hyperextension. With prolonged hyperextension, the joint can sublux. This is most commonly seen in the MCP and metatarsophalangeal (MTP) joints. Ulnar neuropathy places the patient at additional risk for subluxation of the fourth and fifth digits. For dorsal hand burns, prevention of subluxation is achieved with a combination of splinting and ROM exercises. A dorsal hand burn splint places the MCP joints in 60 to 90 degrees of flexion and the distal interphalangeal (DIP) and proximal interphalangeal (PIP) joints in full extension. Similarly, the MTP joints may become subluxed after contracture of healed wounds, especially in children. Application of surgical high-top shoes with a metatarsal bar keeps the toes in an antideformity position.

Posterior hip dislocation can be a problem in children. Hips maintained in an adducted and flexed position are at risk for dislocation. Anterior shoulder dislocations occur in positions of abduction and extension. Shoulder dislocations may result from positioning in the operating room (116).

**PAIN COMPLICATIONS**

Pain management is integral to the care of the burn survivor. In the acute stage after partial-thickness burn, some nerve endings in the dermis remain intact resulting in significant pain at the site of injury. In contrast, after full-thickness burn the nerve endings are completely destroyed and the burned area is less painful or painless. Treatment of burns with daily dressing changes that debride necrotic tissue causes significant procedural pain. Patients may report unbearable levels of pain (117). It is important to manage the constant background pain of the burn injury itself as well as the intermittent procedural pain.

It is well documented that hospitalized burn patients’ experience of pain varies among patients and across patients over time (118–120). Therefore, it is important to tailor the pain management plan for each individual. A few guiding principles are accepted as good clinical practice. Background pain can be treated with a continuous infusion of opioids, a long-acting oral opioid, or patient-controlled analgesia. Procedural pain is well managed with short-acting opioid medications scheduled prior to the procedure (121,122). In a randomized controlled trial of 79 hospitalized burn survivors, Patterson et al. (123) found that adding lorazepam to the standard opioid pain medication significantly reduced pain levels during procedures for those with high baseline pain.

Nonpharmacological interventions are important adjunctive interventions to standard pain medication regimens. This is an area of increasing interest in recent years. Simple environmental modifications and consistency are helpful. The goal is to create a calm atmosphere that incorporates as much patient autonomy as possible. Patients should be encouraged to direct their own dressing care. Some find the presence of family members and music helpful. The treatment team should provide consistent timing, staff, and routine of procedures. Investigators have found some success in the use of hypnosis to decrease procedural pain in burn injury (124,125). Patterson and Ptacek performed a randomized controlled trial of 61 hospitalized severe burn patients, comparing hypnosis with a control intervention that includes attention, information, and brief relaxation instructions prior to dressing changes. For those with high baseline pain levels, the hypnosis group reported less posttreatment pain (124). Virtual reality has demonstrated success as a distraction technique to help assuage pain during wound care (126–128). Sharar et al. performed a randomized controlled within-subject trial of 146 treatment comparisons among 88 subjects. Compared with standard analgesic therapy alone, the addition of virtual reality distraction resulted in significant reductions in subjective pain ratings, including intensity (20% reduction), unpleasantness (26%), and time spent thinking about pain (37%) (128). Other nonpharmacological interventions that are validated in the literature include massage therapy for both acute and chronic pain in burn patients (73,129).

Neuropathic pain after burn injury is not well categorized in the literature. Neuropathic pain is defined as pain initiated or caused by a primary lesion or dysfunction in the peripheral or central nervous system. Neuropathic pain symptoms consisting of pins and needles, burning, stabbing, shooting, or electric sensations are common complaints of burn patients following healing of their open wounds. In two studies examining a total of 534 burn patients at least 1 year after injury, 71% and 82% reported paresthetic sensations in their burn scar, respectively. These sensations were associated with burn size and skin grafting (130,131). Schneider et al. retrospectively reviewed 72 patients with complaints of neuropathic pain and characterized their clinical course. Neuropathic pain symptoms first occurred at a mean of 4 months and persisted until 13 months after injury. Documented initial pain severity
score was 7 out of 10. Typical exacerbating factors included temperature change, dependent position, light touch, and weight-bearing activities. Common alleviating factors included rest, massage, compression garment use, and elevation (132). Common treatment regimens include gabapentin, opioids, and steroid injections into areas of symptomatic hypertrophic scar. Two case reports discuss successful treatments of neuropathic pain, one using methadone and the other an implanted peripheral nerve stimulator (133,134).

SKIN COMPLICATIONS

Hypertrophic Scarring

Hypertrophic scarring can result from deep partial- and full-thickness burns. The scarring first presents as a firm, red area of healed burn scar. It progresses over weeks to become raised, erythematous and rigid (Fig. 43-14). Scars can contract and if present over a joint, can contribute to contracture formation. Over time scars mature, taking on a pale, more pliable and less thick appearance. This process may take up to 2 years. Scarring can result in significant impairments in function. In addition to physical impairments, hypertrophic scarring may lead to psychosocial consequences. Scarring can impact self-esteem, social isolation, body image, and community reintegration (135,136). It is one of the most significant long-term complications of burn injury.

Histologically, scars demonstrate a whorled collage pattern, in contrast to the parallel array of collagen fibers in normal skin. Scar tissue exhibits a proliferation of fibroblasts and capillaries, thickened epidermis, and a lack of rete pegs. Endothelial cell proliferation results in occlusion of the microvasculature leading to a local hypoxia. With maturation, the microvasculature degenerates and capillaries are reabsorbed (137–139). There exists a broad range of clinical presentations of scars that include varied thickness, color, rigidity, and corresponding symptoms. The prevalence of hypertrophic scarring is estimated at greater than 60% of white and greater than 75% of nonwhite survivors with severe burns (140). Interestingly, scarring is not documented in neonatal, elderly, and morbidly obese burn survivors. Prevalence data is confounded by lack of a standardized objective measure of hypertrophic scarring. The Vancouver Scar Scale is the most common method of measurement; however, it is comprised of subjective ratings of pigmentation, vascularity, and pliability and as a result has a relatively low interrater reliability (141). Risk factors for the development of hypertrophic scarring include open wounds for greater than 2 to 3 weeks, burns that require grafting, and heavily pigmented individuals (142). The etiology of hypertrophic scarring is largely unknown.

There exists little evidence in the literature to support specific treatment strategies for hypertrophic scarring. Current studies limitations include sample size, randomization, adequate follow-up, and objective outcome measures. Given the seriousness of the problem, investigators are actively pursuing research in this area (143). Meanwhile, providers rely on clinically derived treatment approaches. Management begins with early identification of hypertrophic scarring and those at risk of scarring. Initial conservative treatment measures include splinting, scar massage, exercise, and positioning. Pressure therapy

FIGURE 43-14. Hypertrophic scarring of the hand. Note the raised and rigid appearance of the scar.
constitutes the main treatment intervention. It is postulated that pressure (at least 25 mm Hg) inhibits capillary blood flow resulting in local ischemia. A decrement in tissue metabolism leads to impaired fibroelastic activity and enhanced collagenase activity. Apoptosis ensues with release of proteases and liposomal contents (43,144).

The goal of therapy is to arrest scar development and flatten existing scars. Initially, pressure wrappings are applied around the affected areas using plastic elastic (ACE), cotton elastic (tubigrip) or adhesive elastic (coban) bandages. As edema resolves the scarred area assumes a more stable shape and custom-made pressure garments are fit. It is recommended to wear compression garments 23 hours of the day until the scar matures, which can be up to 2 years (145). Compliance with this regimen is difficult. The garments are warm, socially awkward, and difficult to don. They stretch and shift with wear, and replacement garments are recommended every 3 months. Concave areas are poorly compressed. To improve fit or pressure over areas with challenging contour features, such as the digital web space, silicone sheets or gels, sponges, or conformed pieces of plastic are used to improve the delivery of pressure and limit shear over the new and fragile epithelium (43,146,147).

Chang et al. performed a randomized controlled trial comparing pressure garment therapy to no pressure garment therapy in 122 consecutive patients with hypertrophic scarring. No significant differences in Vancouver Burn Scar Assessment Scale were found between the two groups (148). In another randomized controlled trial, investigators compared high-pressure (elastic nylon) to low-pressure (elastic cotton) treatment of hypertrophic scarring and found no significant difference in scar outcome (149). Both studies utilized subjective outcome measures to determine the number of days to scar maturation or the number of days for pressure therapy. The efficacy of pressure therapy is unknown; however, there are no major treatment alternatives. Future research in this area will greatly advance the care of burn survivors.

In addition, pressure therapy may have detrimental effects. Obstructive sleep apnea has been described in two patients secondary to pressure garments used to treat severe facial and upper-body burns. Obstructive sleep apnea was confirmed by complex sleep polysonography, and documented desaturations significantly improved with removal of the garments (150). In children, hypertrophic scarring of the head and neck and pressure garment treatment have been reported to cause deformities. Knowledge of such potential complications helps inform treatment decisions. Scar contractures of the mandible may lead to dental occlusion and difficulty with lip closure and subsequent drooling. Pressure garments at the mandible may cause mandibular hypoplasia. Severe neck scarring and resultant contractures may elongate the mandible. Mouth burns are associated with dental changes, including cross bite, crowding, and bite retrusion. Pressure garments may affect facial growth and development of normal contours (151).

Other treatments, silicone gel and massage, have been examined in the literature as well. Ahn et al. investigated the efficacy of an 8-week course of topical silicone gel in ten adults with hypertrophic scars. Each subject acted as their own control; mirror image or adjacent scars were untreated controls. Based on elastometry, skin biopsy, texture, color, thickness, durability, and itching, the silicone-treated areas demonstrated greater improvements than control scars at 4, 8, and 12 weeks (152). This suggests that silicone gel may be an effective intervention for hypertrophic scars. A study of 30 pediatric burn patients examined the use of frictional massage 30 minutes a day over a 3-month period. No significant effects were demonstrated with regard to the vascularity, pliability, and height of the hypertrophic scar (153).

**Blisters**

Blisters are a common complication of partial-thickness burns. Blisters result from inflammatory changes after injury that increases capillary permeability, thereby permitting fluid accumulation between the epidermis and dermis. Management of blisters is a controversial topic in burn care. Sargent performed an extensive literature review on the topic, examining issues related to the care of blisters, including infection, healing, function, aesthetics, patient comfort, dressing care, and cost-effectiveness. The author issued clinical practice guidelines for management of blisters in partial-thickness burns. Small blisters (<6 mm) may be left intact, as they are unlikely to rupture spontaneously, damage underlying tissue, or impede healing. Large blisters (>6 mm) should be debrided. Thick-walled blisters on palms and soles of feet need not be debrided as they are less likely to become infected, and debridement may cause patient discomfort and impaired mobility. Debridement of blisters is associated with faster wound healing and decreased scarring. Decreased scarring is also associated with use of temporary skin substitutes to cover the debrided site (154).

**UV Sensitivity and Skin Pigmentation**

Sun protection is essential for burn survivors. The area of burn injury is susceptible to further damage from ultraviolet radiation from the sun. It is recommended that burn survivors of deep partial- and full-thickness burns avoid and protect against sun exposure for the first few years after injury. Especially after burn injury, people of all different skin pigmentation are at risk for sunburn from ultraviolet radiation. Avoiding direct sun exposure, especially during midday hours when ultraviolet exposure is highest, minimizes risk of sunburn. Covering sites of burn injury with clothing for at least the first year after injury is recommended. In addition, sunscreen with a sun protection factor of 15 or greater should be applied to healed burn sites prior to any sun exposure (155).

Pigmentation changes are common after burn injury. Two studies followed patients longitudinally after superficial partial-thickness burn injury and found that subjects developed hypopigmentation by spectrophotometry measurements at the burn site. Hyperpigmentation correlated with premorbid skin color, age, sun exposure, and time after injury (156,157). Deep partial- and full-thickness burn injuries may result in hypopigmentation or depigmentation. Dyspigmentation
after burn injury can be treated surgically. Al-Qattan reported treating 15 subjects with hyperpigmented skin grafts of the palm and digits with surgical excision and split-thickness skin grafts. He also treated eight subjects with hypopigmented burn sites of the palm, digits, and forearm with dermabrasion and thin split-thickness skin grafts. The hyperpigmentation group exhibited color and texture match and all subjects were satisfied. The hypopigmentation group demonstrated slight hyperpigmentation of the graft site; however, all subjects were satisfied with the results (158). Other investigators demonstrated good results by treating depigmented burn scar using a carbon dioxide laser for dermabrasion followed by split-thickness skin grafting (159).

Malignancy
Development of malignant tumors in chronic burn wounds or scars is extremely rare, but is a frequently reported complication. Most tumors are squamous cell carcinoma; basal cell carcinoma and malignant melanoma are less common. Diagnosis ranges from 20 to 30 years after burn injury. Two recent large cohort studies followed 16,903 and 37,095 burn survivors, respectively, for a mean of 16 years. There was no increased risk for squamous cell carcinoma, basal cell carcinoma, or malignant melanoma in the burn survivors compared to the general population. In addition, subgroup analysis of those with more severe burns and longer follow-up exhibited no increased risk for skin cancer (160, 161).

**METABOLIC COMPLICATIONS**

**Catabolic State**
Patients with burns greater than 40% TBSA experience a hypermetabolic response for at least 1 year after injury. Catabolism contributes significantly to morbidity and mortality. The catabolic state in burn injury is associated with impaired wound healing, increased infection risk, loss of lean body mass, slowed rehabilitation, and delayed community reintegration. Pharmacologic and nonpharmacologic strategies are implemented to help reverse the effects of catabolism. Nonpharmacologic interventions include early burn wound excision and closure, aggressive treatment of sepsis, maintenance of thermal neutrality by elevation of the ambient temperature, high carbohydrate, high protein diet, and early institution of resistive exercises. Pharmacologic interventions may include use of recombinant human growth hormone, low-dose insulin infusion, synthetic testosterone analog (oxandrolone), and β-blockade (29, 101, 162, 163).

The benefits of oxandrolone on hypermetabolism in burn injury are well supported by multiple well-designed studies in recent years. Jeschke et al. performed a prospective randomized controlled trial of 235 burned children with greater than 40% TBSA. Subjects receiving oxandrolone for at least 7 days during acute hospitalization exhibited shorter length of intensive care unit stay and higher lean body mass than controls (164). When oxandrolone was given to children for 1 year after severe burn, subjects demonstrated continued improved lean body mass, bone mineral content, muscle strength, height, and weight compared to controls (165). A separate multicenter prospective randomized controlled trial of 81 adults with 20% to 60% TBSA was stopped early because of significantly shorter length of stay exhibited by the oxandrolone group compared to controls (166).

**Temperature Regulation**
Full-thickness burns damage the sweat glands of the dermis. Despite treatment with skin grafting, the sweat glands are not replaced or regenerated. Impaired sweating may affect thermoregulation (167), particularly with those with larger TBSA burns. Patients with large burn injuries often report overheating and exaggerated sweating response in areas of unburned skin with exercise and heat. Such complaints may interfere with burn survivors exercise tolerance, overall fitness and health, as well as occupational reintegration (168).

Austin et al. studied the effects of exercise on thermoregulation in six burn survivors with greater than 30% burn injuries and three nonburned controls. Subjects exercised for 1 hour on a cycle ergometer with an ambient temperature of 35 degrees centigrade. Subjects with burn injuries tolerated moderate exercise. The burn subjects maintained their heat tolerance by increasing sweat rates in areas of unburned skin (169). In another study, subjects with larger burns (mean 49% TBSA) performed bench stepping in a 40 degree centigrade environment for 3 hours. Despite a high sweat rate from their unburned skin, the severely burned subjects were unable to maintain body temperature compared to nonburned controls (170). Children appear to tolerate a short duration of moderate exercise. McEntire et al. studied 15 children with severe burns (>40% TBSA) and compared them with 13 healthy children controls. Subjects completed 20 minutes of treadmill exercise at approximately 75% of their peak aerobic power. There was no significant difference in body temperature between the severely burned children and controls (171).

**PSYCHOSOCIAL ISSUES**
Psychosocial complications after burn injury are common and present major obstacles to burn survivors’ rehabilitation and community integration. Burns are traumatic, functionally impairing, excruciatingly painful, and disfiguring. All of these factors contribute to the prevalence of psychosocial complications after burns. In addition, premorbid psychiatric history is higher in burn patients than in the general population. Approximately one third of burn injuries are associated with concomitant alcohol or drug use (6). Preinjury psychiatric problems, including anxiety, depression, and other problems, are more common in burn patients than in the general population and have been documented at 28% to 75% (172). Common psychosocial issues after burn injury include posttraumatic stress, depression, anxiety, sleep disorders, and...
community integration. The research community has excelled in identifying and describing these important psychosocial issues after burn injury; however, at the time of this publication few studies have examined treatment interventions. Future investigation in this area is needed.

Posttraumatic Stress
Posttraumatic stress occurs after one experiences a traumatic event with the threat of injury of death to self or others. Three types of symptoms characterize posttraumatic stress disorder (PTSD): reexperiencing the event, avoidance of reminders of the event, and an increased state of arousal. For the diagnosis of PTSD symptoms must be present for greater than 1 month, and for acute stress disorder symptoms must be present for greater than 2 days during the first 4 weeks after the event. Over one half of burn survivors report posttraumatic stress symptoms at 1 day, 1 month, and 1 year after injury. Commonly reported symptoms include sleep disturbance, recurrent and intrusive recollections of the injury, avoidance of thoughts or feelings associated with the burn, and distress at reminders of the burn. Twenty-one percent of subjects at 1 month and 19% of subjects at 1 year met full criteria for PTSD. Overall, posttraumatic stress symptoms decrease over time. It is common for burn survivors to endorse some posttraumatic stress symptoms but not fulfill the diagnostic criteria for posttraumatic stress and acute stress disorders (173). Risk factors for the development of PTSD after burn injury include posttraumatic stress symptoms at earlier time points, female gender, social support, and the size and location of burn injury (174). Treatment interventions for posttraumatic stress have not been validated in the literature. Nonetheless, screening for posttraumatic stress symptoms and referral for pharmacological and nonpharmacological treatment interventions are recommended.

Depression
Depression is a similarly common psychiatric problem after burn injury. One group of investigators reported the prevalence of moderate to severe depression symptoms using the Beck depression inventory at 53% at 1 month, 34% at 1 year and 45% at 2 years after injury (175). Premorbid psychiatric problems, head or neck burns, length of hospital stay, and female gender are risk factors associated with depression (175,176). Edwards et al. followed 128 burn survivors and found that one third of subjects reported some form of suicidal ideation during the first year after injury. Pain severity at discharge was the only significant predictor of suicidal ideation (177). The frequency of suicidal ideation highlights the importance of early identification of symptoms in suicide prevention. Furthermore psychosocial outcomes are complex and multifaceted. Pain management may impact depression and other psychiatric complications. There are no treatment studies to date for depression in burn injury. Clinicians should identify those at risk for depression and those with depressive symptoms and provide appropriate treatment options.

Sleep Disturbances
Sleep disturbances may be related to a psychiatric problem, medical issue, or a direct consequence of the burn injury. As many as 74% of burn survivors reported sleep problems at 1 week after discharge from the hospital. Common problems include nighttime awakenings, daytime napping, nighttime pain, and difficulty with sleep onset (178). Sleep problems, such as insomnia and nightmares, persist at 1 year after injury (173). Gottschlich et al. examined polysomnography in 11 severely burned inpatient children over forty-three 24 hour sessions. Deep sleep (stages III and IV) and rapid eye movement were completely missing in 40% and 19% of recordings, respectively. This demonstrates altered sleep architecture in the acute period after burn injury (179). In a separate study of hospitalized burned adults, investigators found that sleep duration was 5.5 hours with numerous awakenings. In addition, significant temporal relationships were documented between sleep, pain, and analgesic medications. A lack of sleep one night was associated with increased pain complaints the following day and increased analgesic intake. High levels of pain and analgesic intake during the day were associated with poor sleep the following night (180).

It is important to ask burn patients about their sleep. Since other psychiatric problems can contribute to sleep problems, clinicians should consider these issues as well when treating sleep disorders in burn injury. Treatment of pain, pruritis, depression, and posttraumatic stress and anxiety may improve burn survivors’ sleep.

Community Integration
The ultimate goal of rehabilitation is reintegration into society, including a return to one’s work, school, and recreational and community activities. The various physical and psychological complications of burn injury detailed above may result in significant impairments that hinder community integration. In a study of 463 burn survivors who completed the Community Integration Questionnaire, researchers found significant problems in home integration, social integration, and productivity. Home integration was best predicted by gender and living situation; social integration was best predicted by marital status; productivity was best predicted by burn severity, age, and preburn job satisfaction (181). Brych et al. followed 363 burned adults who were employed at the time of injury and found that the mean time off work was 17 weeks. Sixty-six and ninety percent of survivors returned to work at 6 and 24 months, respectively. However, a subgroup analysis found that only 37% returned to the same job, with the same employer, without accommodations (136). This more detailed data underscores the severity of job disruption after burn injury. Factors associated with unemployment after burn include severity of injury, extremity burns, premorbid psychiatric history, and premorbid unemployment (136,182,183). Esselman et al. examined barriers to return to work in 154 burn survivors using a survey administered 16 times over 1 year. At 1 year, significant barriers included physical abilities, psychosocial factors (nightmares, flashbacks, body image), and working conditions (humidity, temperature, safety) (168).
Some burn survivors experience permanent impairment as a result of their injury. The American Medical Association publishes guidelines for determining the extent of impairment. Impairment is graded as a percentage of whole body impairment. Considerations in Children and Elders

In the sixth edition of the guidelines, impairment in burn injury is primarily dependent upon the following factors: the severity of the skin condition; the frequency, intensity, and complexity of symptoms and the treatment regimen; the ability to perform ADL. This impairment rating may be modified by objective physical examination findings, facial disfigurement, and related impairments of other organ systems including musculoskeletal, respiratory, cardiovascular, endocrine, and gastrointestinal. It is important to remember that these guidelines are intended for evaluation of patients with permanent impairments, those who have attained maximal recovery (184).

**SPECIAL CONSIDERATIONS**

**Pediatric Burns**

The physiologic and anatomic differences between adults and children need to be considered during acute and long-term medical management after burn injury. Children are not small adults. Anatomically, the pediatric trachea is shorter, the glottis is more anterior, and the diameter of the airway is smaller compared to an adult (185). These anatomic differences are important during intubation and in the presence of facial burns, upper airway damage, or edema. Children have an increased risk of bronchospasm (186). The smaller airways make them more susceptible to occlusion if there is pulmonary debris, such as after an inhalation injury.

After a significant burn injury, children often demonstrate impaired cardiac function. This is particularly relevant if the child is less than 1 year of age, has underlying cardiac anomalies or sustained an inhalation injury (187).

**Pediatric Exercise**

The goals of exercise are similar to adults, but the methods are different. Exercises are designed to be compatible with the child’s level of development and incorporate play. Toys should be developmentally appropriate for the child. For instance, nesting blocks for the child less than 2 years old are used to enhance hand function. In an older child, wheelbarrow walking or crab walking stretches the shoulder muscles and strengthens the shoulder girdle. Bicycling and soccer ball kicking are appropriate for lower extremity stretching, strengthening and endurance training in older children. A wide variety of products, such as theraband and therapy putty can provide resistance and facilitate strengthening. Handheld computer games are valuable for fine motor function of the hand. Interactive computer games, such as the WII, utilize visual monitors and encourage gross motor participation.

**Splints for Children**

The skin of infants and children differs from that of adults. The skin of children is thinner, full-thickness wounds are more likely to occur (188). Stretching techniques, splint fabrication, and pressure garments (Table 43.5) are designed to accommodate the thinner and more fragile skin of children. Therefore, the treatment of contractures in children is similar to that of adults, with a few modifications.

It is often difficult for parents to stretch a child’s joints. Aggressive stretching of contractured joints is often difficult for children to tolerate. Stretching under anesthesia and pain medication administration prior to therapy may be indicated.

When using splints in the pediatric population, one must take into consideration the smaller size of body parts and the quality of the skin. Extra padding or foam may be needed. Compared to adults, children have hypermobile joints; this is also a consideration. Splints in children require frequent evaluation by caregivers because of normal growth.

Many suggest that splints for infants are applied when the child is sleeping and allow active use of the extremity during wakeful periods. If this is insufficient, wearing time is increased during the daytime and splints are removed for exercise or therapeutic activities. Children tolerate prolonged joint immobilization and do not develop joint contractures like adults.

Dynamic splints are an option for adolescents but are generally not tolerated by infants or small children because they are too difficult to keep in place. Small children are better with static splints or serial casting. By age 3 or 4, children may use

<table>
<thead>
<tr>
<th>TABLE 43.5 Compression Garments: Considerations in Children and Elders</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pediatrics</strong></td>
</tr>
<tr>
<td>Limited understanding and reasoning</td>
</tr>
<tr>
<td>Decreased attention span</td>
</tr>
<tr>
<td>Decreased cooperation</td>
</tr>
<tr>
<td>Preexisting behavioral problems</td>
</tr>
<tr>
<td>Locate zippers in locations not easily accessible by the child (e.g., on the back not the chest, use colored garments, blue, etc.)</td>
</tr>
<tr>
<td>Need assistance to don or doff</td>
</tr>
<tr>
<td>Small body parts require careful measuring</td>
</tr>
<tr>
<td>Rapid growth and high activity (frequent refitting)</td>
</tr>
</tbody>
</table>
adjustable three-point splints at the knee or elbow. Dynamic ankle splints work in this older group as well.

For young children, hand and ankle splints are made with longer proximal extensions to secure the splint from sliding out of position. For instance, a hand splint may extend to the wrist and forearm. Soft straps that cross the volar aspect of the wrist and forearm secure the device. Splints may need to be covered with a garment, such as a sock, to prevent the child from removing the device. The hand of a young child or infant is typically splinted flat with the wrist in extension, fingers extended and abducted, and the thumb in extension and abducted radially. Sandwich splints lined with foam with a dorsal and ventral component are useful

**Pediatric Reconstructive Surgery**

Children often require surgery to restore function. In addition, surgery may be indicated to correct deformities and functional impairments that develop as a result of growth. Scar tissue and some grafts do not elongate with growth and consequently reconstructive surgery is required. If ROM is normal as the child grows, surgery is not indicated. The neck and axillae require close monitoring. Reconstructive surgery is performed approximately 1 year after menarche for girls with significant scarring of the anterior chest.

**Geriatric Burns**

In the geriatric population, preexisting physical limitations or medical problems have a greater impact on rehabilitation than in younger survivors. Premorbid functional level and health are particularly important in establishing a geriatric rehabilitation plan. Older adults often have cardiovascular disease. In the face of a serious burn injury, an older patient may develop worsening cardiac function or myocardial infarction (189). There is a baseline decrease in pulmonary function that is associated with aging. Also, many older individuals have underlying pulmonary or cardiac disease that negatively impacts respiratory reserve and function. A burn injury coupled with inhalation injury markedly increases the risk of mortality and morbidity. Principles of cardiopulmonary rehabilitation are included in the exercise plan for these individuals.

**Skin and Wound Healing**

The atrophic skin of the elderly is characterized by the absence of rete pegs, a thinner dermis, and a reduced number of skin appendages (190). As a consequence, elderly sustain more severe burns than younger adults from the same injury (191). Additionally, the cellularity of skin decreases with age, leading to a reduced number of macrophages and fibroblasts, thereby prolonging healing time. The skin of the elderly exhibits less turgor than that of younger people and is due in part to a decrease in glycosaminoglycans. The redundant tissue can be harvested for use as full-thickness grafts and improved functional outcome.

Early excision and grafting is a well-established approach to burn injury in children and adults; however, the risk-to-benefit ratio is not as clear in the elderly (192,193). The burn wound and donor site are additive, effectively increasing the surface area of open wounds resulting in increased rates of morbidity and mortality.

**Exercise**

During normal aging, one experiences sarcopenia, or muscle mass loss as well as loss of strength (194). Strength declines at a rate of 15% per decade after age 50%, and 30% after age 70. There is evidence that resistance training can reduce and reverse some of the muscle changes and strength loss of aging (195). However, sarcopenia, and the associated weakness, is aggravated by disuse or bed rest. As a result, elderly are at greater risk for loss of function after burn injury.

A geriatric burn exercise program takes into account one’s preexisting musculoskeletal and neuromuscular conditions that influence the exercise prescription. One must determine if passive, active-assistive, or active exercise therapy is appropriate. An exercise regimen includes ROM, stretching, and strengthening.

Stretching principles are similar to those of any adult. Each joint is stretched individually. This progresses to stretching of an entire limb. With redundant tissue present in some body regions, scarring and contracture development are less frequent than in other adults.

ROM exercise is modified in the presence of underlying joint disease, such as degenerative arthritis. A burn injury involving a chronically arthritic shoulder is treated differently than a nonarthritic joint. One may need to modify goals to achieve a functional range rather than full ROM.

Strength training starts slowly and gradually increases in intensity. It is important to closely monitor cardiopulmonary status. As in other age groups, strengthening usually begins with muscles that oppose scar contracture formation. Exercise should advance judiciously beginning with nonresistance exercises and progressing to resistance bands and light weights. Maximal exercise capacity and oxygen consumption decrease with age; however, there is evidence that aerobic capacity can increase with endurance training in the elderly.

**FACIAL BURNS**

Skin contractures and scars of the face are challenging to prevent. Facial tissue is highly mobile and has few points of fixation. Common facial deformities include ectropion of the lower eyelid and microstomia of the mouth. The development of ectropion prevents eyelid closing leading to eye irritation that may result in corneal ulcers. Microstomia is a contracture of the mouth that impairs mouth function.

Current review of the literature indicates that rehabilitation interventions for face burns lack general agreement. No generally accepted medical or rehabilitation protocols exist. However, a recent survey indicates that the common treatment methodologies include positioning, splinting, exercise, stretching, and pressure therapy (196).
During the acute phase many begin treatment with splints for the neck, mouth, nose, and ears. Active ROM of the face, including eyes and mouth, is initiated as soon as the patient can participate. For ear injuries, pillows are used to prevent chondritis and tissue adhering to the bed linens. Devices are fabricated to relieve pressure for an injured pinna.

Microstomia is a contracture of the oral aperture of the face (Fig. 43-15). It is associated with impaired oral hygiene, eating, and speech. In addition, microstomia may disrupt muscular and dental development in children (197). The contracture may involve the skin of the mouth or involve the perioral musculature that creates the sphincter of the mouth. The orbicularis oris muscle is separated from the surface by a thin subcutaneous layer and from the mucosa below by a thin submucosal layer.

Oral stretching splints are used to maintain the normal dimension of the mouth for those at risk of developing microstomia. These splints are often custom made and tailored to maximize comfort and benefit. Commercial devices are also available. Splint designs vary by direction of force (horizontal, vertical, or circumoral stretch) and orientation (intraoral or extraoral). Monitoring is required for skin irritation and breakdown. Factors to consider in prescribing a microstomia splint include age, dentition, stage of dental development in children, and location and depth of injury. Devices should be comfortable to wear, insert, and clean. Compliance is often challenging. To maximize success and achieve optimal patient compliance, a graduated wearing schedule is advised and more than one device is indicated (198,199).

The timing of initiating pressure therapy in clinical practice for facial burns is variable (200). Pressure is an accepted treatment for facial burn scarring and deformity. Face masks or pressure garments fitted to the face are widely used. In 1979, Rivers et al. (201) first described the fabrication of a transparent face mask for application of pressure to prevent scarring. Transparent rigid face mask is the most common (196) facial pressure device (Fig. 43-16). There is better patient compliance and social acceptance of a transparent face mask than a fabric face garment (202,203). Accuracy of fit is critical to successful scar suppression. The transparent face mask allows for better fit, as one can easily monitor for scar blanching (201). Allely et al. described the use of laser Doppler imaging for mask fabrication. This technique holds the potential for increased precision and efficiency in fabricating transparent face masks (204).

There exists concern that pressure will alter the dental, maxillary, and mandibular features of a child. Some recommend the involvement of an orthodontist in all children with facial burns (205,206). Literature review reveals limited support for the effects of pressure therapy on preservation of craniofacial features after facial burns. In a literature review on the topic, investigators were unable to draw any conclusions because of a lack of statistically robust data (207). One study met inclusion criteria and included six children, three with total face masks and three with partial face masks of the lower face. Those with the lower face mask exhibited anterior protrusion of the teeth and little reduction in the inferior growth of the mandible. These changes persisted after pressure therapy was discontinued. For subjects with full face masks, maxillary horizontal growth was more affected than vertical growth. Radiographs in both groups after pressure therapy was discontinued demonstrated resolution in the bony abnormalities (96,205).

CONCLUSION

With more patients surviving burn injuries, care is increasingly focused on the complications, rehabilitation, and long-term outcomes of burn survivors. Burn injuries may present as major catastrophic injuries, with a complex array of problems that include contractures, hypertrophic scarring, pain, neuropathy, and psychosocial problems. Physiatry is an integral component of the specialized multidisciplinary burn care team throughout the continuum of care. Rehabilitation interventions begin immediately after injury in the intensive care unit aimed at preventing long-term complications of burns. Splinting, positioning, and exercise are hallmarks of early rehabilitation. Burn care may continue for years after hospital discharge, managing physical and psychosocial impairments after burn injury and ultimately promoting maximal reintegration into the community.
REFERENCES


101. Klein GL, Wolf SE, Langman CB, et al. Effects of therapy with recombinant human growth hormone on insulin-like growth factor system components and serum levels of biochemical markers of bone forma-

maintenance of pamidronate on the conservation of bone mass fol-

103. Prekora R, Hrenkon DN, Sherard DJ, et al. Pamidronate preserves bone mass for at least 2 years following acute administration for pediat-

104. Evans EB. Bone and joint changes secondary to burns. In: Lewis SR,


115. Carrougher GJ, Pateck JT, Sharer SR, et al. Comparison of patient satis-


118. Weinberg K, Binshall C, Vail D, et al. Pain and anxiety with burn dress-


120. Robert R, Brack A, Blakeney P, et al. A double-blind study of the anal-


124. Patterson DR, Pateck JT. Baseline pain as a moderator of hypnotic anal-


126. Hoffman HG, Patterson DR, Carrougher GJ. Use of virtual reality for adjunctive treatment of adult burn pain during physical therapy: a con-

127. Hoffman HG, Patterson DR, Carrougher GJ, et al. Effectiveness of vir-


Rehabilitation for Patients with Cancer Diagnoses

GENERAL ASPECTS OF REHABILITATION FOR CANCER PATIENTS

Historical Background
Patients with cancer diagnoses are living longer because of a combination of early detection, a broader selection of cancer treatment options, and better general medical management. While the treatment of cancer has moved toward less invasive and more preservation-oriented techniques, there remains a high incidence of disability in individuals with cancer and survivors (1–3). Census data indicates cancer as the 13th most common cause of self-reported disability (4).

Rehabilitation Expectations in the Cancer Population
While rehabilitation goals for cancer patients have historically been organized into restorative, supportive, preventive, and palliative categories (5), another important concept is that the clinical course can be both chronic and dynamic (6), resulting in need for rehabilitation at various points in the disease trajectory. The heterogeneity of cancer types creates very diverse rehabilitation needs in this population. The more common issues are highlighted here.

Demographics of Cancer and Its Relevance to Rehabilitation
Cancer is the second leading cause of death in the United States, accounting for approximately one in every four deaths among both children and adults (7). Nearly 60% of all cancers occur in individuals age 65 and older, and age-adjusted incidence is ten times greater in individuals over age 65 than younger age groups (8). The overall 5-year relative survival rate for 1996 to 2002 was 66% (Table 44-1), up from 51% in 1975 to 1977, and 35% in the 1950s (9). Five-year survival for childhood cancers has improved from less than 50% prior to the 1970s to about 80% today, though there is variability by site (7). The lifetime risk of being diagnosed with cancer is approximately 40%, and there are greater than 10.7 million individuals alive in the United States with a history of cancer diagnosis (2004 estimate) (9).

Restorative services may not be appropriate for those with a markedly reduced life expectancy. However, survival often varies markedly within a given tumor type because of factors such as stage of disease and histology, so each case must be approached individually. Radically improved survival rates now compel rehabilitation professionals to examine the needs of long-term survivors. Of particular interest are the more common cancers with large survivorship populations, such as breast and prostate cancer (7), and cancers known to have a high incidence of disabling complications (see next section) (10). Since cancer is most common in older age groups, the impact of cancer within the geriatric population is receiving increasing attention. In general, cancer in the elderly may not be more disabling than other common medical conditions such as diabetes or congestive heart failure (11), but more severe symptoms or extensive treatment is associated with greater loss of function (12).

CANCER REHABILITATION SERVICE DELIVERY
One seminal study of the rehabilitation needs of individuals with cancer identified that 54% had physical medicine problems (10), with very high incidence (70% or greater) among those with central nervous system (CNS), breast, lung, or head and neck tumors. There was a large gap between the identified rehabilitation needs and the services actually delivered, which improved dramatically with a program for patient education, automatic screening of patients for rehabilitation needs, and introducing a physiatrist into the clinical oncology team. Organized cancer rehabilitation programs, while not widespread, have been described (13). Core components include a committed administration, physiatrist as medical director, and effective marketing, as well as practical considerations including accessibility.

Because of the heterogeneity of rehabilitation needs across the cancer spectrum, as well as the complexity of care in individual cases, screening and surveillance tools should be employed which are both systematic and clinically practical. Issues may be present at symptom, impairment and function levels. Visual analog scores have been described for pain, fatigue, appetite, mood, and sleep (14). Office-based tests that have been applied...
### TABLE 44.1

**Estimated New Cancer Cases and Deaths by Sex, United States, 2008**

<table>
<thead>
<tr>
<th>Condition</th>
<th>Both Sexes</th>
<th>Male</th>
<th>Female</th>
<th>Both Sexes</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Estimated New Cases</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All sites</td>
<td>1,437,180</td>
<td>745,180</td>
<td>692,000</td>
<td>565,650</td>
<td>294,120</td>
<td>271,530</td>
</tr>
<tr>
<td>Oral cavity and pharynx</td>
<td>35,310</td>
<td>25,310</td>
<td>10,000</td>
<td>7,590</td>
<td>5,210</td>
<td>2,380</td>
</tr>
<tr>
<td>Tongue</td>
<td>10,140</td>
<td>7,280</td>
<td>2,860</td>
<td>1,880</td>
<td>1,210</td>
<td>670</td>
</tr>
<tr>
<td>Mouth</td>
<td>10,820</td>
<td>6,590</td>
<td>4,230</td>
<td>1,840</td>
<td>1,120</td>
<td>720</td>
</tr>
<tr>
<td>Pharynx</td>
<td>12,410</td>
<td>10,060</td>
<td>2,350</td>
<td>2,200</td>
<td>1,620</td>
<td>580</td>
</tr>
<tr>
<td>Other oral cavities</td>
<td>1,940</td>
<td>1,380</td>
<td>560</td>
<td>1,670</td>
<td>1,260</td>
<td>410</td>
</tr>
<tr>
<td>Digestive system</td>
<td>271,290</td>
<td>148,560</td>
<td>122,730</td>
<td>135,130</td>
<td>74,850</td>
<td>60,280</td>
</tr>
<tr>
<td>Esophagus</td>
<td>16,470</td>
<td>12,970</td>
<td>3,500</td>
<td>14,280</td>
<td>11,250</td>
<td>3,030</td>
</tr>
<tr>
<td>Stomach</td>
<td>21,500</td>
<td>13,190</td>
<td>8,310</td>
<td>10,880</td>
<td>6,450</td>
<td>4,430</td>
</tr>
<tr>
<td>Small intestine</td>
<td>6,110</td>
<td>3,200</td>
<td>2,910</td>
<td>1,110</td>
<td>580</td>
<td>530</td>
</tr>
<tr>
<td>Colon</td>
<td>108,070</td>
<td>53,760</td>
<td>54,310</td>
<td>49,960</td>
<td>24,260</td>
<td>25,700</td>
</tr>
<tr>
<td>Rectum (deaths included with colon)</td>
<td>40,740</td>
<td>23,490</td>
<td>17,250</td>
<td>680</td>
<td>250</td>
<td>430</td>
</tr>
<tr>
<td>Liver and intrahepatic bile duct</td>
<td>21,370</td>
<td>15,190</td>
<td>6,180</td>
<td>18,410</td>
<td>12,570</td>
<td>5,840</td>
</tr>
<tr>
<td>Gallbladder and other biliary</td>
<td>9,520</td>
<td>4,500</td>
<td>5,020</td>
<td>3,340</td>
<td>1,250</td>
<td>2,090</td>
</tr>
<tr>
<td>Pancreas</td>
<td>37,680</td>
<td>18,770</td>
<td>18,910</td>
<td>34,290</td>
<td>17,500</td>
<td>16,790</td>
</tr>
<tr>
<td>Other digestive organs</td>
<td>4,760</td>
<td>1,470</td>
<td>3,290</td>
<td>2,180</td>
<td>740</td>
<td>1,440</td>
</tr>
<tr>
<td>Respiratory system</td>
<td>232,270</td>
<td>127,880</td>
<td>104,390</td>
<td>166,280</td>
<td>94,210</td>
<td>72,070</td>
</tr>
<tr>
<td>Larynx</td>
<td>12,250</td>
<td>9,880</td>
<td>2,570</td>
<td>3,670</td>
<td>2,910</td>
<td>760</td>
</tr>
<tr>
<td>Lung and bronchus</td>
<td>215,020</td>
<td>114,890</td>
<td>100,330</td>
<td>161,840</td>
<td>90,810</td>
<td>71,030</td>
</tr>
<tr>
<td>Other respiratory organs</td>
<td>5,000</td>
<td>3,510</td>
<td>1,490</td>
<td>770</td>
<td>490</td>
<td>280</td>
</tr>
<tr>
<td>Bones and joints</td>
<td>2,380</td>
<td>1,270</td>
<td>1,110</td>
<td>1,470</td>
<td>820</td>
<td>650</td>
</tr>
<tr>
<td>Soft tissue (including heart)</td>
<td>10,390</td>
<td>5,720</td>
<td>4,670</td>
<td>3,680</td>
<td>1,880</td>
<td>1,800</td>
</tr>
<tr>
<td>Skin (ex-basal and squamous)</td>
<td>67,720</td>
<td>38,150</td>
<td>29,570</td>
<td>11,200</td>
<td>7,360</td>
<td>3,840</td>
</tr>
<tr>
<td>Melanoma of the skin</td>
<td>62,480</td>
<td>34,950</td>
<td>27,530</td>
<td>8,420</td>
<td>5,400</td>
<td>3,020</td>
</tr>
<tr>
<td>Other nonepithelial skin</td>
<td>5,240</td>
<td>3,200</td>
<td>2,040</td>
<td>2,780</td>
<td>1,960</td>
<td>820</td>
</tr>
<tr>
<td>Breast</td>
<td>184,450</td>
<td>1,990</td>
<td>182,460</td>
<td>40,930</td>
<td>450</td>
<td>40,480</td>
</tr>
<tr>
<td>Genital system</td>
<td>274,150</td>
<td>195,660</td>
<td>78,490</td>
<td>57,620</td>
<td>29,330</td>
<td>28,490</td>
</tr>
<tr>
<td>Uterine cervix</td>
<td>11,070</td>
<td>11,070</td>
<td>3,870</td>
<td>3,870</td>
<td>3,870</td>
<td></td>
</tr>
<tr>
<td>Uterine corpus</td>
<td>40,100</td>
<td>40,100</td>
<td>7,470</td>
<td>7,470</td>
<td>7,470</td>
<td></td>
</tr>
<tr>
<td>Ovary</td>
<td>21,650</td>
<td>21,650</td>
<td>15,520</td>
<td>15,520</td>
<td>15,520</td>
<td></td>
</tr>
<tr>
<td>Vulva</td>
<td>3,460</td>
<td>3,460</td>
<td>870</td>
<td>870</td>
<td>870</td>
<td></td>
</tr>
<tr>
<td>Vagina and other genital, female</td>
<td>2,210</td>
<td>2,210</td>
<td>760</td>
<td>760</td>
<td>760</td>
<td></td>
</tr>
<tr>
<td>Prostate</td>
<td>186,320</td>
<td>186,320</td>
<td>28,660</td>
<td>28,660</td>
<td>28,660</td>
<td></td>
</tr>
<tr>
<td>Testis</td>
<td>8,090</td>
<td>8,090</td>
<td>380</td>
<td>380</td>
<td>380</td>
<td></td>
</tr>
<tr>
<td>Penis and other genital organs, male</td>
<td>1,250</td>
<td>1,250</td>
<td>290</td>
<td>290</td>
<td>290</td>
<td></td>
</tr>
<tr>
<td>Urinary system</td>
<td>125,490</td>
<td>85,870</td>
<td>39,620</td>
<td>27,810</td>
<td>18,430</td>
<td>9,380</td>
</tr>
<tr>
<td>Urinary bladder</td>
<td>68,810</td>
<td>51,230</td>
<td>17,580</td>
<td>14,100</td>
<td>9,850</td>
<td>4,150</td>
</tr>
<tr>
<td>Kidney and renal pelvis</td>
<td>54,390</td>
<td>33,130</td>
<td>21,260</td>
<td>13,010</td>
<td>8,100</td>
<td>4,910</td>
</tr>
<tr>
<td>Ureter and other urinary organs</td>
<td>2,290</td>
<td>1,510</td>
<td>780</td>
<td>780</td>
<td>380</td>
<td>320</td>
</tr>
<tr>
<td>Eye and orbit</td>
<td>2,390</td>
<td>1,340</td>
<td>1,050</td>
<td>240</td>
<td>130</td>
<td>110</td>
</tr>
<tr>
<td>Brain and other nervous system</td>
<td>21,810</td>
<td>11,780</td>
<td>10,030</td>
<td>13,070</td>
<td>7,420</td>
<td>5,650</td>
</tr>
<tr>
<td>Endocrine system</td>
<td>39,510</td>
<td>10,030</td>
<td>29,480</td>
<td>2,430</td>
<td>1,110</td>
<td>1,320</td>
</tr>
<tr>
<td>Thyroid</td>
<td>37,340</td>
<td>8,930</td>
<td>28,410</td>
<td>1,590</td>
<td>680</td>
<td>910</td>
</tr>
<tr>
<td>Other endocrine</td>
<td>2,170</td>
<td>1,100</td>
<td>1,070</td>
<td>840</td>
<td>430</td>
<td>410</td>
</tr>
<tr>
<td>Lymphoma</td>
<td>74,340</td>
<td>39,850</td>
<td>34,490</td>
<td>20,510</td>
<td>10,490</td>
<td>10,020</td>
</tr>
<tr>
<td>Hodgkin’s lymphoma</td>
<td>8,220</td>
<td>4,400</td>
<td>3,820</td>
<td>1,350</td>
<td>700</td>
<td>650</td>
</tr>
<tr>
<td>Non-Hodgkin’s lymphoma</td>
<td>66,120</td>
<td>34,450</td>
<td>30,670</td>
<td>19,160</td>
<td>9,790</td>
<td>9,370</td>
</tr>
<tr>
<td>Myeloma</td>
<td>19,920</td>
<td>11,190</td>
<td>8,730</td>
<td>10,690</td>
<td>5,640</td>
<td>5,050</td>
</tr>
</tbody>
</table>
to assess impairments and basic function include manual muscle testing, grip dynamometry, range of motion, limb girths, up and go test (15), timed walking, single foot balance, tandem walking (16), modified sit-and-reach test (for flexibility), and stand and sit test (for strength) (17). Karnofsky (Table 44-2) and Eastern Cooperative Oncology Group (ECOG) scales have been employed by oncologists as a measure of performance status, less so for functional outcome; hence, these measures are of uncertain value in rehabilitation. Some questionnaire-based tools developed for oncology patients incorporate both functional and quality-of-life measures. Examples include the Functional Assessment of Cancer Therapy (FACT), European Organization for Research and Treatment of Cancer Quality of Life Questionnaire (EORTCQ), Cancer Rehabilitation Evaluation System (CARES), and Functional Living Index-Cancer (FLIC). The 36-Item Short Form Health Survey (SF-36), a health status instrument, has also been applied to the cancer population (18–20).

### TABLE 44.2 Karnofsky Scale

<table>
<thead>
<tr>
<th>Karnofsky Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Able to carry on normal activity; no special care is needed.</strong></td>
</tr>
<tr>
<td><strong>10 Normal; no complaints, no evidence of disease</strong></td>
</tr>
<tr>
<td><strong>9 Able to carry on normal activity; minor signs or symptoms of disease</strong></td>
</tr>
<tr>
<td><strong>8 Normal activity with effort; some signs or symptoms of disease</strong></td>
</tr>
<tr>
<td><strong>Unable to work; able to live at home; cares for most personal needs; varying amounts of assistance is needed.</strong></td>
</tr>
<tr>
<td><strong>7 Cares for self; unable to carry on normal activity or do active work</strong></td>
</tr>
<tr>
<td><strong>6 Requires occasional assistance, but is able to care for most of own needs</strong></td>
</tr>
<tr>
<td><strong>5 Requires considerable assistance and frequent medical care</strong></td>
</tr>
<tr>
<td><strong>Unable to care for self; requires equivalent of institutional or hospital care; disease may be progressing rapidly.</strong></td>
</tr>
<tr>
<td><strong>4 Disabled; requires special care and assistance</strong></td>
</tr>
<tr>
<td><strong>3 Severely disabled; hospitalization is indicated, although death is not imminent</strong></td>
</tr>
<tr>
<td><strong>2 Very sick; hospitalization necessary; active supportive treatment necessary</strong></td>
</tr>
<tr>
<td><strong>1 Moribund, fatal process progressing rapidly</strong></td>
</tr>
<tr>
<td><strong>0 Dead</strong></td>
</tr>
</tbody>
</table>

**TABLE 44.1 Estimated New Cancer Cases and Deaths by Sex, United States, 2008 (Continued)**

<table>
<thead>
<tr>
<th>Leukemia</th>
<th>Acute lymphocytic leukemia</th>
<th>Chronic lymphocytic leukemia</th>
<th>Acute myeloid leukemia</th>
<th>Chronic myeloid leukemia</th>
<th>Other leukemia</th>
<th>Other and unspecified primary sites</th>
</tr>
</thead>
<tbody>
<tr>
<td>44,270</td>
<td>5,430</td>
<td>15,110</td>
<td>13,290</td>
<td>4,830</td>
<td>5,610</td>
<td>31,490</td>
</tr>
<tr>
<td>25,180</td>
<td>3,220</td>
<td>8,750</td>
<td>7,200</td>
<td>2,800</td>
<td>3,210</td>
<td>15,400</td>
</tr>
<tr>
<td>19,090</td>
<td>2,210</td>
<td>6,360</td>
<td>6,090</td>
<td>2,030</td>
<td>2,400</td>
<td>16,090</td>
</tr>
<tr>
<td>21,710</td>
<td>1,460</td>
<td>4,390</td>
<td>8,820</td>
<td>450</td>
<td>6,590</td>
<td>45,090</td>
</tr>
<tr>
<td>12,460</td>
<td>800</td>
<td>2,600</td>
<td>5,100</td>
<td>200</td>
<td>3,760</td>
<td>24,330</td>
</tr>
<tr>
<td>9,250</td>
<td>660</td>
<td>1,790</td>
<td>3,720</td>
<td>250</td>
<td>2,830</td>
<td>20,760</td>
</tr>
</tbody>
</table>


Cancer (FLIC). The 36-Item Short Form Health Survey (SF-36), a health status instrument, has also been applied to the cancer population (18–20).

### Inpatient Rehabilitation

Several studies have shown that cancer patients and noncancer patients achieve comparable functional gains from inpatient rehabilitation programs, as measured by the Functional Independence Measure (FIM). Patients with neoplastic spinal cord injury (SCI) (21) and brain tumors (22,23) have been shown to have shorter rehabilitation lengths of stay but similar discharge rates to home when compared to age matched controls. Reasons may include higher initial FIM score (seen in some studies), and possibly fewer behavioral sequelae, better social support, and expedited discharge planning due to poor long-term prognosis in some cases (24). Functional improvements made during acute rehabilitation are maintained 3 months after discharge (25). Chemotherapy, radiation therapy, and specific tumor type have not been shown to adversely affect rehabilitation outcome (26,29).

The incidence of transfer back to acute care from rehabilitation is higher than noncancer patients in most (26–28) but not all (29) series. One study that examined reasons for transfer found that infection was more common in cancer patients than in controls (28). Low albumin, elevated creatinine, and use of feeding tube or indwelling bladder catheter have been reported to be risk factors for transfer (30). Prognosis and the patient’s general tolerance of rehabilitation therapies must be weighed in the decision for inpatient rehabilitation. However, poor expected long-term survival is not a contraindication if substantial functional gains are likely to be made in the short or intermediate term. Functional gains for patients with advanced disease should be defined broadly enough to include family/caregiver training that will allow terminally ill patients to remain home with hospice services if that is their choice.

### Outpatient Rehabilitation

Outpatient care, typically, addresses specific musculoskeletal or soft-tissue problems, such as lymphedema, contracture, and pain, as well as mobility and self-care issues. Often there is need
for surveillance of symptoms and function, both at critical points in care (e.g., in association with surgery) and over an extended period of time. One study of individuals with advanced breast cancer and remediable disabling impairments found that outpatients were markedly less likely than inpatients to receive rehabilitation services (31). This suggests the need for improved rehabilitation systems for outpatients and perhaps especially for those with advanced disease. Home health care may be needed if mobility is a significant obstacle to treatment.

Increasing attention is being paid to appropriate models for outpatient care, especially survivorship care. The Institute of Medicine has recommended that all cancer patients receive a care plan at the end of treatment summarizing care and also detailing future concerns. While many of the issues such as surveillance for recurrence or new cancers, medical late effects, reproductive issues, genetic testing, and economic factors, are beyond the typical rehabilitation scope, other surveillance issues may be highly relevant for good functional outcomes and clearly do pertain to rehabilitation care. For example, some physical impairments, such as contracture or lymphedema, can occur as late effects that warrant ongoing physiatrist’s assessment and management. In addition, the beneficial effects of physical activity are becoming increasingly convincing, both for cancer survival and reducing disability from comorbidities (32). The clinical focus and program structure of rehabilitation services for survivors vary widely but they tend to be most readily found, either as part of a comprehensive multidisciplinary outpatient cancer clinic or through referrals for outpatient consultations by physiatrists specializing in cancer care (33).

Consultation During Acute Care
In the acute care setting, consultation is most frequently requested for evaluation and treatment of mobility and self-care needs, as well as for the assessment of cognitive status, communication, and swallowing. The physiatrist will be asked to participate in decision making about the setting for future rehabilitation efforts. Services for pain control or the provision of orthotic/prosthetic devices may also be indicated. One study administering the FIM instrument to acute oncology inpatients found that 87% of patients had rehabilitation needs on admission and 84% still had needs upon discharge (16). Another study applying organized interdisciplinary rehabilitation care to oncology inpatients reported significant functional improvement per Barthe Mobility Index and Karnofsky Performance Scale (34).

Precautions
The physiatrist should routinely check for the following conditions, some a result of treatment such as chemotherapy, which empirically have an impact on the ability of the patient to safely tolerate some rehabilitation services, such as exercise or therapeutic heat.

1. Hematologic profile: hemoglobin less than 7.5 g, platelets less than 20,000, white blood cell count less than 3,000
2. Metastatic bone disease (see the section “Bony Metastatic Disease”)
3. Compression of a hollow viscous (bowel, bladder, or ureter), vessel, or spinal cord
4. Fluid accumulation in the pleura, pericardium, abdomen, or retroperitoneum associated with persistent pain, dyspnea, or problems with mobility
5. CNS depression or coma, or increased intracranial pressure
6. Hypokalemia/hyperkalemia, hyponatremia, or hypocalcemia/hypercalcemia
7. Orthostatic hypotension
8. Heart rate in excess of 110 beats/min or ventricular arrhythmia
9. Fever greater than 101°F

CANCER-RELATED PAIN

General Approach to Assessment and Treatment
An estimated 60% of patients with cancer experience pain, with 25% to 30% having severe pain (35). Presence of pain, as well as other symptoms such as fatigue and insomnia, is associated with decrease in functional status, particularly in elderly cancer patients (12).

The World Health Organization (WHO) analgesic ladder, which has been validated and is considered the cornerstone of cancer pain management, matches treatment to the pain intensity. The first line of treatment is the nonopioid analgesics (aspirin, acetaminophen, and nonsteroidal anti-inflammatory agents, etc.). If insufficient, an opioid (codeine, oxycodone, morphine, fentanyl, methadone, etc.) should be added. In addition to intensity, one must consider multiple other factors, including acuity (acute, crescendo, chronic), pathophysiology (somatic, visceral, neuropathic), and temporal (continuous, intermittent, breakthrough) (35,36). Visceral pain is typically poorly localized, cramping, or deep aching. Somatic pain is well localized to discrete anatomic areas, often sharp or stabbing, and neuropathic pain has a burning, tingling, or throbbing quality. While the WHO ladder remains fundamental, increasing attention is being paid to other treatments, such as early use of interventional procedures when clinical assessment suggests a high chance of success, not just when all other measures have failed. Medication regimens should be tailored to specific pathophysiologic pathways. For example, when pain is due to direct tumor spread, antitumor therapy is most likely to be effective. Edema or antibody-mediated neurologic compromise is often managed with corticosteroids, inflammatory pain with nonsteroidal anti-inflammatory medication or corticosteroids, and neuropathic pain with antidepressants, anti-convulsants, and topical preparations (36).

Pain intensity can be measured by numerical (1 to 10 rating), categorical (none, mild, moderate, severe), or pictoral (Wong-Baker FACES) methods (36). Historically, complete pain relief has been the goal, and even described as a “patient right,” but there is increasing recognition that it may not
always be possible, and that in most cases 33% to 50% pain reduction is clinically meaningful (37). Factors associated with difficulty attaining adequate pain control include neuropathic quality, psychologic distress, history of addiction, and impaired cognition (35).

Patient wishes should be included in the treatment plan, and use of a pain dairy can assist in optimizing treatment (37). Technological innovations such as an interactive computer program for education about pain and other symptoms have also been developed (38).

**Opioid Strategies**

Opioid agents that are commonly prescribed in the setting of cancer include oxycodone, morphine, hydromorphone, and fentanyl (Table 44-3). Meperidine and propoxyphene should be avoided due to toxic metabolites that can lead to seizures or cardiac arrhythmias, especially in the setting of dehydration or renal dysfunction (36). Methadone may be desirable in the setting of renal failure; however, because of its high potential for interaction with other medications and marked individual variation in pharmacokinetics, it should only be prescribed by physicians highly experienced with this drug (35). While oral administration predominates in most physiatric settings, increasing options have become available including parenteral routes such as transdermal, epidural, and intrathecal administration (39). In general, dosing is advanced to the level at which pain is controlled or at which toxicities preclude higher dosing. Daily effective dose can be established with short-acting preparations, and then converted to longer-acting forms. There should be additional dosing available for breakthrough, intermittent, or incident pain (including that associated with rehabilitation therapies) consisting of the equivalent of a patient’s 4 hours dosing needs, 25% to 50% of that dose, or 5% to 10% of the total daily opioid dose (35).

Management of side effects is crucial. An effective bowel program, including stool softeners and laxatives should be prescribed. Sedation is often transient, but if persisting greater than 1 week, measures such as caffeine intake or use of a stimulant such as methylphenidate can be helpful (36). However, in the setting of delirium, a neuroleptic may be necessary after other

**TABLE 44.3 Pharmacologic Management of Pain**

<table>
<thead>
<tr>
<th>Analgesic</th>
<th>Route</th>
<th>Duration of Analgesic</th>
<th>Dosage</th>
<th>Side Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aspirin</td>
<td>Oral</td>
<td>4–6 h</td>
<td>650 mg every 4 h</td>
<td>Gastritis, tinnitus</td>
</tr>
<tr>
<td>Acetaminophen</td>
<td>Oral</td>
<td>4–6 h</td>
<td>650 mg every 4 h</td>
<td>Hepatotoxicity</td>
</tr>
<tr>
<td>NSAIDs</td>
<td>Oral</td>
<td>Varies by agent</td>
<td>Varies by agent</td>
<td>Gastritis</td>
</tr>
<tr>
<td>Tramadol</td>
<td>Oral</td>
<td>6–8 h</td>
<td>50–100 mg every 6 h</td>
<td>Sedation, nausea, constipation</td>
</tr>
<tr>
<td>Morphine</td>
<td>Intravenous</td>
<td>1.5–2 h</td>
<td>2–10 mg</td>
<td>Sedation, respiratory depression, constipation, confusion, pruritus</td>
</tr>
<tr>
<td></td>
<td>Epidural/intrathecal</td>
<td>Up to 24 h</td>
<td>5 mg</td>
<td></td>
</tr>
<tr>
<td>Delayed release MS</td>
<td>Oral</td>
<td>2–4 h</td>
<td>15–60 mg</td>
<td>Sedation, respiratory depression, constipation, confusion, pruritus</td>
</tr>
<tr>
<td>Methadone</td>
<td>Oral</td>
<td>8–12 h</td>
<td>15–60 mg</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Oral</td>
<td>24 h</td>
<td>Varies</td>
<td></td>
</tr>
<tr>
<td>Oxycodone</td>
<td>Oral</td>
<td>3–6 h (standard)</td>
<td>5–10 mg every 4–6 h</td>
<td>Sedation, respiratory depression, constipation, confusion</td>
</tr>
<tr>
<td>Hydromorphone</td>
<td>Oral</td>
<td>12 h (sustained release)</td>
<td>7.5 mg</td>
<td>Sedation, respiratory depression, constipation, confusion</td>
</tr>
<tr>
<td>Hydrocodone</td>
<td>Parenteral</td>
<td>2–4 h</td>
<td>1.5 mg</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Rectal</td>
<td>6–8 h</td>
<td>3 mg</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Oral</td>
<td>3–5 h</td>
<td>30 mg</td>
<td></td>
</tr>
<tr>
<td>Fentanyl</td>
<td>Transdermal</td>
<td>72 h</td>
<td>50 μg/h</td>
<td>Sedation, respiratory depression, constipation, confusion</td>
</tr>
<tr>
<td></td>
<td>Transmucosal (buccal)</td>
<td>4 h (variable)</td>
<td>200 μg</td>
<td>Use transmucosal form only in opioid tolerant patients, for breakthrough; only for cancer patients</td>
</tr>
</tbody>
</table>

*NSAIDs, nonsteroidal anti-inflammatory drugs. Numerous options, including COX-2 inhibitors (rofecoxib, celecoxib) with reduced incidence of gastritis.
*Dosing of this agent and other opioids will be highly variable, depending on degree of opioid tolerance. Dosing can be advanced.
metabolic causes have been excluded. Myoclonus related to opioid use may respond to baclofen, benzodiazepines, dantrolene, or valproate (35). While tolerance of a particular opioid may develop, reduced cross tolerance between different agents makes rotation of opioid drugs an effective way of avoiding escalating dosage requirements and the resulting side effects (35).

Nonpharmacologic Pain Management Approaches

Physical modalities such as cryotherapy, biofeedback, iontophoresis, transcutaneous electrical nerve stimulation, and massage are well tolerated and believed to be safe, though the latter two are not performed directly over areas with known tumor (39). Deep heat such as ultrasound is contraindicated directly over an area of tumor. Data are limited, but one study in mice showed increase in tumor size (but no increase in rate of metastasis) with application of ultrasound (40). Routine physiatric procedures such as trigger point injections may be helpful. Psychologic techniques including imagery, distraction training, relaxation techniques, and coping strategies are encouraged (36). Interventional options can include nerve blocks, vertebroplasty, spinal analgesia (including long-term catheter systems), dorsal column stimulators, and neuroablative procedures (neurectomy, rhizotomy, cordotomy). Complementary and alternative medicine strategies are widely used, with increasing acceptance of massage and acupuncture, especially when other modalities have failed to achieve adequate pain relief (39).

BONY METASTATIC DISEASE

Metastatic disease to the skeleton is one of the most problematic situations for clinicians managing musculoskeletal disorders. The skeleton is the third most common location for systemic metastatic disease (41) (Fig. 44-1). Breast, lung, prostate, kidney, and thyroid cancers account for 80% of malignancies to bone (42). The biology of bone metastasis is thought to involve the interaction of cellular adhesion molecules with the architecture and circulatory supply of the bony apparatus (43). Bone metastases are osteolytic (primarily osteoclastic activity), osteoblastic (primarily osteoblastic activity), or mixed (43). Lymphoma, multiple myeloma, thyroid and renal cell malignancies have the highest rates of osteoclastic activity and therefore high levels of structural damage to bone and fracture risk. However, even in conditions where osteoblastic changes predominate, such as prostate cancer, pathologic fractures can occur. Early and aggressive management is imperative in maintaining function (44).

Pain is the most common clinical presentation of bone metastases (45). The pain is insidious, unrelenting, not associated with trauma or activity, and may be present or intensify at rest (46). The pain is frequently located in less common locations such as the thoracic spine or femoral shaft. Although pain is a common presentation, more than 25% of bone metastases are asymptomatic and found on routine imaging. Classic findings on physical examination include weight loss, exquisite point tenderness over the involved bone, and possibly neurological impairment. Failure to respond to initial treatment and progressive symptoms are “red flags” that require further scrutiny (47).

The assessment of patients with suspected bone metastasis requires an efficient structured approach, including a detailed history and physical examination. Functional assessment and social history are imperative for establishing rehabilitation goals and the need for family support. Initial laboratory evaluation in those with suspected metastatic disease should include a complete blood count, serum protein electrophoresis, urinalysis, C-reactive protein, and a comprehensive metabolic panel including calcium and alkaline phosphatase (47). Plain radiographs, though inexpensive and easily accessible, have limited utility in identifying metastatic bone disease because greater than 50% of the cortex needs to be involved before metastatic disease will be identified (47). The most sensitive imaging study for the identification of bone metastases is the triple phase bone scan, as only 5% to 10% of cortical involvement is required to identify abnormalities (48). Bone scans identify osteoblastic activity in bone and therefore may produce normal results (false negatives) in patients with primarily osteolytic disease such as myeloma or lymphoma. In addition, bone scans have poor specificity. For those patients with localized bone pain, equivocal bone scans, or neurological impairment, magnetic resonance imaging (MRI) with gadolinium is the most appropriate test, particularly for suspected spinal lesions (49). The recent advent of PET scanning has helped to detect tumor activity in cases when the above imaging studies are equivocal or when the primary lesion is osteolytic (50). In some cases, biopsy may be indicated to guide treatment (46).

The median survival for patients with isolated bone disease from cancer of the breast, prostate, or from multiple myeloma is 21 to 33 months (51). During this time, the appropriate use of supportive measures to decrease morbidity and pain, and improve function should be employed. Multidisciplinary
management involves collaboration among physiatry, orthopedic surgery, medical and radiation oncology, with care goals that encompass systemic disease management, pain control, skeletal stabilization, and rehabilitation.

Systemic management options, usually prescribed by a medical oncologist, include chemotherapy, hormonal therapy, monoclonal antibodies, and anti-angiogenesis agents. The administration of bisphosphonates is usually initiated when bone metastases are first detected, although in some cases they may be administered prophylactically. Intravenous bisphosphonates decrease skeletal morbidity, fracture rate, and pain through the inhibition of osteoclastic activity and suspected modulation of local tumor activity (52,53). Radiation therapy, including direct beam and radiopharmaceutical options, frequently can be effective in decreasing local tumor burden and controlling pain (54).

Multimodal pain control begins with the management of systemic malignancy as described above. Nonsteroidal anti-inflammatory agents are employed to decrease periosteal bone reaction, and opioids are used for general pain control (48). In some cases, more aggressive interventional measures may be necessary.

Stabilization of the skeleton is imperative for pain control and function. No system for predicting stability of long bone has been universally accepted. Generally, the greater the amount of cortex involved with metastatic disease the greater the risk of fracture (55). Size criteria for pathologic fracture risk in lower limb long bone include lesions measuring more than 2.5 cm, involvement of more than 50% of the bony cortex, and the Mirels scoring system incorporating pain, size, location, and radiographic appearance (49) (Table 44-4). A recent study comparing various methods found only axial cortical involvement of greater than 30 mm and circumferential cortical involvement predictive of fracture; the former measure has the advantage of being more readily reproducible (56). In practice, it is often difficult to gauge the size of bony lesions, especially lytic ones, which may be irregular, permeative, and difficult to distinguish from surrounding osteopenia. Apart from radiographic assessment, pain that increases with weight bearing may be an indicator of an unstable bony structure (46), warranting early surgical assessment. Fracture risk may actually increase during the first 6 to 8 weeks after radiation as a result of tumor necrosis and softening of bone. Therefore, surgical stabilization is typically done prior to radiation of unstable lesions.

In the spine, the stability of the bone and the presence of neurological impairment guide the assessment in Harrington's classification of vertebral metastases:

I. No significant neurological involvement
II. Involvement of bone without collapse or instability
III. Major neurological impairment without significant bone involvement
IV. Vertebral collapse without neurologic impairment
V. Vertebral collapse with neurologic impairment

For class III to V involvement, surgical intervention is warranted (54). Surgical management of the unstable skeleton is very effective in reducing pain and increasing function. Harrington found good or excellent pain relief in 96% of long bone and 88% of spinal fractures (57), and improved function in 82% of spine stabilization cases (57). A recent advance has been the NOMS algorithm, incorporating neurologic (cord compression), oncologic (radiosensitive or not), mechanical (movement-related pain; fracture/subluxation >5 mm, or angulation >11 degrees with subluxation >3.5 mm), and systemic factors (medical risks of surgery) into decision-making for surgery (58). In refractory cases or nonsurgical candidates, bracing can be considered, but wearing tolerance remains a significant barrier.

The rehabilitation of patients with bone metastasis is based on protection, pain control, energy conservation, and maintenance of function. Protection and pain control can be obtained through the use of bracing, mobility aids, and activity precautions. Some patients with exclusively lower extremity disease may be able to maintain mobility with the use of a cane or walker. Those with more diffuse (including upper limb), or bilateral disease may require a wheelchair or power mobility. Neutral spine techniques preserve function and minimize pain in patients with spinal metastases. It is essential to assess the weight-bearing status of all limbs when prescribing assistive devices for patients with known or suspected bone metastases as bony metastases usually occur at multiple sites, with 20% of metastases present in the upper limbs, especially the humerus (49). Exercise prescriptions should focus on increasing strength, endurance, and function with minimal loading or torsion of the affected bone. A typical exercise program may include aquatic therapy, non-weight-bearing exercise such as cycling and isometric exercise for strength maintenance. Compensatory techniques can decrease the biomechanical load on affected bones and maximize function. These include the use of reachers for activities of daily living, neutral spine techniques, and a step-to-gait pattern when climbing stairs. When metastatic disease limits independence, family training and education are beneficial to reduce the risk of injury to both caregiver and patient, and to identify needs for durable medical equipment.

<table>
<thead>
<tr>
<th>TABLE 44.4 Fracture Risk (&gt;8 Points High Risk)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Points Assigned</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>Anatomic site</td>
</tr>
<tr>
<td>Lesion type</td>
</tr>
<tr>
<td>Lesion size</td>
</tr>
<tr>
<td>Intensity of pain</td>
</tr>
</tbody>
</table>
CANCER-RELATED FATIGUE (CRF)

Fatigue is a normal physiological response to exertion. It becomes pathological when it persists, occurs during routine activities, and does not respond to rest (59,60). Stringent clinical studies routinely find that the majority of cancer patients will meet criteria for CRF at more than one time during their disease continuum (61). High prevalence, impact on function and quality of life, and caregiver burden make the assessment and treatment of CRF a central goal of almost every cancer rehabilitation program (62,63).

Numerous fatigue assessment tools have been validated for oncology patients (64). Busy clinicians, however, may find it easiest to screen their patients using a mild/moderate/severe designation based on a 0 to 10 Likert type scale. Patients reporting fatigue intensity of 1 to 3 are considered as having mild CRF, 4 to 6 as moderate, and 7 to 10 as severe. The National Comprehensive Cancer Network (NCCN) recommends screening for fatigue at the time of diagnosis, consistently during treatment, and as part of the long-term follow-up care (65), even after the completion of successful oncologic treatment.

It is not yet clear whether CRF is a specific physiologic process or “a final common pathway to which many predisposing or etiologic factors contribute” (66,67). Nonetheless, clinical studies have been able to recognize specific factors that are consistently associated with CRF and are therefore thought to precipitate it or intensify its impact. The most common associated factors are pain, emotional distress, sleep disturbance, anemia, nutritional deficiencies, deconditioning, and medical comorbidities (65) (Table 44.5). Identification of lead factors guides the treatment process.

Successful management of CRF requires the coordinated collaboration between clinicians who can address etiologic factors affecting a specific patient. The NCCN guidelines recommend four types of treatment interventions: (a) education and counseling, (b) general strategies, (c) nonpharmacologic, and (d) pharmacologic (65). Since CRF affects patients throughout the cancer continuum, the NCCN provides guidelines for three types of patients: (a) patients on active treatment, (b) patients on long-term follow-up, and (c) patients at the end of life (65).

General education about the nature and management of CRF provides reassurance and leads to earlier recognition and mitigation of its effects. General strategies, as opposed to cause-specific interventions, are intended to minimize the impact and intensity of existing CRF after reversible causes have been addressed. Energy conservation strategies developed for cardiac and pulmonary patients are equally effective for patients with CRF (68). Beneficial interventions include strengthening and endurance programs, psychosocial interventions, nutritional management, and sleep optimization.

The prevalence of disrupted sleep among persons with cancer diagnoses makes it one of the most common and intuitively obvious factors in CRF (69). Many cognitive and behavioral strategies exist for promoting restorative night time sleep and minimizing daytime sleepiness (69). Addressing the underlying anxiety and depression often improves sleep as does increasing physical activity. Sleep can also be addressed with judicious use of pharmacological agents.

Anemia is a common cause of CRF that responds to medical management. Blood transfusions may be useful for more rapid correction of profound anemia, particularly after tumor resection or myeloablative chemotherapy. Several large scale studies demonstrated the utility of erythropoietin for both increasing hemoglobin and reducing fatigue scores in patients with anemia associated with chemotherapy (70). However, the use of erythropoietin in cancer patients was reassessed in light of data concerning increased risk of thrombotic events in dialysis patients receiving erythropoietin (71). Some studies have also shown a decreased survival rate in cancer patients treated with erythropoietin that was not associated with thrombotic events (72). Recent data suggest that target hemoglobin levels of 12 g/dL can confer symptomatic benefit without increasing risk (65).

In addition to medical management of factors contributing to CRF, physicians may use a range of prescription medications to treat CRF directly. Psychostimulants such as methylphenidate and modafinil have been used to treat CRF but their efficacy has not been definitively established (73–76). Corticosteroids are used for numerous purposes in cancer patients and may modulate CRF (77).

Numerous studies have explored the safety and efficacy of exercise interventions for patients with CRF. A meta-analysis

### TABLE 44.5 Interventions for Fatigue

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medications</td>
<td>Nutritional and vitamin supplementation</td>
</tr>
<tr>
<td></td>
<td>Correct endocrine dysfunction (thyroid)</td>
</tr>
<tr>
<td></td>
<td>Stimulants (methylphenidate, d-amphetamine)</td>
</tr>
<tr>
<td></td>
<td>Analgesics</td>
</tr>
<tr>
<td></td>
<td>Antidepressants (bupropion, SSRIs, TCAs)</td>
</tr>
<tr>
<td>Exercise</td>
<td>Regulate sleep/wake</td>
</tr>
<tr>
<td></td>
<td>Glucocorticoids</td>
</tr>
<tr>
<td></td>
<td>Investigational—cytokine-targeted therapy (including NSAIDs)</td>
</tr>
<tr>
<td>Energy conservation</td>
<td>Education</td>
</tr>
<tr>
<td></td>
<td>Adaptive equipment</td>
</tr>
<tr>
<td>Psychologic/coping</td>
<td>Recreational activities</td>
</tr>
<tr>
<td></td>
<td>Relaxation techniques</td>
</tr>
<tr>
<td></td>
<td>Support groups</td>
</tr>
<tr>
<td></td>
<td>Spiritual supports, participation</td>
</tr>
</tbody>
</table>

SSRIs, selective serotonin reuptake inhibitors; TCAs, tricyclic antidepressant agents; NSAIDs, nonsteroidal anti-inflammatory drugs.
identified the greatest statistically significant impact in studies limited to a specific disease population such as breast cancer with less effect demonstrated when patients with heterogeneous diagnoses were recruited (78). Even when exercise interventions do not directly reduce fatigue scores, they nonetheless play an important role in the management of CRF by stemming the cycle of deconditioning that occurs as patients with CRF reduce their activity.

Cancer patients should be screened by a physician prior to the prescription of a moderate intensity exercise program. Clinical literature supports the use of therapeutic exercise for CRF in patients with stage I to III disease but fewer guidelines exist for patients with advanced disease, particularly bony metastases. Walking has been the most frequently used mode of exercise in studies assessing the effect of physical activity on CRF although a recent investigation studied the effect of higher intensity training utilizing cycle ergometry (79). None of the studies reviewed reported any serious adverse events related to the prescribed exercise program. Strength training has been less well studied than aerobic exercise in oncology populations. One randomized prospective study of 155 men with prostate cancer receiving androgen deprivation therapy demonstrated significant reduction in fatigue as measured with the FACT-F scale after participation in a 12-week resistance exercise program (80).

Overall, moderate intensity aerobic exercise can help ameliorate CRF during and after treatment and minimize its effect on the overall function and quality of life. Strength training is beginning to show positive effects as well for fatigue management although there are fewer studies supporting its role. It is best if exercise interventions are prescribed and administered by clinicians who have experience with oncologic patients and are familiar with relevant precautions and contraindications. Exercise studies targeting CRF in populations with comorbidities as well as metastatic disease are needed to enable safe and effective exercise recommendations for these patients.

Effective management of CRF requires a holistic approach that is best achieved through an integrated, interdisciplinary team including physicians, oncology nurses, physical, occupational, and speech language therapists, nutritionists and psychologists. M.D. Anderson pioneered a dedicated fatigue clinic and demonstrated high rates of patient satisfaction as well as some clinical improvement (62). Less formal CRF programs have been established elsewhere and are easily integrated into physiatric practices with established access to interdisciplinary team members who have oncologic experience.

EXERCISE FOR THE CANCER PATIENT

General Aspects

Exercise studies performed in the cancer population have consistently substantiated gains in numerous parameters, including cardiopulmonary fitness, fatigue (see “Cancer-Related Fatigue” section), quality of life, depression, and anxiety (81–83). There may also be benefits of exercise on immune function, such as improved natural killer cell activity, monocyte function, proportion of circulating granulocytes, and duration of neutropenia; however, the implications of these effects are not well delineated on a clinical level, and not all studies have shown immune effects of exercise (84). This section focuses on general exercise principles in the setting of cancer; activity recommendations related to specific impairments will be covered in the respective sections.

Of note, physical activity appears to exert a protective effect against the development of some types of cancers, most notably colon and breast cancers (85). Obesity has been associated with increased risk of death from cancer in both men and women, with one large prospective study estimating effect of overweight and obese status as accounting for 14% of cancer deaths in men and 20% of those in women, with the strongest associations seen in gastrointestinal, kidney, breast, prostate, gynecologic, and some hematologic malignancies (86). Cancer prevention recommendations developed by the American Cancer Society include at least 30 minutes of moderately vigorous physical activity on 5 or more days of the week for adults (preferably 45 to 60 minutes), and at least 60 minutes for children and adolescents (87). Physician recommendation to exercise has been shown to improve self-reported total exercise among newly diagnosed cancer survivors (88).

While exercise studies overall have best explored breast cancer and the aerobic form of exercise, wider evidence is beginning to emerge, among a range of cancer types, also with favorable outcomes (17,89,90). Among exercise forms, cycle ergometry is often favored, with its advantages of positioning options, and relative ease of use by individuals with balance or coordination deficits (91). Preferences for exercise programs have begun to be explored, with patients often expressing interest in home or fitness-center-based exercise, especially walking. However, at least one study suggests that a supervised program may be more likely to yield measurable gains in physical performance (92). Patients voice an openness to, and even preference for, technically based approaches for receiving information (internet or CD-ROM). Many perceive better readiness to start an exercise program after than during treatment (93,94).

Most hematologic parameters for exercise are empiric. The concern with exercise in the thrombocytopenic state lies in the potential for increased blood pressure, which occurs most dramatically with isometric exercise, to result in intracranial hemorrhage, and for high-impact activities to result in muscular or intra-articular hemorrhage. The risk of hemorrhage correlates with the platelet count, but is mitigated by other systemic factors. In a study of acute leukemia patients (95), grossly visible hemorrhage was rare with a platelet count greater than 20,000, and no intracranial hemorrhage occurred with a platelet count greater than 10,000. In general, unrestricted exercise can be pursued with platelet counts greater than about 30,000 to 50,000. Aerobic, but not resistive, activities can be considered with platelet counts greater than 10,000 to 20,000. Active therapy is not advocated with platelet counts less than 10,000.
Exercise for Patients Undergoing Chemotherapy or Postchemotherapy

Cancer patients treated with cardiotoxic agents such as anthracyclines can sustain permanent cardiac damage that affects physical performance. Patients treated with significant doses of these agents (>100 mg/m²) can have reduced exercise time, reduced maximal oxygen uptake, abnormal heart rate response, ST- and T-wave changes, and exercise-induced hypotension (96). However, exercise time, peak oxygen uptake, and ventilatory anaerobic threshold can still improve with an exercise program, due to peripheral adaptation, despite the fact that cardiac parameters, such as exercise heart rate and stroke volume do not increase (96). A controlled study of patients undergoing aerobic exercise (supervised daily, supine bicycle ergometry) during treatment found multiple benefits in the exercise group, including less decrement in performance per treadmill testing, less pain, decreased duration of neutropenia, and shorter hospitalization (97). A study of home-based unsupervised exercise among early-stage breast cancer patients who had completed treatment (about half with treatment including chemotherapy) found higher self-reported physical activity and 1-mile walk test performance than controls, but no difference in accelerometry or anthropomorphic measures (98). A supervised program among breast cancer patients, before and during treatment, focusing on both strength and aerobic training at 40% to 60% maximal twice weekly for 21 weeks, did result in improved lean body mass, reduced body fat, and improved strength per submaximal muscle endurance protocol (99). Another supervised program among breast cancer patients that had completed treatment, focusing on aerobic training three times a week for 15 weeks, found improved peak oxygen consumption in the exercise group but no significant anthropomorphic differences (100).

Exercise for Patients Undergoing Bone Marrow Transplantation

Exercise programs have been developed for bone marrow transplant recipients to counteract the debility occurring with medical morbidity and prolonged hospitalization, as well as to address other factors such as depression and social isolation. Supine or sitting exercise is generally well tolerated, but standing exercise should be attempted at least for brief periods to minimize gastroc soleus tightness. Supine exercise may be most comfortably performed with the head of the bed slightly elevated. Exercise programs emphasize range of motion, aerobic activity, such as walking or cycle ergometry, light resistive activities such as bridging and use of light weights, and deep breathing to prevent atelectasis and pneumonia (101). In one program, referral to physical therapy is placed when the patient is isolated to the room, or when unable to ambulate approximately 200 ft three times a day (101). For those in isolation, use of a sanitized bedside stationary bicycle or in-bed pedicycle facilitates activity (101). In those with graft versus host disease (GVHD), skin erythema and rash may occur. Protective padding can prevent pain and skin irritation during exercise, especially over the soles of the feet. Attention should be paid to strengthening, due to risk of steroid myopathy. Medications such as quinine, carbamazepine, or baclofen may be useful for cramping symptoms of GVHD (102). One study of treadmill training (interval pattern) 30 minutes daily for 6 weeks in bone marrow transplant recipients with stabilized platelet counts and clinical condition resulted in improved physical performance, measured by distance of treadmill walking (103). Another controlled study of stretching and treadmill training found improved preservation of strength, measured by dynamometer in multiple muscle groups among the exercise group at 6 weeks post-transplant (104).

Cachexia

Endogenous tumor necrosis factor (TNF), or that administered exogenously as antineoplastic therapy, can reduce skeletal muscle protein stores. A low-to-moderate intensity of exercise, which relies mainly on type 1 muscle fibers, which are fatigue resistant, should be encouraged, reducing activity at the onset of fatigue. While evidence is limited, empirically, for individuals with marked cachexia, rehabilitation efforts should focus on energy conservation and methods other than strenuous exercise to achieve functional goals.

NEUROLOGIC COMPLICATIONS OF CANCER

Neurological impairments have been identified in 30% to 46% of cancer patients (105). Neurological complications are the second most common reason for admission to the hospital, after admissions for routine chemotherapy (106). Patients present most commonly with low back pain, limb weakness, headaches, and mental status changes (105). Due to the high rate of associated disability and the vulnerability of the nervous system to reversible and irreversible damage, early identification and treatment can prolong life and diminish disability.

Metastatic Brain Disease

Brain metastases are the most common catastrophic neurological impairment in the cancer population. The incidence of metastatic disease to the brain is ten times greater than the incidence of primary tumors (107). Brain metastases occur most frequently with lung, breast, colorectal, melanoma, and genitourinary cancers (105), with lung and breast malignancies accounting for 60% of all lesions. Approximately 85% of brain metastatic lesions are found in the cerebrum and 15% in the cerebellum (108). The most common presenting complaint is a progressive headache, often worse when recumbent. Hemiparesis, seizures, and mental status changes occur frequently. Evaluation includes a complete neurological and clinical examination to exclude other etiologies such as stroke and central nervous system (CNS) infection. Gadolinium-enhanced magnetic resonance imaging (MRI) is the gold standard. Management includes early treatment with corticosteroids to decrease brain edema, and anticonvulsants to decrease seizure risk. Excision of brain metastasis may be indicated, especially if the metastasis is single, the cancer is otherwise well
controlled, and the lesion(s) appears to be the major factor limiting survival or quality of life. Clinical trials have shown that in certain patients a combination of surgery and radiation therapy is superior to either treatment alone (108). Whole brain radiation, with standard doses of 3,000 cGy given over ten treatment sessions, as well as adjuvant chemotherapy are important components of treatment. Intrathecal methotrexate administered through a reservoir secured beneath the scalp has been used to treat CNS and leptomeningeal disease with mixed results but does avoid some of the more toxic side effects of systemic methotrexate. Prognosis for solitary brain lesions and those able to maintain ambulatory function is good. Prognosis is poorer for those with persistent headache, visual dysfunction, or ongoing mental status changes (109). See “Brain Tumor” section for discussion of specific rehabilitation approaches.

Leptomeningeal disease is a result of the spread of malignancy to cerebrospinal fluid (CSF). It is most commonly associated with breast cancer, small cell lung cancer, and melanoma (108). 75% to 80% of cases involve spine, 50% cranial nerves and greater than 50% brain (110). Leptomeningeal disease causes back pain, radiculopathies, cranial nerve dysfunction, and mental status changes. The diagnosis is made by MRI with gadolinium and/or CSF analysis. Treatment is with intrathecal chemotherapy or radiation. The prognosis is uniformly poor.

Spinal Cord Involvement
Spinal cord compression due to metastatic disease occurs in 5% to 14% of all cancer patients (111). Twenty-five percent of patients with neoplastic spinal cord compression do not have a previous diagnosis of cancer (105). Metastases from prostate, breast, lung, and kidney primaries as well as multiple myeloma are the most common sources (112). The thoracic spine (70%) is the most frequent site of metastases followed by lumbar (20%) and cervical (10%). Metastases to the spine characteristically present as progressive, insidious back pain that is worse when recumbent along with associated neurological impairment (105,112). On average, patients have low back pain for 60 days before diagnosis (113). Clues to metastatic disease include a history of cancer, constitutional symptoms, thoracic level pain, and bowel or bladder dysfunction. Findings on physical examination may include point tenderness, paraparesis, a sensory level, and upper motor neuron findings (111). Motor abnormalities often precede sensory changes as a result of epidural extension preferentially affecting the anterior spinal cord, with recovery in the reverse order.

Early diagnosis is the key. Patients presenting without paresis have longer survival than those presenting with paresis (114). Corticosteroids (usually dexamethasone) are administered with a 100 mg loading dose, followed by 4 mg every 6-hours. Palliative radiation is indicated in those with metastatic disease and stable spinal structures.

Paraneoplastic Neuromuscular Disorders
Paraneoplastic neuromuscular disorders (PNM) are a remote effect of cancer caused by antibody formation by the primary tumor (115). Neuromuscular disorders can precede the diagnosis of cancer by years. These disorders are rare, seen in only 0.01% of cancer patients (116). The most common tumor associated with PND is small cell lung cancer accounting for 50% to 75% of cases (116). While neuromuscular junction disorders will be the main focus of this section, various other paraneoplastic phenomena can occur, of which the physiatrist should be aware. These disorders include cerebellar degeneration, organic dementia, metabolic encephalopathy from electrolyte disturbance, and orthostasis from autonomic instability.

Lambert-Eaton myasthenic syndrome (LEMS) is found in 3% of patients with small cell lung cancer (116). The disorder results from presynaptic inhibition of calcium release at the neuromuscular junction. Proximal weakness, autonomic dysfunction, and improvement with exercise are common clinical findings. Diagnosis is established by electromyography and nerve conduction (EMG/NCS) studies. Myasthenia gravis (MG) is a postsynaptic neuromuscular junction disorder that is a result of autoimmune degradation of the postsynaptic membrane. MG is found in 15% of patients with a diagnosis of thymoma (116). The hallmarks of MG are proximal, disconjugate gaze, fatigue on upward gaze, and proximal weakness worsened by activity. Electrodiagnostic studies, antibody analysis, and an anticholinesterase challenge are used to confirm diagnosis.

Paraneoplastic subacute neuropathies include sensory, sensorimotor, and demyelinating forms (115). The clinical presentation is that of rapid (days to weeks) onset of peripheral neuropathy affecting various fiber types. The sensory and sensorimotor involvement is most commonly associated with small cell lung cancer, whereas demyelinating processes are associated with lymphoma. Monoclonal paraproteins create neuropathies in multiple myeloma, osteosclerotic myeloma, Waldenström's macroglobulinemia, amyloidosis, and γheavy-chain disease. Typically, the neuropathy is distal, mixed sensory and motor, with axonal loss and segmental demyelination. The management of PND includes treatment of the primary tumor and immune mediation (116).

Polyneuropathy
Polyneuropathy can result from numerous factors in the cancer patient, including nutritional deficiencies, paraneoplastic disorders (as discussed above), and medical comorbidities. However, the most common cause is chemotherapy-induced polyneuropathy (CIP). Dose-dependent toxicity occurs in peripheral nerves and is commonly associated with a handful of agents (117) (Table 44-6). The pathophysiology of CIP includes disruption of axoplasmic microtubule transport, axonal “dying back,” and direct effects to the dorsal root ganglion (117). The sensory nerves are typically more affected because of the smaller fibers and the location of the dorsal root ganglion outside the blood-brain barrier. The onset of neuropathy coincides with the administration of chemotherapy, and should not progress after treatment has stopped.
Symptoms of dysesthesias, sensory loss, and allodynia typically begin in the foot and ascend. Motor weakness and autonomic dysfunction are delayed findings and may be a sign of significant toxicity. The differential diagnosis of CIP includes nutritional deficiencies, paraneoplastic disorders, mononeuropathies, radiculopathy, myelopathy, and brain disorders. Definitive diagnosis can be obtained through the use of electrophysiological studies. An accurate diagnosis of CIP is required to avoid unnecessary discontinuation of chemotherapy. The management of CIP includes alteration of dosing, and symptom management with medications (see "Cancer-Related Pain" section). Several agents have been evaluated as potential chemoprotective agents but none have proven beneficial (118). Rehabilitation principles include education, safety awareness, gait and proprioceptive training, and the prescription of orthoses and assistive devices for those with motor as well as sensory deficits.

**RADIATION-INDUCED TISSUE DAMAGE**

Radiation therapy (RT) is commonly used for curative or palliative intent. For those tumors that are significantly radioresponsive (seminoma, head and neck cancers, and prostate cancer), it may be used as the sole treatment. In many cancers such as breast, colorectal, lung, lymphoma and melanoma, radiation is used in concert with surgical resection or chemotherapy. Systemic spread of cancer can be managed with palliative RT to reduce tumor burden, pain, alteration of dose, and neurological impairment. RT is typically given via external beam, brachytherapy (local radioactive-implanted materials), or via stereotactic equipment (in certain brain tumors). The radiation is given in fractions (individual doses) over a set period of time (treatments). The overall effects of radiation are a result of individual dose, cumulative dose, and number of treatments (119). The recent advent of three-dimensional techniques, intensity modulated techniques, and brachytherapy have helped to localize the radiation to the tumor site and reduce surrounding soft-tissue effects (120). In addition, hyperfractionated regimens have helped to decrease acute complications from intense dosing.

The local effects of RT can lead to many complications (Table 44-7) (121). Soft-tissue restriction, swallowing dysfunction, and neurological impairments are the most commonly encountered issues for rehabilitation physicians. Swallowing dysfunction will be covered in the "Head and Neck Cancer" section. Soft-tissue fibrosis as a direct result of RT results in dermal fibrosis, musculotendinous contraction, and joint restriction with resulting loss of range of motion and function. The most commonly affected joints are the glenohumeral joint after axillary RT for breast cancer, the femoroacetabular joint after RT for cutaneous malignancies, and the neck after treatment for head and neck cancer. The loss of range of motion may start with guarding from painful, irradiated skin and is followed by contraction of the musculotendinous unit and joint fibrosis. Prevention through meticulous skin care, soft-tissue mobilization, and range of motion techniques is the primary means of management. The use of pentoxifylline may be beneficial to restore microvascular supply to the soft-tissue structures and thus help preserve mobility (119). The importance of delayed

<table>
<thead>
<tr>
<th>Tumor</th>
<th>Agents</th>
<th>Trade Name</th>
<th>Nerve Fibers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breast</td>
<td>Taxanes</td>
<td>Taxol/Taxotere</td>
<td>Sensory &gt; motor</td>
</tr>
<tr>
<td>Lung</td>
<td>Taxanes</td>
<td>Taxol</td>
<td>Sensory &gt; motor</td>
</tr>
<tr>
<td>Ovarian</td>
<td>Platinum</td>
<td>Carboplatin</td>
<td>Pure sensory</td>
</tr>
<tr>
<td>Myeloma</td>
<td>Taxanes</td>
<td>Carboptatin</td>
<td>Pure sensory</td>
</tr>
<tr>
<td>Lymphoma</td>
<td>Thalidomides</td>
<td>Carboplatin</td>
<td>Pure sensory</td>
</tr>
<tr>
<td>Colon</td>
<td>Vinca alkaloids</td>
<td>Oxiplatin</td>
<td>Pure sensory</td>
</tr>
</tbody>
</table>

**TABLE 44.7 Radiation Therapy Side Effects**

<table>
<thead>
<tr>
<th>Acute</th>
<th>Delayed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fatigue</td>
<td>Soft-tissue fibrosis</td>
</tr>
<tr>
<td>Nausea</td>
<td>Skin atrophy</td>
</tr>
<tr>
<td>Vomiting</td>
<td>Auditory changes</td>
</tr>
<tr>
<td>Anorexia</td>
<td>Pulmonary fibrosis</td>
</tr>
<tr>
<td>Skin erythema</td>
<td>GI stricture</td>
</tr>
<tr>
<td>Desquamation</td>
<td>Thyroid dysfunction</td>
</tr>
<tr>
<td>Mucositis</td>
<td>Brain necrosis</td>
</tr>
<tr>
<td>Xerostomia</td>
<td>Myelitis</td>
</tr>
<tr>
<td>Taste loss</td>
<td>Plexopathy</td>
</tr>
<tr>
<td>Proctitis</td>
<td>Lymphedema</td>
</tr>
<tr>
<td>Cystitis</td>
<td>Secondary malignancies</td>
</tr>
<tr>
<td>Decreased libido</td>
<td>Osteonecrosis</td>
</tr>
<tr>
<td>Sterility</td>
<td>Amenorrhea</td>
</tr>
<tr>
<td>Hematological changes</td>
<td></td>
</tr>
</tbody>
</table>
CHAPTER 44 | REHABILITATION FOR PATIENTS WITH CANCER DIAGNOSES

LYMPHEDEMA

The lymphatic system is a component of the circulatory system designed to mobilize cellular debris, metabolic byproducts, and infectious material (124). The objective is to transport these particles to local and regional lymph nodes for degradation and to return the fluid to the central vasculature. The peripheral lymphatic system is divided into superficial and deep lymphatics (124). The superficial lymphatics are web-like and course directly under the skin. The deeper lymphatics run alongside the vascular system. The lymphatic system mobilizes fluid slowly via muscle contraction and respiration toward the centrally located lymph nodes in the groin and axilla. After lymph node processing, the fluid drains into the lymphatic ducts and central vasculature.

Lymphedema is a disorder of lymphatic fluid accumulation resulting in swelling, typically in an extremity but it may also involve the face, thorax, or abdomen. The most common malignancies associated with lymphedema include breast, melanoma, gynecological malignancies, and lymphoma. The first three disorders are associated with lymphedema, primarily because staging and treatment of these malignancies includes lymph node dissection. In breast cancer, a full axillary dissection typically involves removal of lymph nodes below the axillary vein, whereas in melanoma all regional lymph nodes below the clavicle are removed (124). Sentinel lymph node dissection, which includes sampling of lymph nodes that drain directly from the site of the tumor, leads to smaller numbers of lymph nodes being removed and reduces but does not eliminate risk of lymphedema (125,126). Lymphoma is associated with lymphedema secondary to direct invasion of lymph nodes. The risk of lymphedema approaches 25% in breast cancer patients. Associated risk factors include axillary lymph node dissection (ALND), metastatic involvement of lymph nodes, radiation therapy, obesity, and increasing age (127).

Lymphedema usually presents as painless, gradual, colorless swelling with heaviness and loss of limb contour. For patients presenting with rapid onset of swelling, with pain or erythema, one must exclude metastatic seeding (malignant lymphedema), infection, and thromboembolic disease. Two types of presentation are seen in patients with lymphedema. The first type has a gradual onset of fluid accumulation after lymph node removal secondary to overload of the lymphatic system with normal daily fluid production (normal output lymphedema). The second presentation consists of an acute increase in lymphatic fluid that overwhelms the compromised system (high output lymphedema). Common causes of high
output lymphedema include cellulitis, lymphangitis, trauma, thermal burns, and increased metabolism in the at-risk limb (Table 44-8). Lymphedema progresses in stages. The first stage is a primarily fluid stage in which arm volumes diminish with elevation and use of external compression. As lymphatic fluid accumulates in the extremity, an inflammatory reaction ensues that results in subcutaneous fibrosis and hardening of tissue. This is considered the hallmark for stage two lymphedema. Stage two lymphedema does not resolve with elevation and compression garments. Stage three lymphedema is identified by cutaneous fibrosis and verrucous hyperplastic changes of the skin and is rarely seen in the upper extremity (124). As lymphedema progresses, pain may develop secondary to constriction of underlying soft-tissue structures and overload of the supporting structures such as the shoulder (128) (Fig. 44-2).

The clinical evaluation of lymphedema should include a history of the primary malignancy and the aspects of treatment that have impacted the lymphatic system. The onset, duration, and progression of lymphatic swelling should be identified. In addition, identifying attempted management methods is important. Functional restrictions related to lymphedema (loss of shoulder or hand function) should be identified. Full musculoskeletal and neurological examination is needed to identify any underlying deficits. The skin examination should include general skin and soft tissue characteristics, and a specific focus on surgical and irradiated sites. In addition, identification of cellulitis or lymphangitis is paramount to halt processes that may lead to worsening lymphedema. The lymphatic evaluation includes palpation of all lymphatic territories. Examination should include limb girth measurements, usually performed via tape measure or volumetric displacement of the affected and contralateral extremities (129). Suspicion for recurrent malignant disease, while often unlikely to occur, should be part of care.

The management of lymphedema is performed to reduce symptoms, preserve cosmesis, maintain function, and decrease risk of recurrent infection. Complex decongestive therapy (CDT) is the most effective treatment (130). One component of CDT consists of manual lymphatic drainage, which is a massage technique promoting proximal lymph decongestion by fluid mobilization toward unaffected lymphatic territories. Skin care to reduce risk of infection is important and should be combined with stretching and soft-tissue mobilization techniques to the proximal limb, helping to reduce stasis of the lymphatic system. Wrapping the extremity with short stretch bandages to promote lymphatic flow out of the limb is the next phase and can be augmented by use of high tensile foam to break up fibrosis. The final phase uses exercises with gentle compressive wrapping in place in order to use the physiological muscle pump to propel fluid to the proximal lymphatics. CDT is typically divided into a decongestive phase that consists of 24 hours per day compression usually under the guidance of a trained physical or occupational therapist, and

<table>
<thead>
<tr>
<th>Factor</th>
<th>Example</th>
<th>Prevention</th>
</tr>
</thead>
<tbody>
<tr>
<td>Infection</td>
<td>Cellulitis, lymphangitis</td>
<td>Prevent breaks in skin</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Avoid venipuncture</td>
</tr>
<tr>
<td>Lymphatic constriction</td>
<td>Tourniquets</td>
<td>Avoid blood pressure cuffs</td>
</tr>
<tr>
<td></td>
<td>Tight clothes</td>
<td>Blood draws in opposite arm</td>
</tr>
<tr>
<td>Increased metabolism</td>
<td>Burns/extremes of heat</td>
<td>Scar tissue mobilization</td>
</tr>
<tr>
<td>Anaerobic metabolism</td>
<td>Excessive exercise</td>
<td>Sun screen</td>
</tr>
<tr>
<td>Trauma</td>
<td>Fractures, surgery</td>
<td>Avoid fatigue and soreness with exercise;</td>
</tr>
<tr>
<td>Air travel</td>
<td>Low ambient air pressure</td>
<td>build exercise routine gradually</td>
</tr>
</tbody>
</table>

The clinical evaluation of lymphedema should include a history of the primary malignancy and the aspects of treatment that have impacted the lymphatic system. The onset, duration, and progression of lymphatic swelling should be identified. In addition, identifying attempted management methods is important. Functional restrictions related to lymphedema (loss of shoulder or hand function) should be identified. Full musculoskeletal and neurological examination is needed to identify any underlying deficits. The skin examination should include general skin and soft tissue characteristics, and a specific focus on surgical and irradiated sites. In addition, identification of cellulitis or lymphangitis is paramount to halt processes that may lead to worsening lymphedema. The lymphatic evaluation includes palpation of all lymphatic territories. Examination should include limb girth measurements, usually performed via tape measure or volumetric displacement of the affected and contralateral extremities (129). Suspicion for recurrent malignant disease, while often unlikely to occur, should be part of care.

The management of lymphedema is performed to reduce symptoms, preserve cosmesis, maintain function, and decrease risk of recurrent infection. Complex decongestive therapy (CDT) is the most effective treatment (130). One component of CDT consists of manual lymphatic drainage, which is a massage technique promoting proximal lymph decongestion by fluid mobilization toward unaffected lymphatic territories. Skin care to reduce risk of infection is important and should be combined with stretching and soft-tissue mobilization techniques to the proximal limb, helping to reduce stasis of the lymphatic system. Wrapping the extremity with short stretch bandages to promote lymphatic flow out of the limb is the next phase and can be augmented by use of high tensile foam to break up fibrosis. The final phase uses exercises with gentle compressive wrapping in place in order to use the physiological muscle pump to propel fluid to the proximal lymphatics. CDT is typically divided into a decongestive phase that consists of 24 hours per day compression usually under the guidance of a trained physical or occupational therapist, and
a maintenance phase consisting of wrapping at night (performed by the patient or caregiver) and compressive garments during the day. In the maintenance phase, regular surveillance by a physician should be performed to assure volume reduction is maintained. Pneumatic pumps are also available but are controversial. They may be helpful in situations where the above treatments are not clinically feasible, or response is poor. Pneumatic pump use requires a time-consuming daily routine, and clinical response may be modest (129,130).

Evidence confirming the importance of early detection is emerging. A recent case series employing pre- and postoperative volumetric measurements via optoelectric perometry demonstrated subclinical lymphedema in 49% of ALND patients. Early treatment with provision of a 30/20 mm Hg compression sleeve and gauntlet was effective in reducing limb volume to near baseline levels (131).

Malignant lymphedema, resulting from spread of cancer into remaining lymphatics, presents with rapid swelling and a mottled appearance of the skin in the affected limb. Clinical examination may also reveal lymph node fullness. Evaluation should include exclusion of thrombembolic disease and imaging of the regional lymphatics with computed tomography (CT) or MRI. Identifying locoregional recurrence is imperative to initiate early treatment to reduce functional loss and maximize life expectancy. The primary treatment is management of the occluding tumor.

**IMPACT OF CANCER AND CANCER TREATMENTS ON NUTRITION**

Nutrition is an essential variable for all rehabilitation outcomes but specific concerns apply to cancer patients at different points during the disease continuum. Many cancers as well as oncologic treatments affect nutritional status resulting in delayed wound healing, longer hospital stays, diminished quality of life, lower survival rates, and reduced functional performance (132,133).

Guo et al. recently demonstrated a 50% prevalence of below normal prealbumin (<18 mg/dL) in cancer patients admitted to an inpatient rehabilitation unit (134). Older studies found that close to two thirds of cancer patients developed inanition during the course of their disease, demonstrating severe risk in the absence of systematic intervention (135). Multiple anthropometric and laboratory measures can be used to elegantly characterize malnutrition in individual cancer patients but serum albumin below 3 g/dL and loss of 10% or more of stable pre-illness body weight continue to be the most reliable indicators of malnutrition in clinical practice. Laboratory and radiographic studies are necessary for identifying and responding to reversible components of malnutrition in cancer patients. Folate or B12 deficiencies may precipitate symptomatic anemia while hypercalcemia may induce anorexia and nausea. The Bristol-Myers Anorexia/Cachexia Instrument and the Functional Assessment of Anorexia/Cachexia Therapy provide validated means for assessing the impact of malnutrition on patients and can be used sequentially to track the effects of prescribed interventions (136).

Weight loss and malnutrition in cancer patients occur as the result of primary and secondary processes. Secondary processes that interfere with intake and absorption are common and more easily reversible. Nausea and anorexia may be associated with specific chemotherapy regimens as well as with malignancies involving the gastrointestinal system. Depression may be a cause of reduced intake. Patients with head and neck cancers and other diagnoses frequently present with or develop significant dysphagia. Others develop mucositis as the result of treatment regimens. Radiation enteritis can cause malabsorption, intestinal obstruction, and even fistulas that may require total bowel rest and parenteral nutrition.

Secondary causes of malnutrition in cancer patients should be identified proactively and tailored treatment plans should be introduced as early as possible. Mucositis can be treated with ice chips and oral analgesics although these may decrease enjoyment of food. Avoiding anticholinergic medication and adding oral rinses or pilocarpine to promote salivary flow can limit the effect of dry mouth on oral intake. Oral supplements may be more easily tolerated and appreciated than solid foods. Modifications can be made to minimize the effect of acquired aversions such as the use of plastic utensils for patients complaining of a metallic taste or the elimination of odors that affect smell-sensitive patients. Antiemetics such as ondansetron, prochlorperazine, and trimethobenzamide can be used to mitigate nausea while prokinetic agents such as metoclopramide can reduce early satiety from gastric stasis. Appetite stimulants including megestrol acetate and drobinal are widely used and have been shown to increase intake and promote weight gain (137). It remains unclear, however, whether these gains reflect the actual increase in fat-free mass (FFM) needed for improved function. Concern over the prothrombotic effect of megestrol acetate has lead to successful trials with lower dosages (138). Anabolic steroids such as oxandrolone may be used to promote lean body mass in patients with androgen receptor negative malignancies (139).

Extreme weight loss, specifically the loss of lean body mass, may be a primary disease-related process. Cancer cachexia was previously thought to result from direct tumor effect but more recent investigations have suggested that cytokine-mediated processes alter the overall metabolic balance in cancer patients and animal models. TNF, proteolysis inducing factor (PIF), and interleukin 6(IL-6) dependent mechanisms shift metabolic processes increasing lipolysis, protein consumption, and energy expenditure (140).

Timely nutritional assessment and intervention are essential to effective cancer rehabilitation. Identifying nutritionists with oncologic experience can be invaluable for cancer rehabilitation programs. The American Dietetic Association has published standards of practice and professional performance for specialty accreditation in the area of nutritional management of oncology patients (141). In addition to early assessment and monitoring the impact of nutritional interventions, nutritionists in many cancer programs serve as a de facto clearing house for patients and their caregivers.
CANCER-RELATED SEXUAL DYSFUNCTION

Sexual dysfunction may occur as a primary effect of malignancy but more commonly occurs as a side effect of treatment. Sexual function can be affected by disturbance at many levels, including psychologic, central or peripheral nervous system, endocrine, pelvic, vascular, and local effects on gonadal structures. Physical changes may interfere with the patient’s concept of his or her sexual attractiveness. Depression may result in low sexual drive. Sexual dysfunction in cancer patients may persist long after treatment is completed, and it is important for the practitioner to inquire about it (143).

In men, chemotherapy can have adverse effects on spermatogenesis and also on testosterone production. Neuropathies, including dysautonomia, can interfere with the emission phase of male orgasm. Endocrine effects occur in men with prostate cancer treated with orchectomy or with hormonal regimens to reduce serum testosterone. In women, gonadal toxicity can also occur with chemotherapy, especially alkylating agents, and permanent menopause may occur with combination chemotherapy.

Radiotherapy to the prostate or testicles can produce erectile dysfunction, possibly as a result of acceleration of pre-existing atherosclerosis from postradiation fibrotic changes. In women, radiotherapy to the pelvis produces premature menopause, with a dose of as low as 600 to 1,000 rad permanently destroying ovarian function. Radiation also produces damage to the vaginal epithelium, and the fibrotic process can continue over years. The result can be dyspareunia, postcoital bleeding, and even vaginal ulcers. After local radiation therapy for cervical cancer, stenosis of the upper vagina can occur.

After prostate surgery, reports of erectile dysfunction have varied widely, from 9% to 86%, probably related both to the extent of surgery (and other therapies) and assessment techniques (144). More rarely, urinary incontinence occurs. Similar problems may occur after radical cystectomy and abdominoperineal resection of gastrointestinal tumors.

First-line therapies for erectile dysfunction include education and counseling, oral medications (such as sildenafil, vardenafil, or tadalafil), and the use of a vacuum device. Second-line therapies include use of intrarectal agents such as alprostadil, and intracavernous injection therapy (as with papaverine), and the third-line option is surgery for penile prosthesis. Following retropubic prostatectomy, higher doses of sildenafil may increase smooth-muscle content, with potential benefit of maintaining the proerectile ultrastructure during the recovery process (144).

Among breast cancer survivors, chemotherapy has been associated with greater reported sexual problems (145). Hormonal therapies, especially tamoxifen, may also produce side effects such as menstrual changes, vasomotor symptoms (hot flashes, headaches), and vaginal dryness (143).

Among women presenting to a sexual medicine clinic at a large cancer hospital, the most common problems were dyspareunia (72%) and vulvovaginal atrophy (65%). The most common interventions included vaginal moisturizers and water-based lubricants (89%), psychosexual counseling (46%), and local, minimally absorbed vaginal tablets (35%), with 70% of those attending follow-up sessions reporting improvement (146). In general, greasy lubricants, such as petroleum jelly, should be avoided because they may block the urethral opening. Counseling is usually short term and optimally includes both partners. Disease or treatment-specific effects on sexual function should be discussed with the patient, including the use of drawings or diagrams when appropriate. For women who have undergone resection of pelvic tumors, vaginal dilators or pelvic floor exercises may also be indicated. Preservation of fertility is an increasing focus in many programs.

Patients should be encouraged to maintain a good personal image as well as physical closeness and intimacy. The American Cancer Society’s Look Good, Feel Better and Reach to Recovery programs can be very helpful. Peer support through ostomy groups can be of benefit.

RETURN TO WORK AFTER CANCER

Return to work after cancer is usually favorable. However, a large study exploring employment status of cancer survivors who had been working at the time of diagnosis, found 20% reporting disabilities, and about 13% had stopped working for cancer-related reasons, mostly after the first year. Diagnoses most likely to affect long-term employment status include CNS tumors, head and neck cancers, and advanced hematologic malignancies (147). Data from the Childhood Cancer Survivor Study, a very large registry with sibling controls, found 19.6% incidence of performance limitations overall, with 7.9% unable to attend work or school for health reasons (risk ratio 5.9), and most pronounced effect after brain and bone cancers (148). Cancer survivors may be hampered by symptoms and also experience insurability concerns (especially affordability). One study of 253 long-term cancer survivors, many of whom were approaching retirement age, found 67% were actively working 5 to 7 years later, but some individuals reported that the cancer interfered with performing physical
tasks (18%), lifting heavy loads (26%), stooping, kneeling or crouching (14%), prolonged mental concentration (12%), analyzing data (11%), keeping pace with others (22%), and learning new things (14%) (149). Individuals may require extended or intermittent leave from work, and flexible work schedules. Organizations including the National Cancer Institute, American Cancer Society, Lance Armstrong Foundation, and the National Coalition for Cancer Survivorship have educational materials and programs related to employment. For those with legal questions related to employment or health insurance, the American Cancer Society’s information line (1-800-ACS-2345) refers callers to the Cancer Legal Resource Center of the Western Law Center for Disability Rights. Other resources, while not specifically cancer-related, can include the Equal Employment Opportunity Commission (http://www.eeoc.gov/) and the Job Accommodation Network (http://www.jan.wvu.edu/portals/individuals.htm) (150).

### SPECIFIC TUMORS AND THEIR REHABILITATION NEEDS

#### Breast Cancer

Breast cancer is the most common malignancy in women, with over 200,000 new diagnoses per year in the United States (124). The standard for treatment of breast cancer is surgical resection of the tumor with axillary lymph node assessment. The surgical procedures may be followed by any combination of chemotherapy, radiation therapy, hormonal therapy, and plastic surgical reconstruction. With approximately 2 million survivors in the United States, the rehabilitation needs of the breast cancer population have been the most studied and most documented. The most common problems include postsurgical pain, regional shoulder dysfunction, axillary pain syndromes, lymphedema (see “Lymphedema” section), and psychologic adjustment needs. In addition, for those with systemic disease, palliative care includes management of pain, bone metastases, and neurological impairments.

Postsurgical pain can occur with any breast procedure but is frequently mislabeled as postmastectomy pain syndrome. The various procedures for tumor resection include lumpectomy, partial mastectomy, quadrantectomy, simple mastectomy, modified radical mastectomy, and radical mastectomy. Axillary assessment is performed both for staging and treatment. Axillary assessment may include sentinel lymph node dissection (sampling a few lymph nodes that immediately drain the tumor) or full axillary dissection (removing all axillary nodes below the axillary vein) (125,126,151). Overall, approximately 40% of patients develop pain after the above surgical procedure that extends beyond the expected healing time (2 months). Risk factors include young age, more extensive surgery, greater immediate postoperative pain, and anxiety (126,152–154). Incidence of lymphedema is 3% to 5% for those requiring only sentinel lymph node sampling compared to 13% to 19% of patients requiring full ALND (125,151).

Postsurgical pain syndromes include phantom breast pain (155), incisional alldynia, neuroma formation, pectoralis muscle pain, and intercostobrachial neuropathy (156). Incisional alldynia is related to the dense cutaneous innervation of the breast, resulting in hypersensitivity approximating the incision. Neuroma formation typically occurs at the end of the incision and results in a focal area of pain and sometimes a small palpable mass. The pectoralis muscle may be impacted mechanically from surgery and presents with diffuse chest wall pain, worse with shoulder flexion and external rotation (157). The intercostobrachial nerve is the lateral cutaneous nerve that derives from the dorsal primary rami of the second thoracic root. It traverses across the chest wall, enters the axilla and provides sensation to the lateral chest wall and posteromedial portion of the arm. It is frequently sacrificed during axillary dissection and results in sensory impairment and occasionally alldynia.

The assessment of patients with postsurgical pain requires a history that focuses on the type of pain, functional loss, and sleep disruption from pain. The clinical exam should focus on ruling out other causes of pain, including radiculopathy, rib fracture, intercostal muscle strain, cellulitis, and systemic infection. The primary management consists of self-applied techniques including cutaneous desensitization, soft-tissue mobilization, stretching of regional muscle groups and shoulder range of motion. These techniques usually can be commenced once the surgical incision is healed. For those with persistent pain, severe loss of function, or sleep loss, additional care, including a more extended course of physical or occupational therapy, pharmacoeutical management, even interventional procedures may be warranted (157). Thermal modalities should be used with caution secondary to concern of provoking lymphedema.

Breast cancer-related shoulder dysfunction is very common, affecting greater than 50% of patients and is the most common cause of long-term morbidity (157). Shoulder dysfunction typically manifests itself as rotator cuff dysfunction, frozen shoulder, or myofascial pain. The causes of the above issues are multifactorial and can be related to immobilization, muscle weakness, atrophy, mechanical alterations in the shoulder girdle musculature, radiation therapy and neurological impairment. Initially, after surgery there is a period of relative immobility to promote wound healing (see next paragraph). In addition, local irritation to the pectoralis muscles results in shortening and contraction of the muscle groups. This alteration results in aberrant length tension relationships in the shoulder girdle musculature.

The assessment of breast cancer-related shoulder pain begins with a history including postoperative complications, duration of immobilization, pain complaints, functional restrictions, weakness, and sensory symptoms. The history should also identify red flags for recurrence such as constitutional symptoms, progressive pain complaints, concurrent plexopathy or unusual presentations of lymphedema. The physical examination should identify atrophy, restriction in regional musculature; assess scapulohumeral and glenohumeral mobility, regional muscle strength and include a comprehensive neurological exam. The management of shoulder impairment...
Head and Neck Cancer

Head and neck cancer is the sixth most common cancer diagnosis in the United States, with approximately 40,000 cases per year (120,162). Squamous cell carcinoma is responsible for 95% of cases, and there is a strong association with alcohol and tobacco use (163). The primary treatment modalities for head and neck cancer include surgery, radiation therapy, and chemotherapy. Radiation and chemotherapy were used in most cases of local disease. Surgery is used for locally invasive disease and lymph node positive disease along with radiation. The morbidity of head and neck cancer is very high. Common issues include disfigurement, difficulty swallowing, impaired communication, regional shoulder girdle dysfunction, pain, soft-tissue fibrosis, and lymphedema (Table 44.9).

Swallowing dysfunction results in progressive loss of oral intake, weight loss, fatigue and ultimately decreased survival, and can affect the oral, pharyngeal and esophageal phases. Surgical removal of the tongue, mandible, or teeth results in decreased oral food processing and propagation to the pharynx. Compensatory strategies, such as forward or backward head tilts and strengthening of buccal musculature, may facilitate deglutition. Oral transit can be achieved with the use of a syringe or pusher (placement feeding) spoon. Radiation therapy can result in mucositis, xerostomia, trismus or osteoradionecrosis (120,162). Mucositis is irritation of the mucosal barrier and results in pain; this can be a result of either chemotherapy or RT. The primary management is topical agents for pain control, oral hygiene, and artificial lubrication. For severe oral pain, opioids may be utilized to maintain oral intake. After surgery, pendulum exercises and range of motion exercises below 90 degrees of shoulder abduction and flexion are advocated until postsurgical drains are removed (158). Once drains are removed, aggressive pain management and daily shoulder stretching, targeting the pectoralis complex and latissimus dorsi, are initiated. Once full range is obtained, strengthening of the scapular stabilizers, rotator cuff, and deltoid musculature should begin. For those with persistent or progressive shoulder dysfunction, advanced imaging with MRI and referral to a skilled therapist who understands cancer principles is warranted.

Intercostobrachial neuropathy (see above) and axillary web syndrome (axillary coring) are common causes of axillary pain after breast cancer treatment. In addition, seroma formation in the axilla can frequently complicate postoperative recovery. Axillary web syndrome is a hand-like sensation that begins in the mid-axilla and can extend into the antecubital fossa and lateral wrist (157,159). It is due to sclerosed lymphatics and blood vessels that have lost their continuity secondary to axillary dissection (159). Patients complain of pain and shoulder restriction that is aggravated with shoulder flexion and forearm supination. This is self-limited, and usually results in a return of normal function and pain resolution with minimal intervention. The onset is usually within 2 months after axillary dissection and may be aggravated by radiation therapy. The main clinical consequence is loss of shoulder and/or elbow range of motion. Management includes massage and pressure to break up scar tissue along with shoulder mobility (160). A seroma consists of fluid that accumulates in the cavity after mastectomy and axillary dissection. Drains are placed postoperatively and typically remain until drainage is less than 30 cc over a 24-hour period, on an average 2 weeks. Persistent seroma formation leads to increased risk for infection and prolongs the restriction of activity in the upper extremity. Close surveillance and judicious use of the upper limb are standards for management.

Locoregional spread of breast cancer typically manifests itself with chest wall lesions, brachial plexopathy, or lymphedema. Change in skin contour, color, or an atypical rash in breast cancer survivors requires prompt attention and biopsy. Brachial plexopathy results from malignant seeding of the brachial plexus by tumor cells and affects 4% to 5% of breast cancer patients (161). The lower trunk is affected first resulting in shoulder, medial arm, medial forearm and fourth and fifth digit pain and sensory loss. The symptoms can progress and result in a pan-plexopathy. MRI of the brachial plexus with gadolinium is the imaging test of choice in suspected malignant spread. Radiation plexopathy may be in the differential diagnosis; presentation includes greater likelihood of paresthesias, upper trunk involvement, and myokymic discharges, and lower incidence of pain.

### TABLE 44.9  Head and Neck Cancer

<table>
<thead>
<tr>
<th>Intervention*</th>
<th>Impairment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glossectomy</td>
<td>Communication; dysphagia</td>
</tr>
<tr>
<td>Mandibulectomy</td>
<td>Chewing</td>
</tr>
<tr>
<td>Radical neck dissection</td>
<td>Shoulder depression and protraction, facial lymphedema; neck asymmetry</td>
</tr>
<tr>
<td>Radiation therapy</td>
<td>Contracture (neck, jaw); osteonecrosis; dental changes; decreased salivary flow; dysphagia (usually mild); trismus; CNS complications</td>
</tr>
<tr>
<td>Partial laryngectomy</td>
<td>Dysphagia; communication</td>
</tr>
<tr>
<td>Total laryngectomy</td>
<td>Communication (severe); dysphagia (mild)</td>
</tr>
<tr>
<td>Soft-palate resection</td>
<td>Communication; dysphagia</td>
</tr>
<tr>
<td>Nasopharyngeal tumors/treatment</td>
<td>Cranial nerve abnormalities</td>
</tr>
<tr>
<td>Salivary gland tumors/resection</td>
<td>Facial nerve palsy</td>
</tr>
</tbody>
</table>

*See text for treatment strategies.
breads. In addition, xerostomia can compromise oral hygiene and increase risk for dental caries and oral fungal infections. Radiation oncologists attempt to diminish xerostomia by parotid sparing techniques (the parotid glands produce 60% of salivary secretions), localizing RT to tumor sites and using preventative medications such as amifostine. Treatment of xerostomia consists of strict oral hygiene, including sodium fluoride gel, rinses with saline, salt and baking soda (one teaspoon of each in a quart of water) or diluted peroxide, artificial lubricants or use of a spray mister, and salivary gland stimulating agents such as pilocarpine or sugarless gum or lozenges. Diet should emphasize fluids, moist foods, and high-calorie supplements. For edentulous patients, one should wait 6 months after radiation therapy before replacing dentures secondary to gingival shrinking and remodeling. RT also may result in trismus, secondary to fibrosis of the jaw musculature. This is usually a delayed effect and can result in severe jaw mobility restriction. Prevention is through daily jaw mobilization exercises during and after RT. Treatment is through the use of therabite (Therabite Corp, Newton Square, PA), which is a utensil that provides progressive oral excursion through addition of plates. Osteoradionecrosis is a rare complication of RT that results in destruction of the jaw. This can produce severe pain and dental loss. Management is supportive. After unilateral mandibular resections, strengthening of masticatory muscles is needed to prevent drift to the nonsurgical side.

The pharyngeal phase of swallowing is impacted primarily by surgical disruption of the pharyngeal structures and by radiation fibrosis (164). Surgical disruption may be permanent and require long-term enteral feeding. Total laryngectomy does not lead to aspiration risk, as the airway and food pathway are anatomically separated by surgery; however, partial laryngectomy or sacrifice of the superior laryngeal nerve can produce aspiration risk. Many patients experience postoperative dysphagia that is transient, with as many as 85% of supraglottic laryngectomy patients attaining a functional swallow (improved from only 39% in the early postoperative period) (165). Radiation to the superior and middle constrictors results in abnormal elevation of the larynx and pharynx and resultant loss of airway protection (166). Assessment of laryngeal dysfunction is via modified barium swallow to identify aspiration. The primary management of pharyngeal dysphagia is airway protection techniques, such as breath holding during the swallow and throat clear afterward, chin tuck, and dietary modification. The esophageal phase of swallowing is primarily impacted by radiation fibrosis resulting in esophageal stricture. The primary management is mechanical dilatation of the esophagus.

Communication difficulties in head and neck cancer patients can include dysarthria, dysphonia, and aphonia. Basic strategies include use of good eye contact and appropriate gestures, as well as oral exercises for articulation and modification of pitch, loudness, and voice quality. Dysarthria is typically a result of tongue dysfunction from surgical resection or radiation fibrosis. Primary management includes early tongue mobilization and prosthetic devices. Dysphonia relates to vocal cord dysfunction. This can result from RT effects or from mechanical effects of tumor, surgical resection, or neurological impairment impacting the vocal cords. Management consists of retraining primarily through a speech therapist. Aphonia is one of the most feared complications in cancer treatment and is usually a result of total laryngectomy (TL) (163). Communication after TL is through esophageal speech (propelling air into the esophagus and expelling it, making sound), electrolarynx (hand held device to the larynx that makes sound), or tracheoesophageal puncture (TEP; includes placement of a valve that allows for speech production; this approach is also known as tracheoesophageal prosthesis or fistula) (167). Esophageal speech is difficult to master with only 26% becoming facile (167). Electrolarynx results in a mechanical computerized sound that is unnatural to patients. TEP is the preferred method for communication after TL (Fig. 44-3). It results in phonation that is harsher but closer to normal than the other techniques (167). Criteria for TEP placement include no oncologic compromise, normal swallow function, reliable voice, and anatomical integrity for placement (167). Frequently, patients will require dedicated training from a speech therapist.

Salivary gland tumors often result in facial nerve impairment, either from the tumor itself or after surgery. Facial nerve palsy often recovers over a few months; nerve grafting techniques have also been employed (168).

Radical neck dissection (RND) is the most common means of surgically staging and treating locally advanced head
and neck cancer. The traditional RND removes all lymph nodes from the mandible to the clavicle, the sternocleidomastoid, accessory nerve, and external jugular vein. Recently, accessory nerve sparing procedures called functional neck dissection (FND) have become common (169). The effects of the RND include accessory nerve dysfunction, greater auricular neuropathic pain, lymphedema, and neck restriction.Accessory nerve dysfunction directly affects the function of the trapezius muscle and results in regional shoulder dysfunction. The rhomboids and levator scapula become overstretched and the pectoralis major shortened. A recent study comparing RND to FND identified less functional loss in the FND population but persistent strength and functional loss even in those receiving accessory nerve sparing procedures, suggesting stretch, compression and hypovascularity may contribute to nerve dysfunction (169). The result of trapezius loss is abnormal medial scapular stability and excessive loading of other scapular stabilizers with myofascial pain, restriction in shoulder mobility and weakness with above head activities (170). Clinically, patients present with a dropped ipsilateral shoulder, inability to shrug and forward flex the arm, restriction with above head activity and myofascial pain most notably in the levator scapulae. Management includes recumbent range of motion exercises to stabilize the scapula, followed by compensatory shoulder girdle training and education to prevent shoulder injuries. Prognosis is variable and improvement may occur over an 18- to 24-month period.

Pain is common after head and neck cancer treatments, affecting 25% to 50% of patients (162). The most common etiologies for pain include oral pain, soft-tissue fibrosis, and neuropathic pain. Oral pain can be a result of surgery, mucositis, dental caries, and osteoradionecrosis all covered in prior sections. Soft-tissue fibrosis is considered a late effect of radiation (171). Fibrosis is mediated by microvascular disruption and hyperactivation of TGF-β, resulting in scar tissue formation (119,164). The primary effect is painful, restricted range of motion of the neck and jaw. Neck extension, rotation, and lateral bending can be severely impacted. Management is via primary prevention with daily stretching and soft-tissue mobilization techniques. Pentoxifylline and vitamin E have been used in concert to decrease fibrosis and maintain microvascular supply with variable results (119). Neuropathic pain is commonly seen from RND disruption of the greater auricular nerve. Patients can present with exquisite pain that extends from the incision site to the posterior portion of the ear. Common techniques for management include manual desensitization, topical anesthetics, and neuropathic pain medication.

Lymphedema may result from lymph node removal secondary to RND in combination with radiation changes to the lymphatics. Lymphedema of the face with the above techniques (172). In cases of severe acute swelling, thrombosis of the superior vena cava or jugular veins should be excluded.

**Hematologic Malignancies**

Hematologic malignancies often produce fatigue (173), and additional disabling effects can also occur. The most common pain syndromes include neuropathic pain from the primary malignancy, chemotherapy, or herpes zoster, bone pain related to bone marrow expansion or to osteolysis, and headache related to CNS involvement (see “Pain,” “Bony Metastatic Disease,” and “Peripheral Neuropathy” sections). Peripheral polyneuropathy occurs in 10% of patients with multiple myeloma and 15% of those with amyloidosis (174).

Lymphomas include Hodgkin’s disease and non-Hodgkin’s lymphomas. Non-Hodgkin’s lymphomas, including CNS lymphoma, are seen with increased frequency in patients with human immunodeficiency virus (HIV) and in individuals who are immunosuppressed after transplant surgery (175). Bone marrow disorders include acute and chronic leukemias and myeloproliferative disorders, such as polycythemia vera, myelofibrosis, and primary thrombocytopenia. Immunoproliferative diseases include multiple myeloma and Waldenström’s macroglobulinemia.

Lymphomas characteristically invade lymphatics but can readily invade any tissue, with non-Hodgkin’s lymphoma having a greater propensity to invade extranodal sites than Hodgkin’s disease. In addition to the bony and peripheral nerve disorders noted above, patients with history of thoracic (mantle) radiation therapy may have complications such as radiation pneumonitis, cardiac abnormalities (constrictive pericarditis, cardiomyopathy), and hypothyroidism (121). Aseptic necrosis of the femoral heads is described in 10% of long-term survivors of inverted y-field (abdominopelvic) irradiation and MOPP chemotherapy (121). Mycosis fungoides is a cutaneous T-cell lymphoma involving helper T-lymphocytes. Whirlpool treatment may be beneficial for the intractable pruritus and widespread skin scaling that can occur with this lymphoma.

Immunoproliferative diseases occur in older individuals (average age 60 at diagnosis). The skull, vertebrae, ribs, pelvis, and proximal long bones are the most common sites of involvement (121). Osteoblastic lesions are rare (incidence of <2%) and often occur with neuropathy (121). Painful pathologic fractures are common. Spinal cord or nerve root compression may occur from epidural plasmacytoma, amyloidosis, or vertebral body collapse. Cranial nerve palsies occur from tumor occluding calvarial foramina. Focal peripheral neuropathies, such as carpal tunnel syndrome, can occur from soft-tissue amyloid infiltration (121).

The leukemias can produce fatigue and bone pain. Neurologic changes may occur when white blood count is more than 100,000, due to leukostasis, and in the presence of leukemic meningitis. In children, the skeleton is often the first body system to manifest abnormality, and acute leukemia must be included in the differential diagnosis of unexplained musculoskeletal pain or bony pathology.
Rehabilitation emphasizes active exercise, including walking, to maintain endurance (stamina) during the treatment course, which is often protracted. Precautions surround the risk of bleeding in thrombocytopenic patients; pressure palsies, especially in patients receiving neurotoxic drugs or with severe immobilization; and orthostasis. Restricted bony impact is advised for extensive bony infiltration or aseptic necrosis.

Bone marrow transplant recipients have additional needs. High-dose steroids are needed to counteract GVHD, with potential complications including steroid myopathy, aseptic necrosis, and osteoporosis. The exercise program should include proximal muscle strengthening, back extension, and upper-extremity weight bearing (see “Exercise for the Cancer Patient: General Aspects”).

Lung Cancer
Lung malignancies remain very common and are challenging for the physiatrist. Cure is possible in some cases, especially when the disease is detected early and in an operable location, but generally the prognosis remains poor. This often poses a dilemma for the physiatrist because lung cancer produces fatigue and deconditioning that occur with virtually any chronic illness (plus the fatigue of pulmonary disease), as well as many potential neurologic complications, including peripheral polyneuropathy, brain and spine metastasis, or myasthenic syndrome (see “Neurologic Complications of Cancer” section). Apical lung tumors, such as Pancoast tumor, can invade the brachial plexus or lower cervical nerve roots by direct extension, producing pain as an early symptom that is often referred to the shoulder; appropriate diagnosis may be delayed while the patient is treated for a nonexistent musculoskeletal shoulder problem. The neurologic complications often occur early in the clinical course and may even be the presenting feature. Despite a historically pessimistic outlook, in one recent study, lung cancer patients on acute rehabilitation were among the least likely to require transfer back to acute care (28). Improved exercise capacity has been reported in lung cancer patients with severely impaired pulmonary function with an 8-week intensive rehabilitation program (176).

Strategies for pain management as well as education about pulmonary hygiene and breathing techniques may be useful. Patients treated with surgical resection (either lobectomy or pneumonectomy) should be instructed in techniques to promote maximal chest expansion, including coughing, pursed-lip breathing, diaphragmatic breathing, and segmental (hand over upper or lower portion of the rib cage during inhalation) breathing exercises. Postoperatively, cough may be best tolerated with the knees flexed, and the incision splinted by firm pressure from a pillow or towel roll. Trunk posture and mobility exercise, lower-extremity exercise, and early progressive ambulation are recommended.

Gastrointestinal Malignancies
Colorectal cancer is the fourth most prevalent carcinoma in the United States, the second leading cause of cancer death, and the only major malignancy that affects men and women almost equally. Surgery is considered the only curative treatment, at times combined with adjuvant chemotherapy or radiation. When a stoma is present, the enterostomal nurse is a focal person in the rehabilitation process. Acceptance of the stoma is important for long-term quality of life (177). Literature on quality of life among patients with or without stoma has shown inconclusive results (177), and little data is available on long-term function in this population.

Brain Tumors
Brain Tumor Types: Characteristics and Prognostic Implications
Intracranial neoplasms vary widely with respect to rate of growth, impact on function, and overall prognosis. Even a benign or relatively low-grade brain tumor may have enormous functional consequences, depending on location and the extent to which decompression is feasible. Brainstem lesions, for instance have a profound impact on motor function, specific cranial nerves, swallowing, and coordination with little possibility of resection without risk of causing even greater deficits. Communication with referring surgeons, correlation of operative reports, imaging, treatment history, and clinical examination help predict physiologic and functional recovery and allow for the selection of achievable rehabilitation goals.

Benign Brain Tumors
In adults, meningiomas are the most common benign brain tumor, comprising about 20% of all primary brain tumors (178). Other lesions include pituitary adenomas, acoustic neuromas, craniohypophysealomas, epidermoid tumors, third ventricle colloid cysts, and the hemangioblastomas of Von Hippel-Lindau disease. Untreated pituitary adenomas cause endocrine abnormalities, progressive bitemporal visual field loss, and headaches. Prolactinomas precipitate acromegaly and corticotrophin-secreting adenomas are the leading cause of hyperadrenocorticism, so endurably described by Cushing in 1910 (179). Many are effectively treated with one-time surgical resection. Acoustic neuromas grow on the nerve sheath of the vestibular nerve and occur sporadically as well in patients with neurofibromatosis. Patients can develop hearing loss, vertigo, facial palsy, dysphagia, facial numbness, and even hydrocephalus. Results of microsurgery are generally good, especially with tumors measuring less than 2 cm, but local recurrence is common. Craniohypophysealomas arise from pharyngeal epithelium and cause growth failure in children and endocrine dysfunction with or without visual loss in adults as well as significant hydrocephalus. New techniques now permit endoscopic resections to avoid open craniotomy. Radiation has improved the long-term survival.

Malignant Brain Tumors
Ninety percent of malignant brain tumors belong to one of the three subdivisions of gliomas: (a) astrocytoma, (b) anaplastic astrocytoma, and (c) glioblastoma multiforme, each carrying a progressively poorer prognosis (178). While the incidence of primary cerebral lymphomas had increased in association with
HIV, in recent years incidence and prognosis are improving, probably due to more effective HIV regimens (175).

Low-grade astrocytomas of the cerebral hemispheres can be resected but radiotherapy is often used when tumor location precludes safe surgical excision. More aggressive surgical excisions are required for higher grade lesions unless they involve the pons, hypothalamus, or deep frontal white matter. Adjuvant radiotherapy, typically delivered to a limited field, prolongs survival. Chemotherapy appears to augment long-term survival, possibly limited to certain as yet unidentified subgroups. Current studies are assessing the synergistic efficacy of oral temozolomide, which is relatively well tolerated, given during focal radiation for glioblastoma multiforme. Younger age, higher performance rating at diagnosis, and favorable histology are associated with better outcomes (180).

**Brain Metastasis**

See “Neurologic Complications of Cancer,” “Metastatic Brain Disease,” and “Radiation-Induced Tissue Damage” sections.

**Rehabilitation Principles for Brain Tumor Patients**

Impaired cognition (80%), focal motor weakness (78%), and visual-perceptual impairment (53%) compromise function in brain tumor (181). Most patients have impairments in more than one domain and medical comorbidities that warrant a medically supervised rehabilitation program. Current neurologic status as well as expected clinical course determines the selection of realistic goals and rehabilitation interventions. Functional gains are possible for many patients with brain tumors but rehabilitation services for patients with steadily worsening neurological status should focus on caregiver training for palliative goals.

Rehabilitation outcomes for patients with brain tumors compare favorably to gains made by patients with other types of brain disorders such as traumatic brain injury (TBI) or stroke. Standard rehabilitation strategies for the other diagnoses (see Chapters 23 and 24) can be effectively used for brain tumor patients. Brain tumor patients participating in comprehensive inpatient rehabilitation programs have been shown to achieve meaningful gains in FIM scores with characteristically shorter lengths of stay (23,24). Concurrent radiotherapy treatment does not significantly alter these gains (26). Most studies show a higher than average rate of transfer back to an acute medical setting, but with the majority (69%) of these patients returning to meet their rehabilitation goals prior to discharge to a private home/community setting (24,26). Brain tumor patients show similar gains between those with primary brain tumor and metastatic disease (181,182); functional gains are better during the initial presentation to rehabilitation than with recurrence (183). One study examining brain tumor patients found better outcomes in meningioma patients than in those with other diagnoses and in those with left brain lesions (23).

Depression and anxiety must be treated but care should be taken when prescribing medications that may further affect cognition. Data collected from patients with gliomas have documented improvements in function with 10 mg twice daily methylphenidate and has led to the recommendation of this agent as an adjuvant in the treatment of brain tumors (183).

Because the radiation treatments may produce fatigue, sessions should be scheduled later in the day after rehabilitation therapies have been completed (24). Continuation of oral corticosteroids is needed during radiation treatment, as edema due to radiation can precipitate or intensify neurological and cognitive symptoms. Corticosteroid regimens ranging from 2 to 24 mg/d in divided doses of oral dexamethasone or equivalent agent are typically given. Treating clinicians should be aware of potential side effects including steroid psychosis, gastrointestinal bleeding, avascular necrosis, and proximal myopathy. Anticonvulsant therapy appears to be justified in patients with a history of seizure and possibly for short-term use perioperatively, but efficacy of long-term prophylaxis for patients without history of seizure has not been established (184). Incidence of deep vein thrombosis in brain tumor patients is high, and prophylaxis, such as intermittent pneumatic compression with or without low-dose heparin, is warranted. Incidence of postoperative deep vein thrombosis may be highest in meningioma patients (72%) and lower in brain metastasis (20%) (185). Anticoagulation must be judiciously restarted in collaboration with the treating neurosurgeon following any kind of intracranial tumor resection.

**Sarcomas of Bone and Soft Tissue**

Sarcomas are mesenchymal tumors that develop in bone and soft tissue. Although they constitute less than 0.1% of all cancer diagnoses, an incidence of over 12,000 new cases annually (186) and robust survival rates across a broad range of ages create a steady need for specialized acute as well as long-term rehabilitation services for this population. Mobility and self-care deficits are common as 45% of sarcomas involve the lower limb and 15% the upper limb (187). The remaining 40% involve the head, neck, or trunk and may result in functional impairment as well as cosmetic concerns.

Overall, soft-tissue sarcomas are three to four times more common than bone sarcomas (7). Twenty types of soft-tissue sarcoma have been identified but most types are treated similarly based on disease stage rather than tumor type. Tumor size and histologic grade are the most significant prognostic factors. Benign sarcomas including enchondromas, osteochondromas, desmoplastic fibromas, and hemangiomas produce various degrees of impairment as space occupying lesions that may require surgical intervention. Primary malignant bone tumors, as opposed to metastases or plasmacytomas (discussed elsewhere), can present in childhood, adulthood, or old age. Osteosarcoma, chondrosarcoma, Pediatric Ewing’s sarcoma, and malignant fibrous histiocytoma (MFH) are the most common primary malignancies of bone. Fortunately, survival rates for both bone and soft-tissue sarcomas have increased dramatically over the past 20 years.

The key to successful oncologic and functional outcome is good local control. Early referral to centers with dedicated sarcoma programs is highly recommended as specific biopsy techniques and staging protocols increases the number of effective treatment options. Preoperative tumor cytoreduction including radiation for soft-tissue sarcomas and chemotherapy for both bone and
soft-tissue sarcomas may permit less radical surgical treatment including limb sparing surgeries in lieu of amputation (188).

Amputation
The need to achieve wide tumor-free margins can require extensive amputations for sarcoma patients including hemipelvectomies as well as hip and shoulder disarticulations. Initial prosthetic ambulation exceeds conventional expectations for such proximal amputations because many sarcoma patients are young, with few comorbidities even after completing multi-drug chemotherapy regimens. These regimens, however, can cause transient or enduring neurologic injury to the contralateral limb in patients of all ages. As with diabetic and dysvascular lower extremity amputees, the vascular, neurological, and musculoskeletal status of the nonoperative limb has an enormous effect on balance and prosthetic ambulation. Normal aging as well as specific survivorship issues will continue to produce new rehabilitation challenges.

Many elements of upper and lower extremity prosthetic management described in Chapter 74 can be applied to patients who undergo amputations as part of their cancer treatment. Key distinctions, however, have to be taken into account (189). Patients should know in advance that recent chemotherapy or radiation often delays wound healing and should be expected to prolong the interval between surgery and prosthetic training. Expected volume fluctuations of residual limbs may be exacerbated by coexisting lymphedema after surgical resection or radiation. Irradiated skin is also more prone to “breakdown” and is less able to heal, requiring meticulous skin inspection and multiple types of socket adaptation. Preliminary socket adaptation should include variable thickness socks, flexible liners, and pressure relief.

Limb-sparing Procedures
Advances in adjuvant treatments and surgical techniques over the past 30 years have made limb-sparing procedures a reasonable alternative for many patients (190). These patients must have a neurovascular bundle that is not compromised by tumor, have no pathologic fracture, have little extraosseous extension of tumor (for bone tumors patients), have no distant spread of tumor, and have a high probability for good cosmetic and functional outcome. Patients must be informed of the possibility of subsequent need for prosthetic revision or amputation due to infection, fracture, or prosthetic loosening (especially in athletic individuals). Close to a quarter of oncologic limb-sparing procedures will require subsequent surgical revision but less than 5% progress to amputation (191). For lower-limb procedures, energy expenditure data tend to favor the limb-sparing option over amputation (192). The cosmetic advantages of limb-sparing procedures are generally considered self-evident but comparison of quality-of-life variables between patients treated with limb-sparing procedure versus amputation has not shown a conclusive advantage to either approach (189). Quality-of-life for limb-sparing patients appears to vary with the success of the procedure; those with problematic results often report relief after the late amputation is performed.

Remobilization must be pursued slowly following limb-sparing procedures to help control edema, promote wound healing, and maintain alignment. Extensive soft-tissue resections often result in high output drainage for 1 to 2 weeks after surgery. In the absence of complications, isometric gluteal and quadriceps contractions, active assisted straight-leg raises, and ankle pumps may be started 1 to 2 weeks postoperatively. Upright activity progresses from non-weight bearing to partial and full weight bearing based on muscle strength and wound healing (193). Rotationplasty may be performed as an alternative to amputation for distal femur lesions. The ankle joint is rotated 180 degrees and replaces the absent knee joint, achieving the functional equivalent of a below-knee amputation. Because of poor cosmesis, however, this procedure is done less frequently than in the past (189).

Limb-sparing procedures may be limited to the removal of involved muscle, typically hip adductors, quadriceps, hamstrings, or gastrocnemius. Patients undergoing these procedures will need a period of bed rest until wound drainage stops. Leg elevation, compression garments, and active range of motion should be instituted as part of the rehabilitation process. Patients with extensive quadriceps excision will need a knee immobilizer postoperatively and may need to ambulate with an ankle-foot orthosis (AFO) in plantar flexion with a dorsiflexion stop so as to create and maintain an extension moment at the knee. Patients with adequate gluteal, hamstring, gastrocnemius, and soleus strength may be able to generate adequate knee extension despite the loss of their quadriceps mechanism (194). Adductor muscle group and hamstring excision rarely require assistive aides or orthotic device use. However, if the sciatic nerve is sacrificed, an AFO will be needed to provide toe clearance. When the gastrocnemius is excised, a rocker sole can be added to the sole of the shoe to promote push-off. Achilles tendon stretching should be performed regularly to maintain ankle range of motion (195).

Prostate Cancer
Prostate cancer is the most common malignancy in men, with greater than 200,000 new cases per year (196). Disabilities from bone metastases and complications of androgen deprivation are major factors in morbidity. Men who have undergone radical prostatectomy or radiation therapy may experience incontinence and/or impotence.

Bone metastasis occurs in 85% of patients who die from prostate cancer, and is the presenting finding in 10% to 20% of prostate cancer diagnoses (197). Prostate-specific antigen (PSA) can be utilized as a screening tool with less than 2% of patients having PSA less than 10 presenting with bone disease. Treatment for advanced disease includes androgen deprivation or radiation therapy. Median life expectancy is 53 months in those with isolated bone disease and good functional scores (198). Spinal cord compression is present in 6% to 12% of patients and requires urgent management to maintain function.

Complications from androgen deprivation therapy include loss of lean body mass, fatigue, and osteoporosis. The loss of lean body mass can be attenuated by resistance training (80). Osteoporosis occurs, with a 6% to 10% loss of bone mineral
density per year (199). Greater than 33% of patients develop an osteoporotic fracture within 5 years of initiating androgen deprivation therapy (200). Preventive management includes bisphosphonates (53), exercise, and bone protection techniques (see “Cancer-Related Fatigue,” “Neurologic Complications of Cancer,” “Cancer-Related Sexual Dysfunction,” and “Bony Metastatic Disease” sections).

CONCLUSION

Rehabilitation interventions for patients with cancer focus on preservation or restoration of function, pain relief, and education about planning and prioritizing life activities to assure quality. The rehabilitation team is in the unique position of being able to recommend practical suggestions to improve function and preserve independence throughout the course of illness whether in the settings of active disease, or the occurrence of late effects that may be experienced by cancer survivors.

REFERENCES


diets (generalists, specialty, and advanced) in oncology nutrition care. *J Am Diet Assoc.* 2006;106:946–951.


167. Cooper JS, Fu K, Marks J, et al. Late effects of radiation therapy in patients with cancer of the oropharynx are significantly affected by the radiation therapy dose to the superior and middle constrictor muscles: a dose effect relationship. *Radiat Oncol.* 2007;85:64–75.


The term vascular disease encompasses a variety of acute and chronic pathophysiologic syndromes caused by congenital and acquired disorders affecting the arterial, venous, and lymphatic systems. Arterial diseases include those acute or chronic disorders that result in partial or complete, functional or anatomic occlusion, or aneurysmal dilation of the arteries. An example of functional occlusion is abnormal vascular reactivity of the arteries supplying a given tissue such as vasospasm. Venous disease includes acute or chronic occlusion of the systemic venous or pulmonary arterial system, usually as a result of thromboembolism. Chronic venous disease is a spectrum of diseases and disorders of the limbs, with spider veins and varicosities on one end of the spectrum and edema, skin changes such as venous hyperpigmentation, and ulceration on the other. The cause is either primary valvular incompetence or previous deep vein thrombosis (DVT; postphlebitic/postthrombotic syndrome). Lymphatic diseases result from congenital or acquired disorders that cause obstruction, incompetence, or disruption of lymphatic vessels or the lymph-conducting elements of lymph nodes. The most frequent form of obstructive lymphatic disease is lymphedema.

In this chapter, we (a) discuss vascular diagnostic testing and focus on selected arterial, venous, and lymphatic disorders; (b) review chronic arterial occlusive diseases including acute arterial occlusions and the most frequent vasospastic vascular disorders; (c) review current evaluation and treatment of acute DVT, varicose veins, and venous insufficiency; and (d) discuss current management of chronic lymphedema.

The rehabilitation professional is often asked to evaluate the patient with a painful, swollen, or ulcerated limb. A thorough understanding of the pathophysiology, available diagnostic testing, and clinical evaluation will help the practitioner choose the appropriate vascular diagnosis and treatment regimen.

VASCULAR DIAGNOSTIC TESTING

Vascular diagnostic testing is typically performed to confirm a clinical diagnosis and document the severity of disease. Other indications for vascular testing include monitoring disease progression, assessing outcome after an intervention, and localizing lesions to specific segments of the limb. Vascular diagnostic studies are generally classified as noninvasive (ankle-brachial index [ABI], transcutaneous oximetry) or invasive (contrast angiography, venography, lymphangiography).

Arterial Testing

Ankle-Brachial Index

The ABI provides objective data about arterial perfusion of the lower limbs (Table 45-1). Pressures are obtained using blood pressure cuffs placed around the patient’s lower calves or ankles. A hand-held Doppler detects systolic blood movement in the dorsalis pedis and the posterior tibial arteries. The brachial (arm) pressure is measured in the standard fashion. In normal individuals, there should be no inter-arm systolic pressure gradient, or this pressure difference should be minimal (<12 mm Hg). If the arm blood pressures are not equal, a subclavian or axillary arterial stenosis is present in the arm with the lower pressure. The higher of the two blood pressures is then used for subsequent blood pressure ratio (ABI) calculations. In a healthy individual, due to peripheral amplification of the pulse pressure, the ankle pressure should be higher than the brachial arterial systolic pressure; the normal ankle-to-arm systolic blood pressure ratio is, therefore, greater than 1.0. ABI values are considered to be low-normal when they are less than 1.0 and more than 0.90, mildly diminished when they are less than 0.90 and more than or equal to 0.80, moderately diminished between 0.50 and 0.80, and severely decreased when less than 0.50. An ABI identifies individuals who are at risk for developing rest pain, ischemic ulcerations, or gangrene, and it is a marker of generalized atherosclerosis (1). The risk for death, usually from a cardiovascular event, increases dramatically as the ABI decreases. The 5-year mortality rate in patients with an ABI less than 0.85 is 10%; when the ABI is less than 0.40, the 5-year mortality rate approaches 50% (2,3).

The ABI is not accurate when the systolic blood pressure cannot be abolished using a blood pressure cuff. The incidence of noncompressible (artificially high), calcified conduit arteries is highest in diabetic, elderly, and chronic renal failure patients. Despite high recorded systolic pressure, these individuals may have severe disease. Patients with severely stenotic or occluded iliofemoral arteries may also have a normal ankle pressure if sufficient collateral circulation is present. If such patients have symptomatic evidence of arterial disease, the test should be repeated after exercise. Other diagnostic tests (segmental pressure measurement, Doppler waveform analysis, or pulse volume recording [PVR]) may also be performed to rule out significant arterial occlusive disease.
Segmental Pressure Measurements
Arterial pressures can be measured using blood pressure cuffs placed at various levels (upper thigh, lower thigh, upper calf, and lower calf above the ankle) sequentially along the limb. Systolic blood pressures obtained in this manner can be indexed relative to the brachial artery pressure in a manner analogous to the ABI. Segmental pressure analysis is often used to determine the location of arterial stenoses. The presence of a significant systolic pressure gradient (>10 to 15 mm Hg) between the brachial artery pressure and the upper thigh systolic pressure usually signifies the presence of aortoiliac obstruction. A pressure gradient located between the upper and lower thigh cuff signifies obstruction in the superficial femoral artery. A gradient between the lower thigh and upper calf cuff indicates distal superficial femoral or popliteal artery obstruction. Segmental pressure measurements may be artificially elevated or unpredictable in patients with calcified or noncompressible vessels (as described with ABI). In such individuals, Doppler waveform analysis, arterial duplex studies, or transcutaneous oximetry studies may be of benefit.

Photoplethysmography
Photoplethysmography is a noninvasive optical technique used to measure changes in the cutaneous microcirculation by detecting the reflection of infrared light. The probe contains infrared light sources and a photoelectric cell to measure reflected light. Blood is more opaque to red light than the other components of the skin and subcutaneous tissue. The amount of blood under the source beam affects the absorption of light. Photoplethysmography can be used in two ways. Using the alternating current component, the pulses recorded resemble those obtained with a strain gauge. These tracings can be used to show distal arterial disease in the hands or feet. The photoelectric cell also documents changes in extremity flow, with positional changes of the arm. This may be used to evaluate possible arterial compression at the thoracic outlet. A second method is the direct current coupling component. The DC component of the signal varies slowly and reflects variation of total blood volume of the examined tissue. With this mode, blood volume changes are recorded without major distortion. This application is used to record venous refilling after exercise (see “Venous Plethysmography”).

Continuous Wave Doppler
Waveform analysis can provide important information that may confirm arterial patency or identify occlusive lesions (Fig. 45-1). In many circumstances, a change in blood velocity or pulse waveform such as a change from a triphasic to monophasic waveform provides reasonable, accurate information about the location and extent of specific lower extremity lesions. Doppler waveform analyses are reliable even in highly calcified vessels that are not amenable to pressure determinations.

### TABLE 45.1 Diagnostic Arterial Testing

<table>
<thead>
<tr>
<th>Diagnostic Tests (Arterial)</th>
<th>Skin Perfusion</th>
<th>Patency</th>
<th>Determine Location of Stenosis/Occlusion</th>
<th>Evaluate Aneurysm</th>
<th>Monitor Disease Progression</th>
<th>Accurate in the Presence of Noncompressible Calcified Arteries</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ankle-brachial index</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Segmental pressures</td>
<td></td>
<td>+/-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Continuous wave Doppler</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pulse volume recording</td>
<td></td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Photoplethysmography</td>
<td>+</td>
<td>+/-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transcutaneous oximetry (TcPO2)</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Duplex scan</td>
<td></td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Computed tomography</td>
<td></td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Magnetic resonance angiography</td>
<td></td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contrast angiography</td>
<td></td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Whether the waveform changes shape or pulse dampening occurs at a particular level (5). When an altered pulse volume waveform is present, it can be inferred that there is a hemodynamically significant lesion proximal to the site of the cuff.
Transcutaneous Oximetry

Transcutaneous oximetry (TcPO₂) determinations provide a very sensitive means to assess skin perfusion and the potential for cutaneous healing at a specific site (6). TcPO₂ measurements are relatively simple and reproducible. Surface oxygen-sensing electrodes calibrated to 45°C are attached to the skin and allowed to equilibrate before recording the TcPO₂ value. The feet are then elevated to 30 degrees for 3 minutes, and the TcPO₂ values are again recorded. Normal TcPO₂ values are greater than 50 to 60 mm Hg. TcPO₂ values less than 20 to 30 mm Hg suggest severe local ischemia and bode poorly for future wound healing (7). A decrease in the TcPO₂ value of 10 mm Hg with an elevation is significant (6) and suggests tenuous perfusion.

Duplex Scanning

Duplex scanning using B-mode imaging combined with directional Doppler can visualize and assess arterial aneurysms and detect flow velocity changes at sites of localized stenosis or occlusion. Duplex studies can assess plaque morphology, surgical graft patency, and establish the presence of arteriovenous fistulae. This technique requires a technically proficient examiner, may require extensive time for a complete examination, and is significantly more expensive than most physiologic noninvasive testing methods. Duplex scanning is particularly helpful in assessing proximal iliofemoral stenosis that may be amenable to angioplasty, providing follow-up data to assess continued patency of both venous and prosthetic arterial grafts, and evaluating the patency of prior angioplasty sites or intravascular stents.

Imaging Techniques

Technological advances are enabling computed tomography (CT) and magnetic resonance angiography (MRA) to replace conventional angiography as a means of identifying arterial stenoses and occlusions.

CT Angiography

During the past decade, CT angiography (CTA) has become a standard noninvasive imaging modality for vascular anatomy and pathology. With continued improvement in spatial resolution, CTA is now the mainstay for preoperative imaging of abdominal aortic aneurysms. It provides accurate information not only of the size of an aneurysm but also the exact location and critical measurements needed for repair.

CT with three-dimensional reconstruction provides a global view of the chest, abdomen, and associated large vessels (Fig. 45-2). The accuracy is less operator-dependent when
compared to ultrasound and is especially useful in patients with a large body habitus. CTA requires radiation exposure for image acquisition.

**Magnetic Resonance Angiography**

MRA can be used to determine the morphology of blood vessels, assess blood flow velocity, evaluate the lumen for the presence of thrombosis, and evaluate for the presence of hemorrhage, infection, or the status of the end organ. MRA has been found to have a sensitivity of 99.6%, a specificity of 100%, a positive predictive value of 100%, and a negative predictive value of 98.5% for detecting patent segments, occluded segments, and hemodynamically significant stenosis in the aorta, iliac, and femoral vessels (8). Unlike ultrasound, MRA is not compromised by overlying bone, bowel gas, or calcification. When MRA is compared with conventional contrast angiography in preoperative studies of the aorta, iliac artery, and femoral artery, the two imaging modalities are concordant in almost all cases (8). MRA is relatively expensive and its use limited in situations in which metallic instrumentation may be required. MRA is the optimum imaging alternative in pregnant patients and patients with severe iodinated contrast allergy.

Magnetic resonance studies using gadolinium have a long safety record with little nephrotoxicity at the doses used. Recent reports that gadolinium may play a role in inducing nephrogenic systemic fibrosis (NSF) are a concern. Although rare, NSF can be catastrophic. Caution is recommended with reduced glomerular filtration rate (GFR) (definitely, a GFR <30, possibly <60) (9,10).

**Contrast Arteriography**

Contrast angiography has been the traditional “gold standard” for lower extremity arterial evaluation. Angiography remains the definitive approach for preoperative evaluation in patients requiring revascularization. Pre-procedure arteriography is an essential part of endovascular procedures. Digital subtraction angiography allows enhanced visualization of the structures of interest.

Conventional angiography is associated with an overall minor and major complication rate of about 8%. Most of the complications result from the side effects of the iodinated contrast material and access site complications. Patients with preexisting renal insufficiency, diabetes, or dehydration are at greatest risk for contrast-induced renal failure. To minimize the risk of nephrotoxicity, bicarbonate hydration and oral acetaminophen can be used starting the day prior to the procedure. The risk for contrast reaction varies from 0.04% to 0.22% (11,12). The arterial puncture necessary for the study may be associated with bleeding, hematoma, pseudoaneurysm, and pain at the site. There is also modest risk for contrast allergy, which can be associated with anaphylaxis, or death, and contrast-induced worsening of renal function (11).

**Venous Testing**

**Continuous Wave Doppler**

Continuous wave Doppler (described previously) is also used clinically to test the integrity of the venous system (Table 45-2). This method can identify the presence of venous obstruction or incompetence, quantify severity of the venous disease, and roughly localize these abnormalities to a particular segment of the limb. The venous flow signal is obtained at several sites in the limb. Normal venous flow is spontaneous and phasic with respiration. With continuous wave Doppler, the patency, spontaneous flow, phasicity, augmentation, competency, and pulsatility of the venous flow are determined and graded. Obstruction of a vein is characterized by the absence of normal spontaneous venous flow or by the loss of phasic variation with respiration. If the Doppler probe is placed directly over an obstruction, there is absence of spontaneous flow. If the probe is placed below the site of obstruction, there is a loss of phasic change in the venous flow with respiration (a monophasic low-frequency signal). Several maneuvers (deep breathing, Valsalva, and distal compression of the calf or forearm) can produce augmentation of venous flow. Because continuous wave Doppler provides subjective information, if positive for obstruction, findings should be followed by an objective test.

Venous Doppler ultrasound examination is portable and inexpensive. Validation studies of continuous wave Doppler ultrasound for lower extremity deep vein thromboses (DVT) report a sensitivity of 31% to 96% and a specificity of 51% to 94% (13,14). Limitations with venous Doppler ultrasound include (a) the need for examiner expertise; (b) only thrombi in the major deep veins in direct continuity with the heart can be detected (an isolated thrombi in

<table>
<thead>
<tr>
<th>TABLE 45.2 Diagnostic Venous Tests</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Diagnostic Tests (Venous)</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Continuous wave Doppler</td>
</tr>
<tr>
<td>Duplex ultrasound</td>
</tr>
<tr>
<td>Plethysmography</td>
</tr>
<tr>
<td>D-Dimer</td>
</tr>
<tr>
<td>Contrast venography</td>
</tr>
<tr>
<td>MR or CT imaging</td>
</tr>
</tbody>
</table>

*a*Acute DVT only.
tributaries such as the internal iliac, deep femoral, peroneal, gastrocnemius, and soleus veins will likely be missed); (c) to be detected, the thrombus must produce a flow disturbance (nonocclusive thrombi may be missed because venous flow is present in anatomically adjacent collaterals); and (d) external compression of the vein cannot be distinguished from internal thrombosis. Because of these limitations, continuous wave Doppler ultrasound has largely been replaced by venous duplex scanning for the diagnosis of DVT (combines Doppler principles with real-time B-mode and color-flow ultrasound imaging).

Venous Plethysmography
Plethysmography is a noninvasive method of detecting blood volume changes in an extremity (Fig. 45-3). Plethysmographic techniques have been developed to measure the changes in limb volume that occur when venous return is enhanced (venous insufficiency) or impeded (venous obstruction). With normal venous outflow, positional change (leg elevation) or release of an externally inflated cuff results in rapid emptying of the leg veins. If the valves are competent, refilling occurs in an antegrade fashion through the arteries and capillaries. In normal individuals, this takes a minute or more.

With venous incompetence, the leg volume returns to baseline more rapidly than normal. If the incompetence is primarily superficial, tourniquets placed around the leg or using a finger to compress an incompetent superficial vein will normalize the refilling time. In venous obstruction, the peripheral venous pressure and baseline venous volume are elevated (increased venous outflow resistance). In this case, leg elevation or a rapid release of the cuff results in slower emptying of the leg veins.

Segmental plethysmography uses a sleeve to demonstrate changes in limb volume. Sequential timed inflation and deflation of a proximal cuff produces changes in limb volume, which are measured to determine the venous capacitance and maximal venous outflow. The presence of a proximal lower extremity DVT causes (a) minimal change in the limb volume with cuff inflation and (b) a decreased change in limb volume when the cuff is released.
**Duplex Ultrasound**

Duplex ultrasound (a) directly visualizes and locates intraluminal obstruction; (b) assesses the characteristics of venous flow distal to the inguinal ligament; (c) identifies the presence of collateral veins around an obstructed venous segment; (d) permits direct detection of valvular reflux; (e) allows visualization of specific venous valves and valve leaflet motion; (f) quantitates the degree of incompetence; (g) locates and assesses veins before harvest for bypass procedures; (h) evaluates venous perforator incompetence; and (i) evaluates conditions that may mimic venous disease. Duplex scanning has become the method of choice for testing individual veins of the superficial, deep, and perforating systems.

**D-Dimer**

D-Dimer is a global indicator of coagulation activation and fibrinolysis. D-Dimer level is measured in plasma by use of well-standardized assays that are widely used for the diagnosis of acute venous thromboembolus (VTE). In patients with a first VTE, measuring D-Dimer level allows a global assessment of their thrombotic tendency and a stratification into high- and low-risk patients with regard to risk of recurrence. Patients with a first spontaneous VTE and a D-Dimer level of less than 250 ng/mL after withdrawal of oral anticoagulation have a low risk of VTE recurrence (15). In patients with thrombosis with a very low risk of recurrence, extensive screening for thrombophilic risk factors may be unnecessary.

**Contrast Venography**

Lower extremity contrast venography remains a powerful, but increasingly used tool, in the evaluation of both acute and chronic DVT. At this time, ascending leg venography is rarely used to confirm acute DVT. With advances in duplex ultrasonography, venography has been largely replaced by duplex scanning to evaluate patients with suspected deep venous obstruction or incompetence. It is indicated when a high clinical suspicion of thrombosis is present and noninvasive testing is negative or equivocal (16). In chronic venous disease, ascending venography demonstrates the location and extent of postthrombotic disease, as manifested by occlusion, venous recanalization, collateral channels, and superficial varicosities. Ascending venography may also help with planning of endovascular and open surgical procedures such as iliac and Inferior vena cava (IVC) recanalizations and venous bypass grafts. Ascending contrast venography is performed primarily in patients with significant chronic deep venous occlusive disease who are candidates for endovascular treatment with stents, surgical venous bypass, venous valve repair, or valve transplantation. Descending venography is used in concert with ascending venography to distinguish primary valvular incompetence from thrombotic disease. Descending venography identifies the level of deep vein reflux and evaluates the morphology of the venous valves. With decreasing experience using this technique, the “gold standard” is uninterpretable in 5% to 15% of patients (17). Unfortunately, the numbers of poor-quality examinations will likely increase in the future as the need for venography continues to decline.

**CT Venous Imaging**

The two advantages of CT in venous imaging are the speed and resolution of image acquisition. The disadvantages include radiation exposure and the necessity for administration of iodinated contrast material. Patients with a significant allergic reaction to iodinated contrast material or decreased renal function should be evaluated with alternate imaging techniques.

**Magnetic Resonance Imaging**

The effectiveness of MRI for detection of DVT has been compared with that of contrast venography in a number of trials. Sensitivity and specificity values as high as 100% have been reported (18). MRI can also be used to distinguish acute from chronic DVT. Because MRI is more expensive than a duplex scan, it is rarely used to diagnose DVT.

**Lymphatic Testing**

**Lymphoscintigraphy**

Lymphoscintigraphy, performed by injecting a radioactive colloid and observing uptake into the lymphatic system, has become the standard evaluation tool to establish lymphatic flow patterns (19,20). This test can be performed in both upper and lower extremities by injecting the colloid between the digits of the hands or toes, respectively. Lymphoscintigraphy assesses the most basic function of the lymphatic system, namely the clearance of interstitial macromolecules that are too large to reenter the blood capillaries directly. When lymphedema is present, the images often show a highly characteristic dermal backflow pattern or re-routing of tracer away from the main lymphatic trunks and into fine collateral lymphatic vessels of the skin. Images may also suggest distal or proximal lymphatic obliteration, hyperplasia, or aplasia/hypoplasia of lymphatic vessels. Because lymphatic disease occurs so rarely, the skills for administration and accurate interpretation of this test are often available only in larger medical centers where higher volumes of testing occur.

**Lymphangiography**

Lymphangiography should not be considered unless specific surgical intervention is being contemplated. It is invasive and may damage the (remaining) lymphatic vessels (21).

**Lymphatic Imaging**

The studies described earlier help distinguish venous disease from lymphatic disease. CT scan of the abdomen is helpful to discover the underlying obstructive pathology and should be part of an evaluation of new edema in the lower extremity (22,23). MR imaging is useful in studying the swollen limb without obvious etiology because fluid, fat, soft tissue, and tumor can all be identified (24). Ultrasound imaging may provide accurate measurement of clinical lymphedema in the future.
Chronic arterial occlusive disease with ischemia, arterial thromboemboli, these conditions account for lower extremity arterial insufficiency (thromboangiitis obliterans, arterial thromboemboli), these conditions account for 2.7% to 5% of limbs with claudication annually. Amputation is required in 1% per year (26).

**Treatment**

**Medical**

The management of patients with intermittent claudication has traditionally focused on relief of symptoms. The goals of medical care should be to reduce cardiovascular risk and alleviate symptoms of intermittent claudication. Medical therapies can effectively modify both the natural history of atherosclerotic lower extremity arterial occlusive disease and significantly reduce the morbidity of this disorder.

Screening for elevated homocysteine should be considered in young patients with peripheral arterial disease (PAD). An elevated plasma homocysteine level is emerging as a prevalent and strong predictor for atherosclerotic vascular disease in the peripheral, coronary, and cerebral vessels and is recognized as an independent predictor of PAD (27). An increased plasma total homocysteine level confers an independent risk for vascular disease similar to that of smoking or hyperlipidemia. It further increases the risk associated with smoking and hypertension (28). Elevated homocysteine levels can be lowered by folic acid and other vitamin supplementation; however, no studies to date have examined how this treatment affects atherosclerosis or intermittent claudication symptoms (29).

Among apparently healthy men, elevated baseline levels of C-reactive protein (CRP), a marker for systemic inflammation, may predict future risk for developing symptomatic peripheral arterial occlusive disease. CRP may serve as a molecular marker for underlying systemic atherosclerosis (30).

**Risk factor management.** All patients presenting for treatment of peripheral arterial occlusive disease should have their risk factors rigorously assessed (31). Patients with known PAD should be treated aggressively with a combination of a HMG CoA reductase inhibitor (statin), an angiotensin-converting enzyme (ACE) inhibitor, an antiplatelet agent, and a β-blocker (if there is a history of coronary disease). They should also control their blood pressure and blood sugar level. Smokers should be encouraged to stop smoking (32). On average, an age-matched control group has an all-cause mortality rate of 1.6% per year. This rate increases to 4.8% per year for patients with PAD. Cardiovascular mortality rates are similarly affected, with an overall event rate of 0.5% per year in controls and 2.5% per year in patients with PAD. The presence of PAD is an independent risk factor for mortality even when other known risk factors are controlled (32–35). Treatment needs to focus on both the effects of atherosclerosis in the peripheral circulation.
and the systemic nature of the disease. Appropriate therapy should be instituted to decrease the risk for both peripheral progression and cardiovascular mortality.

The increased cardiac event rate in patients with PAD underscores the importance of intensive medical management to reduce the risks for cardiovascular morbidity and mortality.

**Diabetic management.** The effect of diabetic management on large vessel arterial occlusive disease has not yet been evaluated in a controlled prospective clinical trial. Optimal diabetic management is presumed to improve the rate of lower extremity disease progression and to decrease the incidence of wound infection, gangrene, and amputation (36).

**Nicotine cessation.** Cigarette smoking is an exceptionally positive risk factor for lower extremity PAD. It is 2 to 3 times more likely to cause lower extremity PAD than coronary artery disease (37). Cigarette smoking nearly doubles the risk for progression of peripheral arterial occlusive disease, independent of other associated risks factors (31). Patients should be informed that continued tobacco use is likely to accelerate disease progression and cause progressive symptomatic worsening. Eighteen percent of patients with claudication who continue to smoke cigarettes develop rest pain over the subsequent 5 years of observation (38). In contrast, in those patients who stop smoking, the development of rest pain is exceedingly rare. The 5-year mortality rate for patients with claudication who continue to smoke may be as high as 40% to 50%.

**Lipid management.** Effective lipid management should be considered a mandatory component of the medical therapy of patients with objective evidence of atherosclerotic peripheral arterial occlusive disease. Patients should be treated with diet and pharmacologic therapy to achieve low-density lipoprotein (LDL) cholesterol levels of less than 100 mg/dL (39). Statins have favorable effects on multiple interrelated aspects of vascular biology important in atherosclerosis. In particular, they have beneficial effects on inflammation, plaque stabilization, endothelial dysfunction, and thrombosis. Statins have also been shown to be beneficial in acute vascular events.

**Hypertension management.** The goal of treated hypertension in patients with PAD should be similar to that in patients who have other cardiovascular disease. Antihypertensive therapies should be administered to hypertensive patients with lower extremity PAD to achieve a goal of less than 140 mm Hg systolic/90 mm Hg diastolic (nondiabetics) or less than 130 mm Hg systolic/80 mm Hg diastolic (diabetics and individuals with chronic renal disease) to reduce the risk of MI, stroke, congestive heart failure, and cardiovascular death (36,37).

**ACE inhibitors.** ACE inhibitors have been shown to reduce cardiovascular morbidity and mortality rates in patients with PAD by 25% regardless of the presence or absence of hypertension (28).

**Antiplatelet therapy.** Antiplatelet therapy may decrease the rate of atherosclerotic disease progression, decrease the incidence of thrombotic events in the limbs, and decrease the rate of adverse coronary and cerebral vascular ischemic events. Aspirin in doses of 75 to 325 mg is recommended as safe and effective antiplatelet therapy to reduce the risk of MI, stroke, or vascular death in individuals with atherosclerotic lower extremity PAD (36). Patients with documented arterial occlusive disease may benefit from antiplatelet therapy unless otherwise contraindicated. The relative benefits of newer antiplatelet drugs on limb ischemic event rates and patient survival are presently under investigation. Ticlopidine and clopidogrel are thienopyridines that selectively inhibit the adenosine diphosphate (ADP) platelet receptor with no direct effects on arachidonic acid metabolism (40,41). Ticlopidine has been evaluated in patients with intermittent claudication. In a multicenter, randomized, controlled trial of patients undergoing treatment with ticlopidine versus placebo (STIMS), the need for subsequent vascular surgery was reduced by about 50% over a 7-year time period (42). The use of ticlopidine has dropped markedly owing to blood aplastic side effects. In platelet-aggregation studies, clopidogrel, 75 mg once daily, produced inhibition of ADP-induced platelet aggregation equivalent to that of ticlopidine, 250 mg twice daily. Long-term administration of clopidogrel in patients with atherosclerotic vascular disease has been reported to be more effective than aspirin in reducing the combined risk for ischemic stroke, myocardial infarction, or vascular death (43).

**Prostaglandins.** Prostacyclin and its analogues prostaglandin E1, prostaglandin I2, and iloprost (a prostaglandin I2 analogue) have been shown to improve perfusion through (a) inhibition of platelet aggregation, (b) decreased leukocyte migration and activation, (c) vasodilation (possibly through an effect on resting sympathetic tone), and (d) profibrinolytic effects. Iloprost is typically infused for 6 hours per day for 28 days (2 ng/kg/min). Iloprost was shown to decrease the probability of dying or requiring major amputation during treatment and the subsequent 3 to 6 months (44). Oral vasodilator prostaglandins such as beraprost and iloprost are not effective medications to improve walking distances in patients with intermittent claudication (36).

**Vasodilator drugs.** In general, vasodilator drugs do not improve symptoms in patients with arterial claudication. Direct-acting vasodilators have minimal effect at the focal atherosclerotic site. Vasodilator drugs do not vasodilate lower extremity collateral vessels. In addition, this class of medications may elicit a fall in blood pressure and limb perfusion pressure if a preferential vasodilatory effect occurs in other nondiseased circulation (45).

**B-Blockers.** Although β-blockers were previously believed to have detrimental clinical effects in patients with claudication, clinical trials have demonstrated a symptom-neutral effect for these agents in most patients (46). Because β-blocker therapy may be efficacious for the treatment of associated coronary artery disease or myocardial infarction, these drugs do not need to be empirically withdrawn from the patient with claudication.

**Agents for intermittent claudication.** Cilostazol inhibits the action of phosphodiesterase and increases the amount of intracellular cyclic adenosine monophosphate. This results
in significant platelet and vasodilatory activity as well as antiproliferative properties. Since antiplatelet and vasodilatory drugs have been shown to have no positive effect on claudication-limiting walking distance, the mechanism by which cilostazol achieves improvement in PAD patients remains speculative (47). Cilostazol (100 mg orally two times per day) is indicated as an effective therapy to improve symptoms and increase walking distances in patients with lower extremity PAD and intermittent claudication (in the absence of heart failure) (36). This medication is contraindicated in patients with congestive heart failure.

Pentoxifylline has received variable reports of success in patients with arterial occlusive disease (48). Minimal efficacy and caffeine-like side effects limit use of this medication.

**Antioxidants.** Antioxidant agents may render LDL cholesterol resistant to oxidation and make that lipid fraction less atherogenic. In addition, antioxidants may improve endothelium-dependent vasodilation by reducing oxidative degradation of nitric oxide (49). A number of descriptive and case-controlled studies have shown an association between antioxidant agents (vitamins E, C, and β-carotene) and reduction in cardiovascular events (50). Further studies are required to evaluate if oxidative stress and antioxidant status are implicated in the clinical progression of disease and to define the formulation and dosing of antioxidant vitamins. Although the basic science is promising, large randomized controlled trials have yet to show a compelling agent that will bring these clinical effects to fruition.

In summary, although there are few widely accepted pharmacologic interventions for PAD, current recommendations are that all PAD patients should receive antiplatelet therapy, stop smoking, exercise, and be screened and treated for hyperlipidemia, hypertension, diabetes, and hypercoagulability in accordance with national guidelines and community standards (51).

**Rehabilitation**

**General self-care measures.** Patients with PAD should be instructed to wear protective footwear at all times (never walk barefoot or in socks) and monitor their extremities carefully for redness or skin break down. Extremes of temperature should be avoided. The feet should be washed carefully with mild soap and warm water. Drying is best performed by blotting or patting with a soft clean towel (rubbing should be avoided because it may injure the skin). The skin between the toes should be carefully dried to avoid maceration. Emollients without preservative or perfume should be used (avoid between the toes) to prevent cracking of the skin. Proper footwear, which avoids producing areas of point pressure, should be used. Whenever new shoes are purchased, the patient should gradually (over a period of a week) wear-in shoes to make sure there are no areas of point pressure with the new footwear. Warm outer footwear should be used in the winter to protect against cold injury.

Decreased activity secondary to symptomatic lower extremity arterial occlusive disease can result in deconditioning, which further contributes to disease impairment. Deconditioning may also be “iatrogenic” as a result of a prolonged period of limited mobility to avoid trauma to ischemic wounds.

**Exercise.** Regular exercise training produces a reduction in the inflammatory markers associated with endothelial damage (37). Evidence suggests that patients following an exercise regimen improve both their claudication distance and cardiovascular risk profile (51). Exercise training may elicit improvements in maximal walking time of 25% to 200% (52). On average, patients can expect to double their intermittent and absolute claudication distances. Exercise training has been postulated to improve performance in patients with claudication by directly augmenting limb flow, improving blood viscosity, biomechanically improving the efficiency of gait, and altering the ischemic pain threshold or tolerance. To optimize the benefits of an exercise program, patients should receive a structured claudication exercise rehabilitation program for at least three sessions weekly over a period of 12 weeks (52–56). Continued improvement can be seen over 24 weeks of training. Strength training, whether sequential or concomitant, does not augment the response to a walking exercise program (57). The optimal exercise program for improving claudication pain distances in patients with PAD is intermittent walking to near maximal pain during a program of at least 6 months. Such a program should be a part of the standard medical care for patient with intermittent claudication (58). Patients should be instructed to walk until claudication occurs, rest until it subsides, and continue, repeating the cycle for 1 hour each day. Improved walking performance has also been demonstrated through upper limb aerobic exercise training in patients with PAD (59).

**Gene Therapy**

Molecular therapies to induce angiogenesis are appealing in the claudicant population because: ischemia is subacute, time is available for angiogenesis to occur, and collateral development is associated with increased walking distance. Molecular therapies that result in increased levels of vascular endothelial growth factor, fibroblast growth factor, and hepatocyte growth factor have been used in claudication populations. The RAVE trial (regional angiogenesis with vascular endothelial growth factor in PAD) concluded that a single IM administration of endoviral vascular endothelial growth factor was not associated with improved treadmill exercise performance or quality of life over placebo (60).

**Revascularization**

Previously, surgical revascularization was considered for patients with rest pain, pending tissue loss, or significant limitations of lifestyle who failed medical treatment. Endovascular intervention coupled with aggressive proactive medical management is replacing these conventional paradigms (51).

**Endovascular.** Endovascular therapy is a broad term that encompasses several treatment modalities: percutaneous transluminal angioplasty (PTA), stenting, stent-grafting (e.g., Viabahn and aortic stent grafts), atherectomy, cryoplasty, cutting-balloon angioplasty, and laser-directed atherectomy.
PTA is indicated for focal stenosis or short segmental occlusions in which the adjacent vessels are relatively free of disease. A localized stenosis of the common iliac artery (<5 cm in length) is the most favorable situation for angioplasty (61). Iliac PTA for focal iliac disease is also a valuable adjunct when combined with distal surgical revascularization in appropriate patients with multilevel disease (62). Angioplasty is a controlled injury to the vessel wall (51). Smooth muscle cell proliferation within the media (normal <1%), increases to more than 20% within 48 hours after angioplasty. After balloon angioplasty, there is thrombosis formation, intimal hyperplasia development, elastic recoil, and remodeling. In contrast after stent placement, elastic recoil and remodeling are eliminated, and thrombosis followed by intimal hyperplasia is the main contributor to in-stent restenosis (51).

Endovascular stents were introduced to help resolve the problems of residual stenosis, elastic recoil, flow-limiting arterial dissection, and to improve patency rates after balloon angioplasty. The response of a vessel to a stent is dependent on the stent design, length, composition, delivery system, and deployment technique (51). Initial results of stents placed for treatment of occlusive atherosclerotic disease in femoral and popliteal arteries showed promising results. Subsequent studies demonstrated that exaggerated neointimal hyperplasia in the stent-treated segment frequently leads to in-stent restenosis. The biology of in-stent restenosis is different than that seen after balloon angioplasty. In a retrospective study, nitinol stents significantly improved primary patency rates in femoropopliteal arteries compared with stainless steel stents (63). In the intermediate term, a randomized controlled study has shown superior results with nitinol stents compared to PTA with the option of secondary stenting (64). The policy of routine primary stenting versus angioplasty for femoropopliteal lesions remains controversial. Primary stenting adds to the cost and treatment failures are more difficult to manage than those with angioplasty alone (65).

The potential role of radiation in the prevention of restenosis has been evaluated. Ionizing radiation inhibits cellular proliferation and has been used in the treatment of neoplastic and non-neoplastic diseases. Endovascular brachytherapy has shown a delay but not an inhibition of restenosis when compared to PTA alone (66).

Cryoplasty is a novel therapy that combines conventional balloon angioplasty with application of cryotherapy. The hope is that cryotherapy will allow more accurate angioplasty; reducing dissections and vessel response to injury (51). This modality may have a role in the treatment of lesions across the knee joint or in patients with restenosis after standard angioplasty.

There is decreasing long-term effectiveness of PTA in more distal vessels because of the reduction in both artery size and flow rate. It is reasonable to consider distal angioplasty for limb salvage in patients at high risk for limb loss who are not surgical candidates. Although the rate of long-term clinical success appears to be less than that for conventional surgical reconstruction, the benefits in terms of decreased morbidity and probable cost savings may well justify the use of PTA in these circumstances. At present, PTA is generally not recommended for patients with diffuse iliac artery disease unless they are extraordinarily poor surgical candidates.

**Subintimal recanalization.** Subintimal recanalization of infrainguinal occlusions is a minimally invasive percutaneous endovascular technique that allows revascularization of occluded infrainguinal arteries by creating a new lumen between the intima and the adventitia of the arterial wall. Unlike conventional PTA, subintimal angioplasty displaces the atheromalous and calcified intima and media to one side of the vessel lumen thereby producing a relatively smooth neolumen. The catheter and wire are redirected to the true arterial lumen distal to the area of occlusion. Subintimal recanalization can be used in patients with critical limb ischemia (67). These patients frequently have multiple comorbidities and consequently long-term survival is poor, irrespective of the technique used to revascularize the ischemic limb. The success rate of subintimal recanalization has been reported to be 78% to 90% depending on the length of the occluded arterial segment treated. Patency rates are approximately 50% at 1-year follow-up, with higher (70% to 80%) limb salvage rate. The most common complications include peripheral embolization, inadvertent ruptures, and bleeding. Other alternative endovascular modalities in patients with long-segment disease include percutaneous atherectomy devices (Silverhawk) and laser-directed atherectomy.

**Endovascular aortic aneurysm repair (EVAR).** Interventions for repair of abdominal aortic aneurysms have undergone significant evolution since the introduction of EVAR more than a decade ago. The aim of open surgical or endovascular treatment of abdominal aortic aneurysms is to prevent rupture. The size criteria and indications for repair are the same for both open or endovascular repair. Initial results with endograft placement for abdominal aortic aneurysm were promising yet issues regarding durability prohibited widespread use of the original devices. Anatomic considerations that traditionally preclude repair are now being overcome with these advancing technologies. The complications of endovascular repair include groin wound complications, stent-graft migration, and leak. As the technology has evolved, less device-related complications are being seen. The need for continuous postoperative surveillance, the high cost of the device, complications, re-interventions, late ruptures, and lack of evidence of improved late mortality or a better quality of life continue to raise questions of the durability and specific indications of EVAR. EVAR should be offered for high risk or elderly patients with suitable anatomy for endograft repair and selected patients with symptomatic, ruptured abdominal aortic aneurysm (68). With the introduction of fenestrated and branched graft technology, an increasing number of patients will undergo endovascular repair in the near future.

**Surgical.** Open repair remains the gold standard for the treatment of abdominal aortic aneurysm. The mortality rate from aortic aneurysm repair has traditionally been very low (<5%). Based on prospective randomized multicenter studies, surgical treatment is recommended over observation for males.
with an abdominal aortic aneurysm of 5.5 cm in diameter or greater and for females who have an aortic aneurysm of 5.0 cm or greater in diameter (69). If the aneurysm is smaller, observation with serial ultrasounds or CT scans every 6 months is warranted. The “normal” growth rate is approximately 3 to 4 mm per year (36). Open surgery (transabdominal, retroperitoneal, or mini-laparotomy) should be offered as the first option to low risk patients less than 70 years of age.

Ischemic rest pain and tissue necrosis, including ischemic ulceration or gangrene, are well accepted as indicators of advanced ischemia and threatened limb loss. Without treatment, most limbs with these symptoms experience disease progression and require major amputation. These symptoms represent an unequivocal indication for arterial revascularization, if anatomically feasible. Large vessel bypass surgery with synthetic graft material is well established and durable. Aortobifemoral grafting continues to be regarded as the gold standard for the treatment of aortoiliac occlusive disease (70). If early and long-term patency is to be achieved, it is important that the site for vascular reconstruction has a relatively unobstructed inflow and a patent distal runoff.

The purpose of profundoplasty is to relieve a significant stenosis or an occlusion of the proximal portion of the deep femoral artery in order to restore its function. The procedure is employed as an adjunct to an inflow procedure or when patients present with an occluded limb of an aortobifemoral bypass. The profunda femoris artery is often the only major outflow vessel in the groin. In this case, a profundoplasty essentially ensures continued patency of the aortofemoral limb. In addition, profundoplasty can be performed to optimize healing after transfemoral amputation.

Judicious selection of the appropriate method of infrainguinal reconstruction for a given patient requires an appreciation of the results obtained with all available approaches. Although PTA may be appropriate for some patients with short-segment lesions, and profundoplasty may be effective in others, most patients with an ischemic limb require conventional surgical bypass. Most claudicants achieve sustained relief, and 80% to 90% of limbs threatened with critical ischemia are salvaged with vascular reconstruction. Of variables influencing the ultimate outcome, the choice of conduit is most important. For optimal results, every effort must be made to use autogenous vein for infrainguinal reconstruction (71,72).

Infrafemoral revascularization (pedal bypass grafting) has become an accepted treatment for patients with severe distal disease, limb-threatening ischemia, and tissue loss, regardless of age or diabetic status (73,74). In situ distal bypass using reverse or intact saphenous vein has shown promising long-term patency. Patient mortality rate following revascularization averages about 5%. Patient survival rates range from 30% to 70% at 3 years (annual mortality rate following recovery from operation is between 10% and 20%) (75). If the arterial anatomy of the foot permits, autologous vein bypass should be offered to patients with severe limb ischemia before a major amputation is considered. Pedal bypass grafting has low perioperative mortality and morbidity. Postoperative surveillance with duplex ultrasonography is warranted to detect hemodynamically significant abnormalities within the graft that can be corrected before thrombosis occurs. With this, failing grafts can be salvaged, secondary patency can be improved, and the rate of late amputations diminished (76). Once thrombosis occurs, attempts to restore long-term patency are unlikely, and only modest limb salvage rates will be achieved. Patients with chronic renal insufficiency have been shown to have a poor outcome following re-intervention. Data suggest that re-intervention should not be considered when these patients present with failed grafts following an initial pedal bypass procedure (77,78).

Attempts at revascularization should be avoided in the absence of life-threatening sepsis, chronic flexion contracture, paralysis, and in patients with markedly reduced life expectancy. A multicenter randomized trial showed that coronary artery revascularization before elective vascular surgery does not alter long-term survival (79). Revascularization should be delayed in most individuals with a significant acute comorbidity (recent myocardial infarction) unless the limb is eminently threatened and a higher perioperative morbidity is acceptable (80).

**Intermittent Pneumatic Compression**

Intermittent pneumatic foot and calf compression has been shown to improve walking distance comparable with supervised exercise (81). This may be a useful device in rehabilitation centers involved in wound management after providers are properly trained in its use. Skin blood flow, as reflected by TcPO2, can be augmented acutely in ischemic limbs by intermittent venous occlusion with an externally applied inflatable cuff (82). External compression briefly raises the tissue pressure, emptying the underlying veins and transiently reducing the venous pressure without occluding arterial blood flow. The proposed mechanism to explain the increased flow is analogous to the pumping action of the calf muscle during walking (83). The transient inflation imitates the effects of normal gait by generating vigorous hemodynamic impulses throughout the veins each time the lower extremity is compressed. The pneumatic impulse enhances the venous return and causes venous pressure to decrease transiently until veins are refilled by forward flow from the arteries. An increase in the hydrostatic pressure gradient during this brief period is thought to be a major mechanism for the enhancement of arterial leg inflow. The altered flow and shear forces generated by the inflation of the pneumatic cuff may mediate the release of endothelial and humoral factors having local and systemic effects (83). A direct reduction in peripheral resistance is also postulated via release of nitric oxide secondary to shear stress across the vessel wall (84).

**Chelation Therapy**

Given its lack of efficacy and important safety concerns, EDTA should not be used to treat patients with intermittent claudication (36). Chelation therapy with EDTA has the potential for significant and even fatal complications. Nephrotoxicity
that may produce renal failure is a recognized complication of EDTA therapy. In addition, rapid infusion of EDTA may produce hypocalcemia with resultant tetany and cardiac arrhythmias.

**Acute Arterial Occlusion**
Most cases of acute arterial occlusion can be attributed to one of three causes: thrombosis, dissection, and emboli.

**Thrombosis**
Thrombosis usually occurs at the site of an underlying vascular abnormality such as an atherosclerotic lesion or an aneurysm.

**Dissection**
Dissections are associated with hypertension, atherosclerosis, connective tissue disorders, trauma, and iatrogenic causes (related to invasive diagnostic and therapeutic cardiovascular interventions) (37).

**Embolii**
Embolii large enough to occlude relatively large arteries typically have a cardiac source. The most common abnormalities causing cardiac-derived emboli include ventricular mural thrombi, valvular diseases, and atrial disorders such as chronic or paroxysmal atrial fibrillation. An unusual cause of arterial embolus includes a paradoxical embolus (a DVT that passes through an atrial septal defect, ventricular septal defect, or patent foramen ovale, and enters the arterial system). In about 5% to 10% of cases, no source of emboli is found. Emboli tend to be multiple and recurrent. Certain hypercoagulable states, such as protein C and S deficiency, the presence of antiphospholipid antibody, and malignancy have been associated with peripheral embolism.

**Clinical Presentation**
The clinical presentation of acute arterial occlusion is described as "six Ps": pain, pallor, paresthesias, paralysis, pulselessness, and polar cold. Some or all of these findings may be present. The limb is at risk if blood flow is not restored quickly. Once the tissues become ischemic, cells compensate for the lack of oxygen by converting to anaerobic metabolism. Lactic and pyruvic acids are produced and released into the circulation. If the ischemia persists, cellular adenosine triphosphate and pyruvic acids are produced and released into the circulation. An unusual cause of arterial embolus includes a paradoxical embolus (a DVT that passes through an atrial septal defect, ventricular septal defect, or patent foramen ovale, and enters the arterial system). In about 5% to 10% of cases, no source of emboli is found. Emboli tend to be multiple and recurrent. Certain hypercoagulable states, such as protein C and S deficiency, the presence of antiphospholipid antibody, and malignancy have been associated with peripheral embolism.

**Treatment of Acute Arterial Occlusion**
Ideally, all acute occlusion warrants immediate repair, although the urgency is governed by the degree of ischemia. If a patient presents with symptoms consistent with the "six Ps," a medical-surgical emergency exists, and immediate evaluation and intervention must be performed. Severe ischemia is suggested by pallor at rest, profound coolness, tender or hard muscles, and loss of motor and sensory functions. When severe ischemia is present, repair must occur within hours to salvage the limb. Immediate measures are needed to protect the lumen and restore the blood flow. The standard therapy for arterial embolies is surgery. The most common site of embolization is the femoral bifurcation. Ideal treatment consists of expeditious diagnosis of acute arterial ischemia, recognition of the embolic source, rapid systemic anticoagulation, and surgical embolectomy. Heparin is given to prevent thrombus propagation and treat the embolic source. Angiography may be required to plan a repair, and once a pre-existing occlusive or aneurysmal disease, or once the etiology is uncertain. Balloon embolectomy is performed without angiography when the embolic source is not certain and the vessel was previously normal. In acute arterial thrombosis, lysis of acute occlusion can be effective (85). If a thrombolytic strategy is elected, one must ensure that the infusion catheter can be positioned within the substance of the thrombus. An open operative approach is necessary when access to the thrombus cannot be achieved. Successful thrombolysis should be followed by endovascular or open surgical revision of any lesion unmasked after dissolution of the thrombus. Endovascular modalities such as balloon angioplasty with or without stenting can be performed at the conclusion of thrombolysis (usually through the same access site used for the infusion).

Historically, it has been thought that 4 to 6 hours (following the onset of symptoms) is the maximal length of tolerable ischemia. Patients with prior chronic limb ischemia tend to tolerate longer period of acute ischemia. The physiologic state of the limb, determined mainly by a balance between metabolic supply and demand, rather than the elapsed time from the onset of occlusion, is actually the best predictor of limb salvage.

**Other Arterial Diseases**

**Upper Extremity Ischemia**
Symptomatic vascular diseases involving the upper extremity are rare in comparison to those involving the lower extremity. However, ischemia caused by vasospasm is much more common in the upper extremities than in the lower extremities. Upper extremity arterial occlusive disease is less common, but more varied in etiology, than disease of the lower extremity; associated vasospastic and microcirculatory disorders are more common. Coldness and color changes are the most common presenting symptoms in upper extremity ischemia. Ischemia may be constant or intermittent; a manifestation of a fixed arterial obstruction, vasospasm, or both; and it may reflect involvement of large proximal arteries, small distal arteries, or the microvasculature. Impairment of the arterial circulation...
to the upper extremities may result in a variety of symptoms including weakness, intermittent vasospasm, or irreversible tissue loss with digital ulcers, skin necrosis, and gangrene.

The manifestation of acute arterial occlusion of the upper extremity depends on the location and extent of the clot, the preexisting status of the vascular bed, and the collateral flow. Occlusion of the ulnar artery may be asymptomatic if the radial and palmar arch circulation is intact. In contrast, occlusion of the axillary artery (before the bifurcation of the brachial and profunda brachial artery) will typically result in severe ischemia. Raynaud’s phenomenon, connective tissue disease, vasculitis, thromboangiitis obliterans (Buerger’s disease), erythromelalgia, acrocyanosis, livedo reticularis, pernio, frost bite, and occupational trauma may result in microvascular disease of the upper extremity (37). Arteriolar constriction is usually well tolerated. When arteriolar constriction is superimposed on fixed arterial obstruction, previously viable fingers may become ischemic.

Establishing the underlying etiology is essential for both definitive treatment and the prognosis of the patient. Upper extremity ischemia may cause a significant impairment due to decreased hand function.

**Raynaud’s Syndrome**

Raynaud’s syndrome is characterized by episodic attacks of vasospasm in response to cold or emotional stress. The fingers and hands are most often affected. In certain patients, the toes and feet may be involved. Classic episodes of vasospasm cause an intense pallor of the distal extremity followed in sequence by cyanosis and rubor on rewarming. Most patients do not experience the complete triple color response. Typically, only pallor or cyanosis is noted during attacks. Generally, the attacks are over within 30 to 60 minutes. Episodes are usually bilateral. Attacks may occur infrequently, for example, some may only have symptoms during the winter, but other patients may have a significant impairment/disability with multiple daily episodes. Digital ulcerations are rare but may occur. Females are affected more commonly than males. The mechanisms responsible for episodes of digital vasospasm remain an enigma. The symptoms may be related to abnormalities in adrenergic function, blood viscosity, or endothelial disorders (86). The normal regulation of blood flow to the fingers is affected by a number of different factors acting by local, humeral, and nervous mechanisms. Disturbance of any of these may predispose to vasospastic attacks in primary Raynaud’s. The exact mechanism of digital artery vasospasm in Raynaud’s syndrome remains unknown; however, it is felt to be due to an exaggeration of the normal thermoregulatory system. In clinical practice, it is helpful to differentiate patients with Raynaud’s syndrome as having either vasospasm or obstructive disease.

Raynaud’s disease refers to a primary vasospastic disorder where there is no identifiable underlying cause. Raynaud’s phenomenon refers to vasospasm, secondary to another underlying condition or disease. Predisposing factors include atherosclerosis, arteritis, cancer, collagen vascular disease, thoracic outlet syndrome, embolic occlusion, occupational disease, and certain medications (86). Secondary Raynaud’s phenomenon is occasionally unilateral and may produce skin breakdown. The distinction between Raynaud’s disease and Raynaud’s phenomenon has important clinical utility as it underscores the different pathologic mechanisms, treatment options, and outcomes of these two groups.

**Treatment**

Treatment of primary Raynaud’s syndrome may be difficult. Fortunately, the majority of patients with Raynaud’s disease have only mild to moderate symptoms that respond well to conservative measures. Dressing warmly, using mittens rather than gloves, heat conservation, and avoiding unnecessary cold exposure may substantially improve symptoms. Nicotine and drugs with a potential for vasoconstriction should be avoided. Biofeedback is effective in certain cases. If a secondary cause of vasospasm is not identified, often reassurance and patient education are all that is necessary. Occasionally, it is necessary for a patient to move to a warmer climate to achieve complete relief.

Pharmacologic therapy is indicated for patients with severe symptoms whose activities of daily living are affected or those who are at risk of ischemic tissue damage. A number of vasodilator medications have proved to be beneficial for patients with Raynaud’s syndrome. Calcium channel blockers are the most commonly prescribed medications for vasospasm associated with Raynaud’s syndrome. Multiple studies have shown dihydropyridines, such as Nifedipine, to be effective in reducing the frequency and severity of vasospastic attacks. Alpha-1 blockers such as doxazosin and terazosin may also be of benefit in reducing the frequency and severity of attacks. Angiotensin-converting inhibitors and angiotensin-II receptor antagonists may be of benefit in both primary and secondary Raynaud’s (86).

Novel therapies for Raynaud’s syndrome with upper extremity ischemia refractory to standard therapy include: (a) prostacyclin administered as a continuous intravenous infusion; (b) a potent endothelial-derived endothelin blocker Bosentan; and (c) an intermittent pneumatic compression (87). Interruption of the sympathetic innervation, either through ganglionic injection or surgical sympathectomy, is reserved for patients with severe symptoms.

**Vasculitic Syndrome**

Vasculitis, or angiitis, is an inflammatory disease of blood vessels. It often causes damage to the vessel wall, stenosis or occlusion of the vessel lumen by thrombosis, and progressive intimal proliferation. Vasculitic symptoms reflect systemic inflammation and the ischemic consequences of vascular occlusion. The distribution of vascular lesions and the size of the blood vessels involved vary considerably in different vasculitic syndromes and in different patients with the same syndrome. Vasculitis can be transient, chronic, self-limited, or progressive. It can be the primary abnormality or secondary to another systemic process. Histopathologic classification does not distinguish local from systemic illness or secondary from primary insult.
Rheumatoid Vasculitis
Rheumatoid vasculitis usually occurs in patients who have severe deforming arthritis and a high titer of rheumatoid factor (88). The vasculitis is mediated by the deposition of circulating immune complexes on the blood vessel wall with activation of complement. Proliferation of the vascular intima and media causes an obliterative endarteropathy. Leukocytoclastic, or small vessel vasculitis, produces palpable purpura or cutaneous ulceration (Fig. 45-5). A systemic necrotizing vasculitis, histopathologically indistinguishable from polyarteritis nodosa (PAN), complicates some cases of seropositive rheumatoid arteritis. Rheumatoid arthritis–associated polyarteritis is three times as common as primary polyarteritis. Type II cryoglobulin and vasculitis may complicate many different connective tissue diseases. Cryoglobulin should be assayed in any patient with an autoimmune disease that develops vasculitis (88).

Cryoglobulinemia
Cryoglobulins are immunoglobulins that reversibly precipitate at reduced temperatures. Type I cryoglobulins are aggregates of a single monoclonal immunoglobulin. Patients with type I cryoglobulinemia are often asymptomatic. Type II cryoglobulins are frequently associated with chronic infection (hepatitis C) and immune disorders. The typical presentation of type II cryoglobulinemia is that of a nonsystemic small vessel vasculitis with palpable purpura, urticaria, and cutaneous ulceration. Peripheral neuropathy, arthralgias, and arthritis are common.

Polyarteritis
Polyarteritis occurs by itself (PAN) or in association with another disease (secondary polyarteritis). PAN is an acute necrotizing vasculitis that affects primarily medium-sized and small arteries. It is a systemic disorder that may involve the kidneys, joints, skin, nerves, and various other tissues. Biopsy typically shows necrotizing changes and disruption of the blood vessels (89). If untreated, the systemic form of PAN is associated with a less than 15% survival rate at 5 years. With steroid therapy, the 5-year survival rate increases to more than 50% (89).

Other Vasculitides
A wide variety of other vasculitides may affect small and medium-size vessels. These include allergic angiitis (Churg-Strauss syndrome), Henoch-Schönlein purpura, various forms of hypersensitivity vasculitis, and numerous nonspecific necrotizing and nonnecrotizing vasculitides. Knowledge of the etiology of an underlying vasculitis is critical for rehabilitation professionals involved in the management of ischemic wounds.

Treatment
General measures should be performed as outlined for patients with PAD. In addition, because vasculitic processes typically involve small vessels (arterioles, venules), gentle compression with a low-stretch wrap or graduated compression stockings (20 to 30 mm Hg) may decrease associated venous congestion and enhance skin perfusion.

Thromboangiitis Obliterans (Buerger’s Disease)
Thromboangiitis obliterans (Buerger’s disease) is a segmental disease of the arteries and veins of the extremity. The first manifestation of Buerger’s disease may be superficial phlebitis. Thromboangiitis obliterans occurs predominantly in young adult male smokers. Few, if any, cases occur in the absence of tobacco use. If smoking is discontinued, the process is frequently arrested.

There are two major distinguishing features of Buerger’s disease compared with other forms of vasculitis. Pathologically, the thrombus is highly cellular, with relative sparing (less intense cellular activity) in the wall of the blood vessel. The disease usually affects the small distal arteries first and progresses proximally if smoking continues (90). In addition, the usual immunologic markers (sedimentation rate and CRP, circulating immune complexes, antinuclear antibody, rheumatoid factor, and complement) are usually normal or negative. It is not known whether the vascular lesions of Buerger’s disease are primarily thrombotic or inflammatory. In either case, the intense inflammatory infiltration and cellular proliferation seen in the acute-stage lesions are distinctive, especially when the veins are involved. Buerger’s disease is segmental in distribution. “Skip” areas of normal vessel between diseased segments are common. The intensity of the periadventitial reaction may be quite variable in different segments of the same vessel. The skip lesions may be observed angiographically and histopathologically (91).

Polyarteritis
Polyarteritis occurs by itself (PAN) or in association with another disease (secondary polyarteritis). PAN is an acute necrotizing vasculitis that affects primarily medium-sized and small arteries. It is a systemic disorder that may involve the kidneys, joints, skin, nerves, and various other tissues. Biopsy typically shows necrotizing changes and disruption of the blood vessels (89). If untreated, the systemic form of PAN is associated with a less than 15% survival rate at 5 years. With steroid therapy, the 5-year survival rate increases to more than 50% (89).

Other Vasculitides
A wide variety of other vasculitides may affect small and medium-size vessels. These include allergic angiitis (Churg-Strauss syndrome), Henoch-Schönlein purpura, various forms of hypersensitivity vasculitis, and numerous nonspecific necrotizing and nonnecrotizing vasculitides. Knowledge of the etiology of an underlying vasculitis is critical for rehabilitation professionals involved in the management of ischemic wounds.

Treatment
General measures should be performed as outlined for patients with PAD. In addition, because vasculitic processes typically involve small vessels (arterioles, venules), gentle compression with a low-stretch wrap or graduated compression stockings (20 to 30 mm Hg) may decrease associated venous congestion and enhance skin perfusion.

Thromboangiitis Obliterans (Buerger’s Disease)
Thromboangiitis obliterans (Buerger’s disease) is a segmental disease of the arteries and veins of the extremity. The first manifestation of Buerger’s disease may be superficial phlebitis. Thromboangiitis obliterans occurs predominantly in young adult male smokers. Few, if any, cases occur in the absence of tobacco use. If smoking is discontinued, the process is frequently arrested.

There are two major distinguishing features of Buerger’s disease compared with other forms of vasculitis. Pathologically, the thrombus is highly cellular, with relative sparing (less intense cellular activity) in the wall of the blood vessel. The disease usually affects the small distal arteries first and progresses proximally if smoking continues (90). In addition, the usual immunologic markers (sedimentation rate and CRP, circulating immune complexes, antinuclear antibody, rheumatoid factor, and complement) are usually normal or negative. It is not known whether the vascular lesions of Buerger’s disease are primarily thrombotic or inflammatory. In either case, the intense inflammatory infiltration and cellular proliferation seen in the acute-stage lesions are distinctive, especially when the veins are involved. Buerger’s disease is segmental in distribution. “Skip” areas of normal vessel between diseased segments are common. The intensity of the periadventitial reaction may be quite variable in different segments of the same vessel. The skip lesions may be observed angiographically and histopathologically (91).
the region of the thoracic outlet. Symptoms are thought to result from nerve compression in 90% to 95% of cases (92). The second most common presentation is venous compression. Repetitive trauma to the artery can lead to intimal damage, embolization, aneurysm formation, or acute thrombosis. Although only a small number of cases of thoracic outlet syndrome are clearly due to arterial compression, thoracic outlet syndrome is the most common cause of acute arterial occlusion in the upper extremity of adults younger than 40 years old and the most common cause of acute upper extremity venous occlusion in the young adult. The cause of arterial thoracic outlet syndrome is often a cervical or rudimentary first rib. Primary venous thoracic outlet syndrome is generally due to the costoclavicular ligament and subclavius muscle compressing the subclavian vein. Arterial involvement is suggested by obliteration of the radial or brachial pulse during provocative maneuvers of the arm (this finding may also be a normal variant). One of the most useful clinical tests is the elevation arm stress test. A patient raises his or her hands above the head and clenches and unclenches the fist for 1 or 2 minutes. The appearance of the patient’s typical symptoms with or without signs of vascular involvement (delayed capillary refilling) is suggestive of this syndrome. Duplex ultrasound or arteriography can be used to document the presence of a functional hemodynamic change with thoracic outlet maneuvers.

**Treatment**

Nonsurgical measures, including stretching (scalenes, pectorals, and trapezius muscle) without recreating symptoms, positioning, and avoidance of aggravating factors may be of benefit. If nonsurgical measures fail and symptoms warrant, surgical resection of the first rib is frequently curative. In those with significant arterial involvement (such as intimal damage or aneurysmal formation), the involved section of artery should be replaced with a graft. Advances in the management of arterial and venous complications include thrombolysis, mechanical thrombectomy, and endovascular stenting. Patients with recent venous thrombosis can have the thrombus cleared with thrombolytic therapy before rib resection. Stenting, if needed, is performed after removal of the first rib and should not be applied if the first rib has not been resected.

**Giant Cell Arteritis**

Giant cell arteritis often involves the arch, thoracic, and abdominal aorta as well as proximal branch artery.

**Temporal Arteritis**

Temporal arteritis affects older individuals and involves branches of the external carotid, brachial, and femoral arteries. Temporal arteritis primarily affects people older than 50 years (93). It predominantly involves secondary or tertiary branches from the aorta. Only about 50% of persons with temporal arteritis have headache or tender temporal arteries. Common symptoms include low-grade fever, jaw claudication, weight loss, anorexia, and other systemic symptoms (93). The erythrocyte sedimentation rate is usually elevated.

In an elderly person with a new onset of mild to moderate temporal headache, temporal arteritis should be considered. If vision loss has already occurred, emergent therapy with corticosteroids is needed. In those with no vision loss, prednisone should be initiated immediately after the diagnosis is made. Temporal biopsy is used to confirm the diagnosis before prescribing long-term prednisone therapy. If the biopsy cannot be performed immediately, prednisone therapy may be initiated until biopsy results are available. High-dose prednisone (40 to 60 mg/daily initially) is usually instituted for 3 to 6 weeks or until the sedimentation rate has stabilized in the normal range. Biopsy results remain accurate, if the biopsy is performed within several days after starting treatment with corticosteroids. Immune-modifying agents may decrease the complications of prolonged steroid suppression. The rate of steroid taper should be modified according to the sedimentation rate and the clinical picture. When surgical correction is necessary, efforts should be made to hold off until the disease has been adequately treated, or preferably until it has “burned out.” Although angiography can be used to make the diagnosis, biopsy is much more commonly used because of accessibility of the temporal artery. Advances in noninvasive imaging (CT and MR angiography) have increased their diagnostic accuracy and provide a marker to follow disease progression.

**Vibration Syndrome**

Vibratory tools such as chain saws, grinders, and jackhammers can induce hand dysesthesias and Raynaud’s phenomenon when used for several years. Symptoms initially occur during use of the instrument. Subsequently, dysesthesias and cold sensitivity persist when the vibratory tool is not being used. After several years, a late stage can be seen with frank digital artery occlusion, cyanosis, ulceration, and gangrene. Ischemia is a rare and late occurrence (94). The exact mechanism of injury is unknown. Repetitive trauma from the vibration of tools is obviously the main cause of the problem. Both the frequency of the vibration and the intensity of the trauma it produces affect the extent of damage to the endothelium (95).

**Treatment**

The treatment of vibration injuries include simple measures and medications outlined for the treatment of Raynaud’s phenomenon. In addition, vibration gloves and modifications to decrease vibratory trauma may be of benefit.

**Hypotenar Hammer Syndrome**

Occlusive disease in the hands can result from trauma to the hypothenar area caused by using the palm of the hand as a hammer in an activity that involves pushing, pounding, or twisting. This results in intimal injury to the ulnar artery as it crosses the hamate bone. When this area is repeatedly traumatized, ulnar or digital arterial spasm, aneurysm, occlusion, or a combination of these lesions can result. This may result in Raynaud’s phenomenon and digital ischemia as the artery either becomes aneurysmal and embolizes or occludes (96).
External Iliac Syndrome in Cyclists

Cyclists with high-level training and competition over several years may report claudication (typically, buttock and thigh) followed by heaviness or numbness of the limb. The symptoms often occur under maximal stress (hill climbing or sprinting) with quick relief as the pace is reduced. Pulse and ankle systolic pressures are normal at rest, but drop when symptoms are reproduced by strenuous ergometric cycling. The history and clinical findings are compatible with subclinical stenosis that becomes hemodynamically significant with maximal stress. Surgical repair includes segmental resection to shorten the artery, endarterectomy, and ligation of a prominent psoas arterial branch (that may enhance arterial lengthening) when present (97).

VENOUS DISEASE

Thrombophilia

Thrombophilia is a hereditary or acquired disorder predisposing to thrombosis.

Deep Vein Thrombosis

DVT is a serious medical condition that continues to plague immobilized patients and those rehabilitation professionals who care for them. The incidence of venous thromboembolism exceeds 1/1,000 (98). Of all venous thromboembolism patients, 30% die within 30 days (99). When a patient presents with the possibility of a DVT, predisposing risk factors such as prolonged immobilization during car or plane trips, use of estrogen, prior DVT, or family history of thrombosis should be elicited. Hypercoagulable states such as those associated with cancer or inherited coagulopathies and those caused by vessel wall injury from surgery or local trauma are major predisposing factors for thrombosis. About 30% of surviving cases develop recurrent venous thromboembolism within 10 years. Independent predictors for recurrence include increasing age, obesity, malignant neoplasm, and extremity paresis. About 28% of patients develop venous stasis syndrome within 20 years. To reduce venous thromboembolism incidence, improve survival, and prevent recurrence and complications, patients with these characteristics should receive appropriate prophylaxis (100).

As initially postulated by Virchow, three factors are of primary importance in the development of venous thrombosis: (a) abnormalities of blood flow (especially reduced flow, or stasis), (b) abnormalities of blood, and (c) vessel wall injury.

The normal hemostatic response requires interactions among the vessel wall, endothelium, platelets, and the coagulation cascade responsible for thrombin generation. Early phases of thrombosis are marked by increases in permeability followed by leukocyte adhesion, migration, and endothelial disruption (101). Stasis may facilitate endothelial leukocyte adhesion and cause endothelial hypoxia leading to a procoagulant state. In addition, stasis allows the accumulation of activated coagulation factors in areas prone to thrombosis.

Although the hemostatic system is continuously active, thrombus formation is generally confined to sites of local injury by a precise balance between activators and inhibitors of coagulation and fibrinolysis. A prothrombotic state may result from imbalances in the regulatory and inhibitory system or from activation exceeding antithrombotic capacity (102).

Thrombosis may occur anywhere but most commonly involves the deep veins of the leg. Once a thrombus forms, several events may occur: (a) the thrombus may propagate, (b) it may embolize, (c) it may be removed by fibrinolytic activity, or (d) it may undergo organization (including recanalization and retraction). An initial inflammatory response leads to fibroblast and capillary ingrowth, which helps to stabilize the thrombus. Organization occurs over weeks to months as the thrombus becomes incorporated into the vessel wall. Once luminal flow is disturbed, prograde and retrograde propagation of the thrombus may also be promoted by hemodynamic factors. The competing processes of recanalization and recurrent thrombosis determine the extent of acute DVT and its sequelae. Venous thrombi rarely lyse completely unless they are subjected to pharmacologic lysis.

The natural history and clinical consequences of a lower extremity venous thrombosis depend on the site of thrombus. Because the tibial and peroneal veins run in parallel, a thrombosis confined to one of these veins is unlikely to cause significant obstruction to the outflow of blood. Small calf vein thrombi frequently occur, especially in postoperative patients. Of those that are asymptomatic, half lyse spontaneously and are unlikely to proceed to deep vein incompetence (103). A calf thrombosis may remain asymptomatic; the thrombus often dissolves without sequelae by natural fibrinolysis. Five to 20% of calf vein thrombi propagate proximally. If extension occurs into the popliteal or more proximal veins, the chance of pulmonary embolism increases from less than 5% to about 50% (104). Thrombus isolated to the calf is less dangerous than thrombus in the thigh, but greater than 20% of such thrombi can extend proximally, and 10% embolize. Laboratory surveillance of lesions is required if anticoagulants are not used (105). Thrombi involving the proximal veins are less likely to undergo spontaneous lysis, even with heparin therapy.

Treatment

Medical

Superficial thrombophlebitis is treated with elevation and superficial heat. Compression wraps or stockings should be worn when ambulating. However, if the thrombotic process is close to the saphenofemoral junction, vascular surgery consultation is warranted for possible high ligation of the saphenofemoral junction.

Standard therapy of acute DVT consists of anticoagulation, elevation, and support of the extremity. Of primary importance is the prevention of thrombosis, with recognition of high-risk groups and effective prophylaxis of thromboembolic disease.
Anticoagulants

Anticoagulants remain the main form of treatment for thromboembolic disease. The two goals of anticoagulation in DVT are to prevent death from pulmonary embolism and to limit venous damage and subsequent postphlebitic syndrome. Prophylactic anticoagulation is warranted in patients with prior venous thromboembolism or known clotting disorders that are traumatized or undergo medical or surgical treatments with prolonged bed rest. If anticoagulation must be interrupted for a surgical procedure, baseline duplex ultrasound and surveillance ultrasounds (every 3 to 5 days) should be obtained.

Heparin. Heparin anticoagulation has been well documented as an effective treatment for thromboembolic disease. Although heparin has no fibrinolytic activity, it allows natural fibrinolysis, which stabilizes the thrombus, avoids further propagation, and relieves acute symptoms in most patients. Early complications with heparin include thrombocytopenia. Benign thrombocytopenia is a common occurrence (platelet counts dropping to 70,000 to 100,000 units/L). Rarely, a severe antibody-mediated thrombocytopenia occurs that may predispose to further thrombosis. In these patients, heparin should be stopped immediately. Heparin-induced thrombocytopenia (HIT) is more common with bovine heparin and less common with prophylactic doses or low–molecular-weight heparin (LMWH). Thrombocytopenia may occur 3 to 15 days after the initiation of heparin therapy. Acceptable non-Heparin anticoagulants include Lepirudin or argatroban. Because of significant cross-reactivity, LMWH is not an acceptable alternative for these patients. Removing Heparin alone is not sufficient to prevent the thrombotic complications of thrombocytopenia (HIT). Immediate substitution of Warfarin for Heparin can result in disastrous complications such as venous limb gangrene, skin necrosis, or worsening of a preexisting thrombosis (106).

LMWH. Comparing the safety and efficacy of LMWH with standard heparin in the treatment of acute proximal DVT, the differences between the groups in recurrent thrombosis, major bleeding, or mortality was not significant (107). LMWH has much less nonspecific binding to plasma proteins, improved bioavailability, and more predictable pharmacokinetics. LMWH dosage is based on patient body weight. Activated partial thromboplastin time monitoring is neither required nor useful. Patients with acute proximal DVT can be treated safely and effectively with LMWH at home during the initiation of warfarin therapy. Consistent evidence demonstrates that LMWH is superior to unfractionated Heparin for the initial treatment of DVT, particularly for reducing mortality and reducing the risk for major bleeding during initial therapy. Additional trials are needed to more rigorously examine the efficacy of LMWH for the initial treatment of pulmonary embolism, but systematic reviews of existing trials indicate that LMWH is at least as effective as unfractionated Heparin for these patients as well (108).

LMWH is an option for patients in whom INR control is difficult, and it may be more efficacious than oral anticoagulants in patients with cancer (109).

Warfarin. Warfarin (Coumadin) acts to inhibit the liver synthesis of vitamin K–dependent coagulation factors II, VII, IX, and X. The effect is delayed until these coagulation factors are cleared from the circulation. Factor VII has the shortest half-life (7 hours), resulting in a prolongation of the prothrombin time. Coagulation is not inhibited until other clotting factors are depleted. The American College of Chest Physicians recommends starting Warfarin with unfractionated Heparin or LMWH for at least 5 days and continuing until a therapeutic INR is achieved (110).

Warfarin also depletes natural inhibitors of coagulation, including protein C and protein S. A rare complication of warfarin therapy is skin necrosis. This has been noted in some patients with protein C deficiency. Hemorrhagic skin necrosis often involves areas where there is extensive adipose tissue (breast, buttock, thigh, or abdominal wall) and may be due to small vessel thrombosis, perhaps related to the reduction of protein C by warfarin. Gradual initiation of warfarin may prevent an abrupt drop in protein C levels.

Warfarin is contraindicated in the first trimester of pregnancy because of teratogenic effects and during the third trimester because of an increased risk for bleeding during delivery.

Thrombolytic Agents

Thrombolytic agents (streptokinase, urokinase, and tissue plasminogen activating factor) rapidly dissolve fresh (<1 week old) clots and may restore patency of the obstructed vein. Thrombolytic agents are not used routinely but may be indicated in the case of a massive proximal DVT. Thrombolysis should be considered for all patients presenting with acute DVT of the iliacaval vein (111). Thrombolysis should also be considered for young, otherwise healthy patients with extensive proximal thrombosis and threat of permanent impairment (104). After experimental acute DVT, thrombolytic therapy and balloon-catheter thrombectomy at 3 hours had similar effects on the valvular competence and the endothelial morphology. In addition, the thrombolytic therapy preserved endothelial responses to platelet products. The data suggest that thrombolytic therapy may preserve vein function after DVT and reduce the long-term potential for recurrent DVT and postthrombotic syndrome (112). If venous insufficiency develops, treatment should proceed as outlined in the section on chronic venous insufficiency.

Vena Cava Filters

Vena cava filters are indicated to prevent pulmonary embolism in patients with a venous thromboembolism who have a contraindication to anticoagulation, failed anticoagulation, VTE during pregnancy, preoperative prophylaxis, bariatric surgery, multiitrauma, neurosurgery, or orthopedic surgery with a recent DVT (1 month). Anticoagulation should be considered when the temporary contraindication to anticoagulation therapy is no longer present. There is insufficient data to support a recommendation that all filter recipients should be treated with indefinite anticoagulation regardless of the risk of recurrent
thrombosis. Until recently, all vena cava filters were permanent implants. There are now filters available that can be removed when the thrombosis risk is over or when anticoagulation can be initiated.

**Phlegmasia Cerulea Dolens**

Phlegmasia cerulea dolens is a rare complication of DVT, characterized by rapid and massive edema, severe pain, and cyanosis (104). Distal cyanosis may indicate extensive blockage to venous return. Phlegmasia cerulea dolens most commonly occurs with proximal, iliofemoral obstruction with extensive distal thrombus of deep and superficial veins. In phlegmasia, the arterial pulses may not be palpable, although anatomically, the arteries are patent. In severe cases, gangrene, which may necessitate amputation, occurs (104). Urgent treatment, including placement of a caval filter, heparinization, and surgical thrombectomy or thrombolysis if possible, is essential to minimize loss of life or limb. Thrombolytic therapy is often given from both the venous and arterial side.

**Chronic Venous Insufficiency**

Chronic venous disease is an important cause of discomfort and disability, and is present in a significant percentage of the population worldwide. A clinical score of Clinical-Etiologic-Anatomic-Pathophysiologic has been developed as a standard for reporting venous disease (113).

Many factors can result in the development of venous insufficiency, including heredity, local trauma, thrombosis, and intrinsic defects in the veins or valves themselves. Venous flow is based on a force that pushes the blood proximally, an adequate outflow, and the presence of competent valves limiting reflux. Any disruption of these components results in chronic venous hypertension (114). Normally, the pressure in the leg veins is equal to the hydrostatic pressure from a vertical column of blood extending to the right atrium of the heart. At the ankle level, the hydrostatic pressure is about 90 mm Hg (115). The pumping action of the calf muscles during exercise reduces this venous pressure by two thirds. Even slight muscular movements during normal standing will lower the pressure (115). Patients with venous insufficiency might fail to reduce ankle pressure or show a rapid return of venous pressure to resting levels at the end of exercise. The time required for the ankle vein pressure to return to resting levels after exercise is an indicator of the degree of reflux in the limb. Elevated ambulatory venous pressure is associated with an increased incidence of ulceration (116). When ambulatory venous ankle pressure is below 30 mm Hg, the incidence of ulceration is close to 0%. The incidence of ulceration increases linearly, reaching 100% when the ambulatory venous pressure is greater than 90 mm Hg (116). The superficial leg veins normally carry 10% to 15% of the venous return. Incompetent valves in the superficial veins alone rarely cause serious venous hypertension, although 10% of patients with venous ulcers have superficial venous incompetence alone.

Postthrombotic syndrome (venous occlusion or valve destruction after thrombosis) develops in 67% to 80% of patients after DVT, although ulcers form in less than 5% of the patients (117). Venous insufficiency may present up to 5 to 10 years after the resolution of the acute episode of thrombosis. Postthrombotic damage in the deep veins is an important cause of chronic venous insufficiency. Advanced venous insufficiency develops only when the valves in the perforator or deep veins are also incompetent. Ultimately, venous hypertension is the result of valve damage and retrograde venous blood flow to the superficial veins. The true mechanism by which venous hypertension leads to induration and fibrosis of the skin (lipodermatosclerosis) and ulceration (Fig. 45-6) remains unclear (118).

**Treatment**

Because venous hypertension in the upright position and during ambulation is the physiologic cause of the damage in chronic venous insufficiency, the first step of treatment should be to reduce the ambulatory venous pressure. Compression therapy is the mainstay of treatment for chronic venous insufficiency.

**Rehabilitation**

**Compression.** To prevent postthrombotic syndrome, compression stockings should be used routinely following diagnosis of a proximal DVT and continued for a minimum of 1 year (108). Compressive dressings aid venous return by compressing the leg and increasing the interstitial tension. Compression of dilated, engorged, superficial, and intramuscular veins indirectly increases the efficiency of the calf pump mechanism (115). With no history of congestive heart failure and no evidence of venous obstruction (on noninvasive studies), lower extremity volume can be stabilized with an intermittent pneumatic compression pump (40 to 50 mm Hg). Compression wraps should be used between pumping sessions. After the volume stabilizes, the patient can be measured for stockings. Elastic stockings should provide graduated compression, exerting most pressure at the ankle, less at the
calf, and least pressure at the thigh. Typically, knee-length graduated compression stockings with a pressure of 30 to 40 mm Hg (at the ankles) are prescribed. When elevation is used for edema control, the extremity is typically elevated above the level of the heart. The patient should lie on a sofa or sit in a recliner to elevate the legs appropriately. The correct duration and frequency of leg elevation should be tailored to the severity of disease. The leg should be elevated whenever possible and long periods of standing or sitting should be avoided. In patients with concomitant arterial disease, lower compression and modified elevation is used to avoid further compromise of arterial inflow.

**Exercise.** The value of exercise in the management of chronic venous insufficiency has not been conclusively demonstrated. Exercises involving the leg musculature, such as walking, bicycling, or swimming, promote muscle tone in the calf and enhance venous return. Exercise, however, produces variable reductions in venous hypertension. Patients with chronic venous stasis due to incompetent deep vein valves generally do not obtain as much reduction in venous pressure as do those in which the primary defect is due to incompetent perforator valves (119).

**Sclerotherapy.** Sclerotherapy (which uses the injection of an irritant substance into the vein) can be used to obliterate incompetent veins in certain settings (114,120). Sclerotherapy is used for the treatment of spider veins, small distal varicose veins, and venous insufficiency resulting from superficial and perforator incompetence. Ultrasound-guided sclerotherapy has been used successfully to treat incompetent perforator veins (121). Complications are rare, usually minor and include skin discoloration, thrombophlebitis, and hematoma formation. The temporary use of elastic compression is suggested after treatment.

**Surgery**

Ablation of superficial reflux is performed by stripping, radiofrequency or endovenous laser therapy of the greater saphenous vein, and avulsion of varicose veins. Subfascial endoscopic perforator vein surgery (SEPS) is included for patients with documented perforator incompetence and severe symptoms of venous hypertension. The stripping of superficial veins and SEPS should be avoided if extensive deep venous obstruction is present because the superficial veins may be acting as collaterals.

**Subfascial endoscopic perforator vein surgery.** SEPS is a minimally invasive technique performed in patients with advanced chronic venous insufficiency. The objective of the operation is to interrupt incompetent medial calf perforating veins to decrease venous reflux and reduce ambulatory venous hypertension in critical areas above the ankle where venous ulcers most frequently develop. Patients with stasis skin changes and healed or active venous ulcerations are potential candidates for the operation (122). Patients with ulcers caused by a primary valvular incompetence of the superficial and perforating veins, with or without deep vein incompetence are most likely to benefit from SEPS. These patients derive maximum benefit in terms of accelerated ulcer healing and 80% to 90% freedom from ulcer recurrence in the long term (122). Despite improved symptoms and objective clinical scores, only 50% of patients with postthrombotic syndrome will remain free of ulcer recurrence in the long term and the role of SEPS continues to be controversial in these patients. Patients with ulcer recurrence after SEPS should undergo duplex scanning to exclude recurrent or persistent perforators (122).

**Deep vein reconstruction.** Surgical bypass should be considered in patients with iliac or iliofemoral venous occlusion. Deep vein valve repair or valve transplantation can be attempted in patients with valvular incompetence and venous ulcers. Stenting of the iliofemoral veins and inferior vena cava has emerged as the preferred treatment of venous outflow obstruction. Severe (>50%) in-stent recurrent stenosis of iliofemoral venous stents is uncommon over the short term (123).

**May-Thurner Syndrome**

May-Thurner syndrome is defined as isolated left lower extremity swelling due to compression of the left iliac vein by the right common iliac artery. Treatment of May-Thurner syndrome has historically involved anticoagulation therapy. Advances in interventional management have allowed relief of the associated mechanical compression by open surgical or endovascular repair. Endovascular stents are the most common treatment.

**LYMPHATIC DISEASE**

**The Lymphatic System**

The management of lymphedema is a challenge for physicians and therapists. This component of the vascular system is often neglected in medical school curricula, yet malfunction can cause significant physiologic and functional impairment. It is important for the practitioner to have an in-depth understanding of the physiology of the lymphatic system and the pathophysiologic mechanisms of impairment to optimize management.

The lymphatic system is a vascular system composed of endothelial-lined channels that parallel the arterial and venous systems. The lymphatics originate in the tissue interstitium as specialized capillaries. These capillaries are porous and readily permit the entry of even large macromolecules such as albumin (124). The distribution of fluid between the peripheral vascular system and the tissues depends on the transcapillary balance between hydrostatic and protein osmotic pressure gradients. Normally, there is a slight hydrodynamic imbalance favoring a small amount of excess capillary filtrate (fluid, salt, and macromolecules) into the tissue spaces. This filtrate or lymph is collected by the lymphatics and returned to the venous circulation. A primary function of the lymphatic system is to return not only fluid, but also high-molecular-weight substances such as protein and particulate matter that are unable to reenter the venules or capillaries. The lymphatic system serves as a buffer system to decrease edema during fluid overload. As interstitial
fluid volume increases, interstitial fluid pressure increases. This results in a marked increase in local lymph flow (125).

**Chronic Lymphedema**

The volume of lymph formed depends largely on the balance between intravascular and extravascular hydrostatic forces, osmotic forces, and the permeability of the filtering capillaries (126). Because peripheral capillaries are not readily permeable to protein, a relatively steep protein osmotic gradient between plasma and interstitium is maintained. The amount of fluid not reabsorbed constitutes the lymphatic load.

When lymphatic blockage occurs (e.g., after lymphadenectomy or radiation-induced lymph node fibrosis), intra-lymphatic pressure distal to the site of the blockage increases. As lymphatic vessels dilate, their valves become incompetent. Increased intralymphatic pressure also decreases lymph formation and increases tissue fluid volume.

Lymphedema may be categorized as either high–lymph-output failure or low–lymph-output failure. High–lymph-output failure results from an overproduction of capillary filtrate (congestive heart failure, ascites, nephrotic syndrome) and leads to a greatly expanded extracellular fluid space. Low–lymph-output failure (decreased lymph absorption) occurs with deficient or obliterated lymphatics.

There is a strong association between the high protein content of lymph and the proliferation of fibroconnective tissue. It has been postulated that fibrin or other specialized protein complexes dispersed throughout the interstitial matrix form an intricate lattice template facilitating the deposition of collagen (127). Fibrosis may also result from the inability of local macrophages to digest the excessive protein load (128). The accumulation of protein promotes chronic inflammation and scar formation even in conditions of adequate lymphatic drainage and normal capillary permeability (129). Although the exact mechanism of this scarring is unknown, interstitial fibrosis results in the brawny, nonpitting soft tissue swelling seen in chronic lymphedema.

**Classification of Lymphedema**

Clinical lymphedema is classified as either primary or secondary (Table 45-3). Primary lymphedema occurs with aplasia, hypoplasia, or abnormal development of the lymphatic system; in addition, primary fibrosis of the lymphatics during puberty or at a later age also results in primary lymphedema. Congenital lymphedema is usually observed at birth with asymmetric presentation in a lower extremity. Several forms of heritable primary lymphedema are presently recognized (130). Milroy’s disease is a familial form of lymphedema present at birth (131,132), which has been found to be autosomal dominant (131). Lymphedema praecox, typically presenting at the time of puberty, can occur anytime from the early teenage years into the third decade. Lymphedema tarda is defined as primary edema occurring after the mid-30s (133). Secondary lymphedema occurs much more commonly. Disrupted lymphatic flow may occur because of infection, trauma, tumor obstruction, surgery, or radiation. Secondary lymphedema from filariasis, usually affecting the lower extremities, is the most common cause of lymphedema worldwide (134). Post-surgical compost radiation upper extremity lymphedema is the most common, with an incidence of approximately 25% (135).

Lymphedema may also be classified into clinical stages. Grade I edema easily pits in response to pressure and is reduced in response to elevation. There is no evidence of thickened spongy fibrosis on examination. Grade II edema does not easily pit with pressure and does not reduce with elevation. Some degree of fibrosis may be present. Grade III edema is irreversible with fibrosis and sclerosis of the skin and subcutaneous tissues. This severe, organized tissue change is often not responsive to mechanical treatment measures (136). Secondary lymphedema commonly presents within the first year after surgery, but can also present many years later (135). In recent years there has been recognition of a fourth grade of lymphedema termed “Grade 0.” This grade refers to a latent or subclinical condition where swelling is not evident despite impaired lymph transport. It may exist months or years before overt edema occurs (137).

**Clinical Presentation**

The differential diagnosis for new-onset unilateral limb lymphedema is important. Acute DVT must be ruled out whether the edema is presenting in the upper or lower extremity (see “Diagnostic Tests”). Other possibilities for diagnosis include postphlebitic syndrome, chronic venous insufficiency, tumor obstruction, chronic infection, and lipedema. Edema secondary to deep thrombophlebitis usually develops suddenly. Chronic venous insufficiency is generally associated with slowly progressing edema. Lymphedema may develop with either of these presentations. Systemic reasons for edema accumulation such as congestive heart failure, liver or renal disease, and fluid-retaining medications such as anti-inflammatory drugs and some antihypertensives should be considered.

**TABLE 45.3 Classification of Lymphedema**

<table>
<thead>
<tr>
<th>I. Primary lymphedema</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Congenital</td>
</tr>
<tr>
<td>1. Nonfamilial</td>
</tr>
<tr>
<td>2. Familial (Milroy’s disease)</td>
</tr>
<tr>
<td>B. Praecox (adolescence to about 35 y of age)</td>
</tr>
<tr>
<td>1. Nonfamilial</td>
</tr>
<tr>
<td>2. Familial (Meige’s disease)</td>
</tr>
<tr>
<td>C. Tarda (onset after 35 y of age)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>II. Secondary lymphedema</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Filarisis</td>
</tr>
<tr>
<td>B. Lymph node excision &amp; radiation</td>
</tr>
<tr>
<td>C. Tumor invasion</td>
</tr>
<tr>
<td>D. Other infection</td>
</tr>
<tr>
<td>E. Trauma</td>
</tr>
<tr>
<td>F. Other</td>
</tr>
</tbody>
</table>
The patient with lymphedema usually presents with asymmetric nonpainful swelling of one or both limbs. The edema often is pitting in its early stages, reduces overnight, and can exhibit a very light pinkish change in color. Chronic lymphedema does not resolve overnight and does not pit. Unlike venous edema, there is often involvement of the foot and toes with loss of dorsal vascular and joint markings. Long-standing edema leads to lymphatic stasis with inflammation and resultant hyperkeratosis with verruca-like appearance and eventual development of papillomas (138). Because the vascularity and elasticity of the skin tissues are retained in early lymphedema, the ulcerations observed in venous disease are usually absent (139). In the postsurgical breast cancer patients, presentation may include edema of the breast, chest wall, and axillary area, as well as, isolated proximal or distal extremity edema.

**Treatment of Lymphedema**

Lymphedema treatment is directed at minimizing the swelling and restoring normal function and avoidance of infection of the affected region (Fig. 45-7).

**Medical Diuretic Agents**

Often, a patient is offered a diuretic for lymphatic edema. Short-term diuretic use, during hospitalization for acute edema reduction, accompanied by elevation and compression bandaging may be useful. Diuretics are usually not helpful in the chronic management of lymphedema because the effect is temporary and the diuresis leaves behind large protein molecules, creating a concentrated state in the interstitium (140).

**Antibiotic Therapy**

Antibiotic therapy is indicated in the management of infection associated with lymphedema (focal cellulitis). No controlled studies are available for a specific regimen, for lymphedema-related cellulitis; however, a program of 1 week per month prophylaxis of Penn VK 250 or 500 mg, four times daily, has been advocated by our group for those patients with several episodes of cellulitis per year (141). For those with a penicillin allergy, a first-generation cephalosporin, clindamycin, or erythromycin may be considered. Inflammatory destruction of remaining lymphatics secondary to repeated infections can worsen the clinical problem. Cellulitis requires prompt management with a 10-day course of oral antibiotics. If symptoms become systemic, intravenous antibiotics may become necessary. Recalcitrant infections may require a program of daily prophylaxis. Antibiotic prophylaxis should be considered for those patients with recurrent cellulitis of the affected limbs. All prophylactic programs must be accompanied by rigorous edema control measures. Patients who have a history of intertriginous

![FIGURE 45-7.](image)

Left upper extremity lymphedema. **A:** Before treatment. **B:** After 2 weeks of specialized lymphatic massage and multilayered low-stretch compression bandaging.
fungal infection or who may be immunosuppressed should use prophylactic topical antifungal foot care.

**Benzopyrones**

Benzopyrones are thought to enhance fluid return by macrophage stimulation with increased proteolysis in the subcutaneous tissues. These agents have been reported to assist in edema reduction in both upper and lower extremities (142). However, a double blind cross-over study did not show any clinical benefit of benzopyrones compared to placebo (143). Reported side effects of benzopyrones include abnormal liver enzyme levels. Benzopyrones, although available elsewhere, are not widely utilized and are not approved by the U.S. Food and Drug Administration (143).

**Rehabilitation**

Comprehensive treatment regimens for lymphedema are referred to by various names, but most often called complex decongestive therapy. These treatment programs have become the standard of care. They include four common components: (a) skin care management and treatment of infection; (b) specialized massage techniques to promote the movement of lymph; (c) compression of the lymphedematous regions; and (d) elevation and exercises to reduce swelling and supplement the massage. After an active reduction phase (Phase I), an ongoing maintenance phase (Phase II) includes daily use of compression garments (and often continued nocturnal compression) to maintain reduction. General contraindications to aggressive lymphedema management include acute infection, cardiopulmonary edema, and ongoing malignant diseases. Conditions such as pregnancy, recent abdominal surgery, radiation fibrosis, DVT, and aortic aneurysm may preclude some portions of the treatment. Palliative lymphedema management is appropriate.

**Skin Care**

A vulnerability to infection exists wherever tissue fluids are poorly drained. Individuals with chronic lymphedema are susceptible to cellulitis from a variety of pathogens (144). Factors such as weak antigenic stimulation of regional lymphocytes and reduced extravascular macrophage function have been suggested as etiological contributors (145). Meticulous skin hygiene is imperative. The limbs should be washed regularly with soap and water and, while still moist, lubricated with an alcohol-free emollient cream.

**Specialized Massage**

Specialized lymphatic massage techniques in the form of manual lymphatic drainage cause collateral lymphatics of both superficial and deep networks to become dilated and carry more lymph and protein. Massage also facilitates the movement of tissue fluid and protein into lymphatic capillaries and along collecting lymphatics (146–149).

When deep lymphatic channels are blocked, there is considerable backflow of lymph into the network of superficial or dermal lymphatics. Where this occurs, massage is performed in a manner that promotes movement of fluid through superficial lymphatic vessels to patent lymph channels.

**Compression**

Compression of a lymphedematous region causes increased total tissue pressure, decreases the hydrostatic pressure gradient from the blood to the tissues, and increases the hydrostatic pressure gradient from the tissues to the initial lymphatics. The pressure gradient along the lymphatic trunks is also increased. Compression of the affected regions is necessary to maintain reductions in edema during and after the treatment. Lymphedematous regions may be compressed with bandages (elastic or low stretch) or graduated compression garments.

*Compression bandages.* There are two basic types of compression bandages. Elastic high-stretch compression bandages have a high resting pressure (from elastic recoil) and low working pressure because they stretch in response to muscle contraction. Low-elastic or low-stretch compression bandages have a low resting pressure and a high working pressure, which increases the total tissue pressure when muscles contract. Lymphatic vessels are compressed between the muscles and the bandages enhancing transport. Proper compression bandaging includes greater pressure application at the distal end of the limb with gradually reduced pressure toward the proximal creating a pressure gradient. Bandages can be used with various types of padding beneath them to reshape the limb (Fig. 45-8).
Compression garments. Once reduced stable limb volume is achieved, graduated compression garments are necessary to prevent fluid reaccumulation. Like compression bandages, compression garments enhance the pumping action of lymphatics and veins, and decrease the hydrostatic pressure gradient from the blood to the tissues. Compression garments are available in various levels of compression. For lymphedema, pressures of 30 to 40 mm Hg often suffice. Recalcitrant edema may require 40 to 50 mm Hg of support. Appropriate compression and fit of the garment along with skill of donning and doffing are critical to a successful management program. The style of the garment, fabric resistance, donning techniques, and wearing schedule should be optimized for each patient. Ill-fitting garments can cause pain and discomfort, leading to noncompliance and fluid reaccumulation. Vendor/industry options include prefabricated and custom garments in various sizes, fabric weaves, and color options.

Elevation

Elevation of a lymphedematous limb can decrease the hydrostatic pressure gradient from the vasculature to the tissues and reduce the amount of fluid and protein moving out of the capillaries. Elevation can also increase the lymphatic flow by increasing the hydrostatic pressure gradient along lymphatic trunks. Patients are encouraged to elevate an involved limb periodically throughout the day. For some individuals, elevation of the affected extremity is the most effective reductive intervention.

Exercise

Muscle contraction increases the total tissue pressure. This results in the active movement of fluid into and along initial lymphatics. The intrinsic pumping action of collecting lymphatics will also be enhanced by compression of these vessels between contracting muscles and surrounding fibrous tissues (150,151). Exercises are critical for fluid mobilization and should be performed while wearing either appropriate compression bandages or compression garments. The external compression augments the increases in tissue pressure and lymph flow. Historically, there has been concern that exercise may exacerbate lymphedema. A recent randomized controlled trial examined exercise and lymphedema in breast cancer survivors and found that a 6-month intervention of resistance exercise did not increase the risk for, or exacerbate symptoms of, lymphedema (152).

Vasopneumatic Compression Therapy

Vasopneumatic compression pumps increase the total tissue pressure in edematous limbs and push tissue fluid back into blood capillaries. Because excess protein is not removed from the tissues (149,153), the concentration of tissue protein may increase, and lead to the reoccurrence of edema. Another limitation of compression pumps in the management of lymphedema is that they cannot apply pressure to areas of the trunk adjacent to draining lymph nodes (154). Pumps may also exacerbate genital edema or produce a collection of high protein fluid proximal to the site of pumping, which may enhance inflammation and fibrosis. When using compression pumps for lymphedema management, pressures of 40 to 50 mm Hg are sufficient for fluid removal and to lessen the risk for tissue damage. It is the clinical observation of the authors that pneumatic compression is more helpful in the management of chronic venous insufficiency than in the management of chronic stage II or III lymphedema.

Surgery

Excisional Debulking

Excisional debulking procedures have occasionally been recommended for patients with significant functional impairment due to excessive lymphedema (133). This may decrease the volume of the affected extremity when irreversible changes in skin and subcutaneous tissues have occurred.

Reconstruction of Lymphatics

Surgical reconstruction of obstructed lymph vessels and lymph nodes has been performed using lymphovenous anastomoses (155–157) or lympholympathic anastomoses (lymphatic grafting) (158,159). Although clinical benefit has been reported by a few surgical groups, the long-term patency and function of such anastomoses is still uncertain. Currently, surgical treatment is reserved for those patients who have lymphangiectasia and primary chylosous disorders (lymphedema, chylosous effusions, and chylosus fistula). Reconstruction of dilated lymph vessels in these patients is done with saphenous vein grafts (160). At present, lymphatic reconstructions are not considered a first-line treatment for chronic obstructive lymphedema.

CONCLUSION

The patient with vascular disease poses a significant challenge to the rehabilitation professional. Arterial, venous, or lymphatic dysfunction may be the primary issue or a critical comorbidity in many patients who present to rehabilitation. These diseases may also be present in patients who present for presumably unrelated problems. A detailed vascular history, examination, and selected diagnostic tests should be inherent in the rehabilitation evaluation. When identified early by the rehabilitation professional, interventions including exercise, appropriate compression, modalities, positioning, protection, and proper footwear may ameliorate the need for more aggressive medical and surgical treatments in patients with vascular disease.

REFERENCES

PART III | MAJOR CONDITIONS


Transplantation Medicine: A Rehabilitation Perspective

INTRODUCTION

History
The modern field of transplantation traces its origins to a successful renal auto-transplant in a dog performed in 1902. The development of vascular anastomosis, which underlies organ transplantation, was developed by Carrel (1,2) who is known as the father of transplantation although he did not pursue this as a career. For this development, he was awarded the Nobel Prize in Physiology. The recognition that allograft rejection is an immunologic process was discovered by Medawar (3). This classic series of studies foreshadowed the future by including descriptions of tolerance, a goal of transplantation (3–5). Progress in the field was fostered by the development of an understanding of the major histocompatibility complex (MHC), which in humans is referred to as the HLA, and the recognition that the HLA is the immunological target of rejection (6). Developments in the understanding of cellular immunity and pharmacologic manipulation has allowed for the growth of the clinical field of transplantation medicine. Transplantation medicine is essentially only 30 years old, with the first successful trial in kidney transplantation in the 1970s and liver transplantation in 1980 by Starzl (7,8). These developments where supported by the development of immunosuppression agents such as cyclosporine and tacrolimus (FK-506) (7,9). In fact, small bowel transplantation has only been recently approved by the FDA as an acceptable treatment for short-gut syndrome. Nonetheless, the field has advanced to such an extent that transplant surgery is almost considered routine.

The Transplantation Process
The reception of a transplant allograft in many ways represents a complete medical renaissance of an individual. The transplant patient is likely an extremely deconditioned, malnourished, cachectic individual with one, if not more, end-stage organ disease, whether it is cirrhosis, lung disease, and/or cardiac disease. Having withstood the often long and unavoidably bloody operation with its attendant risks (i.e., prolonged anesthesia, prolonged paralysis, single positioning), these patients are further deconditioned and commit themselves to a lifetime of immunosuppression with their associated side effects, for example, infection, rejection, steroid-related problems. Such persons often have associated musculoskeletal and neuromotor impairments making rehabilitation an important part of their care.

Despite these seemingly insurmountable obstacles, the transplant patient is often able to return to be a highly functional, productive member of society. Transplant patients are running marathons, having babies, and returning to work as a state governor.

ISSUES RELATED TO TRANSPLANTATION

Of importance to the rehabilitation specialist is a series of medical issues that impact the outcome in persons with transplant.

Immunosuppression
Immunosuppression is a double-edged sword. While rescuing the transplant patient from the wave of rejection, these medications, in addition to exposing the patient to infection, have significant direct toxicities (Table 46-1). Most transplant centers that choose to treat long-term transplant survivors rely on the three-drug cocktail: cyclosporine, azathioprine, and prednisone. At many centers cyclosporine has been replaced with tacrolimus, and the use of prednisone has declined in most patients (1,2). Of interest, many patients at such centers require only monotherapy with tacrolimus and theories regarding the taper of immunosuppressive medications are being espoused (10). Other medications, including rapamycin (11) and mycophenolate mofetyl, are also utilized and may be seen in transplant patients undergoing rehabilitation. Whatever the “primary cocktail” combination, the crucial component to immunosuppression resides in the calcineurin inhibitors, cyclosporine, and tacrolimus. Cyclosporine and tacrolimus work by inhibiting calcineurin, a phosphatase that acts on nuclear regulatory factor, the end effect of which is the inhibition of interleukin-2 (IL-2), a cytokine critical in T-cell activation and thus rejection (11).

Cyclosporine
Cyclosporine, in combination with prednisone and azathioprine, in the early 1980s allowed for the high rate of graft acceptance seen today, and thus it is the mainstay of
immunosuppression in many transplant centers. Cyclosporine is a unique agent that upon introduction nearly doubled the 1-year graft survival among those receiving kidney transplants (12). Working via the inhibition of calcineurin, cyclosporine acts to freeze antigen-activated lymphocytes at an early phase and potentially allows for a deescalation of the cycle of response (11,12). Both the drug and its metabolites are eliminated via the liver as a primary step and the kidney as a secondary clearance step (13).

**Concerns**

Several medications interact with cyclosporine, especially those that act at the cytochrome P450 system. Anticonvulsant therapy leads to decreased levels, while several antibiotics and ketoconazole markedly increase levels of cyclosporine. Of importance is the fact that St. John’s Wort interacts with cyclosporine to cause a decrease in cyclosporine blood levels. In several cases this has lead to transplant rejection.

The main toxicity of cyclosporine is nephrotoxicity, occurring in approximately 18% of patients (14,15). Cyclosporine appears to inhibit renal function via multiple mechanisms, including vasoconstriction via endothelins and thromboxane stimulation, direct toxicity of vascular endothelium causing acute microvascular disease and inducing a chronic interstitial fibrosis (15,16). The two former mechanisms may be reversible, with reduction or cessation of cyclosporine, whereas the latter may progress to renal failure or, fortunately in many cases, stabilize. Loss of renal function remains significant with 9.5% requiring dialysis at 13 years post liver transplant (15,16). There is great variation in patient susceptibility to cyclosporine-based nephrotoxicity, although those with underlying renal impairment appear predisposed to renal failure from cyclosporine. Urine output and creatinine remains the first line of nephrotoxicity detection. Elevation of creatinine should prompt a nephrology or transplantation consultation with a plan to lower cyclosporine if possible and if rejection is not a problem.

Neurotoxicity remains the other major side effect of cyclosporine, ranging from peripheral neuropathy, tremors, and dysethesias, which are very common, to less commonly, seizures and full-blown encephalopathy (17–20). The fine motor tremor associated with cyclosporine therapy can often be diminished by adjustment of the medication and/or accommodation by the patient. Headaches secondary to cyclosporine therapy have been reported, and a demyelinating neuropathy can also occur (21).

Hypertension is associated with the use of cyclosporine therapy. It appears to be secondary to sodium and water retention, and as well as increased intracellular calcium (22,23). Other complications include electrolyte abnormalities such as mild hyperkalemia, which sometimes requires treatment (24,25). Lipid profile dysfunction is also noted with cyclosporine therapy. Hypomagnesemia is also common and may need oral replacement as it predisposes the transplant patient to seizures. Hypertrichosis and gingival hyperplasia are not medically severe, but can be cosmetically troublesome (26).

**Tacrolimus**

Tacrolimus (FK-506) was introduced in the 1990s and is considered more potent than cyclosporine. This agent has been reported to be as much as 100 times more potent than cyclosporine (27). Tacrolimus is a metabolite of the fungus Streptomyces tsukubaensis. It is a macrolide, highly plasma bound, and processed in a variable fashion via the liver in the cytochrome
Tacrolimus immunosuppressive action is similar to that of cyclosporine, working via the calcineurin system. Tacrolimus has allowed for steroid-free therapy, and has permitted small bowel transplantation to be more acceptable (30). Several large trials have examined the efficacy of tacrolimus in preventing rejection after liver transplantation (31). The incidence of acute and chronic rejection appeared to be lowered by the use of tacrolimus. There also appeared to be more retransplantations required in the non–tacrolimus–treated group (32,33). In a study of over 1,000 liver transplant patients by Jain et al., they concluded that chronic rejection occurred rarely among patients maintained long term on tacrolimus–based immunosuppressive therapy (33). In addition, the Multicenter Tacrolimus Rescue Trial has shown that among transplant patients receiving cyclosporine who are experiencing chronic rejection, many can be salvaged by transfer to tacrolimus (34). An additional potential benefit of tacrolimus is that it has been shown to demonstrate a reduction in serum cholesterol when compared to cyclosporine (35,36).

Concerns
Tacrolimus, being also a calcineurin inhibitor, has a toxic profile remarkably similar to that of cyclosporine, although neurotoxicity is more prevalent (37,38). Susceptibility to neurotoxicity may have a genetic relationship. In a study by Yamauchi (39), the genotype of six patients who had experienced neurotoxicity after liver transplantation was examined. The authors found a relationship between neurotoxicity associated with polymorphism in the ABCB1 gene (39). Resting tremor is the most common neurologic adverse reaction although more serious effects such as peripheral neuropathy and encephalopathy have been reported. Debate exists as to how often this profile of side effects occurs among those taking tacrolimus when compared to the population utilizing cyclosporine for immunosuppression (37). Shimono (40) utilized diffusion weighted MRI to evaluate those that developed neurotoxicity following organ transplantation. He noted that 35.7% had white matter abnormalities, 7.1% had putaminal hemorrhage, and 57.1% had normal findings on MRI (40). Barynski (41) evaluated 22 patients with a neurotoxic reaction. He also noted the significance of white matter lesions in this population. Patients often have tremors initially, although personal observation suggests that there may be an acclimation to the drug. For persistent tremor that affects activities of daily living, low-dose clonazepam may be helpful.

Drug interactions are common, especially when using itraconazole or ketoconazole since hepatic metabolism is inhibited, risking toxicity. Because of profound hypomagnesemia, frequent monitoring of magnesium levels should be done and often both oral and intravenous (IV) supplementation is required.

Sirolimus
Sirolimus (Rapamycin) is an immunosuppressive that inhibits the cell cycle and the activation of IL-2 (42,43). This agent is produced by Streptomyces hygroscopicus. Sirolimus is still finding its place in immunosuppression as either a primary drug or adjunct with tacrolimus or cyclosporine. Its mechanism of action is different from either cyclosporine or tacrolimus in that it does not inhibit calcineurin (44). Sirolimus appears to prevent the translation of mRNA impacting cell cycle regulation (45–47). Sirolimus, being nonnephrotoxic, is a viable alternative in patients who develop renal insufficiency caused by calcineurin inhibitors (42,48–50). In a retrospective review, patients who were more than 3 years posttransplantation were selected to evaluate the role of sirolimus. Patients who had proteinuria, those administered any other nephrotoxic agents, and those with a creatinine clearance less than 20 mL/min were excluded. Renal insufficiency was defined as mild, moderate, or severe. In the 16 patients studied, there was significant improvement in serum blood urea nitrogen and serum creatinine levels 6 months after switching to sirolimus therapy. No patient developed cellular rejection or other graft-related complications (51). Among liver transplant recipients with chronic renal insufficiency, conversion to sirolimus–based immunosuppression may allow complete withdrawal of other agents, leading to some improvement in renal function. Another recent study evaluated the safety and efficacy of sirolimus plus steroids as a maintenance regimen with or without small-dose cyclosporine adjunctive therapy in renal transplantation (52). A total of 133 recipients of kidney allograft transplantations recruited in the United Kingdom and Ireland were enrolled into the study. Patient and graft survival were 97.7% and 95.5%, respectively, whereas the biopsy proven acute rejection rate in the first 6 months was 19.5%; incidents of acute rejection rates comprised 22 episodes (16.5%) during the first 3 months of the study and four episodes (3%) after randomization. These data demonstrate that withdrawal of cyclosporine from a small-dose sirolimus maintenance regimen is safe and is associated with an improvement in renal function (52). The study also suggests that the addition of small-dose cyclosporine to a sirolimus maintenance regimen does not increase the immunosuppressive efficacy. A recent study of 26 subjects receiving liver transplantation who experience nephrotoxicity owing to calcineurin inhibitors, suggests successful transfer to sirolimus monotherapy is possible and results in improved renal function (53).

Concerns
Sirolimus causes neutropenia and hyperlipidemia and may predispose patients to more infection. Its antiproliferative properties can also delay healing (54). It is also noted to cause thrombocytopenia, and has also been suggested to have prothrombotic activity. Sirolimus increases the levels of cyclosporine, when used in combination. Even though sirolimus is not primarily nephrotoxic, dual treatment groups have noted an increased incidence of nephrotoxicity (55). Combination sirolimus–tacro-rolimus may cause nephrotoxicity in some patients by mechanisms that are presently unexplained. Sirolimus appears to prolong delayed graft function (DGF), therefore, it may not be the optimal immunosuppressive choice in the DGF setting (56). A report of delayed wound healing in a person treated
with sirolimus who received a liver transplant has been noted, and late wound dehiscence can occur as well (57).

Corticosteroids

Corticosteroids have been used in transplantation medicine since the early days of solid organ transplantation. These agents impact the functional capacity and the concentration of active leukocytes (58). This tends to occur via the regulation of cytokine gene transcription. These agents have value as adjuvant therapy, at times of stress, and in the treatment of mild-moderate acute rejection. Several attempts have been made at discontinuing steroid therapy. The most successful of these have been programs with tacrolimus-based paradigms (59,60).

Corticosteroids have long been a cornerstone of orthotopic liver transplant immunosuppression. Newer, more potent agents have successfully allowed for more rapid tapering and discontinuation of corticosteroids in liver transplant recipients. Washburn (60) hypothesized that corticosteroids can be safely avoided after the first postoperative day using these newer agents. Thirty adult orthotopic liver transplant recipients were prospectively enrolled in a randomized open-label protocol. The incidence of biopsy-proven acute rejection requiring steroid therapy was 6.7% in both the steroid and the "no steroid" groups. Serum cholesterol levels were significantly lower in the "no steroids group" at 6 months after transplantation. Serum triglycerides were also lower, but the difference was not significant. Boots (61) evaluated the role of steroids in renal transplantation, by comparing tapering in 3 to 6 months, with stopping steroids 1 week posttransplantation. Results were noted to be comparable in patients and graft survival, with a similar incidence of acute rejections. The incidence of new-onset diabetes may be reduced among those receiving early taper from steroids (61). The immunosuppressive benefit of adding enteral prednisone to tacrolimus seems to be limited. In a prospective, randomized, double blind, placebo controlled, multicenter study of early steroid withdrawal versus chronic steroid therapy for those with renal transplantation, Woodle et al. (62) noted that early steroid withdrawal is associated with an increase in biopsy-associated rejection (usually mild) yet results in similar allograft survival and function.

The side effects of long-term steroid use are well described and are significant in transplant patients. These include poor wound healing, glucose intolerance, osteoporosis, anasarca, muscle wasting and myopathy, and emotional liability. Lipid abnormalities have also been described. Genotypic testing noted that the presence of an Apolipoprotein E4 allele worsened high-density lipoproteins (HDL), triglycerides, and cholesterol (63). Thus, Apolipoprotein E4 has a larger impact than Apolipoprotein E2 on fasting-lipid profile in transplant candidates. There have been occasional reports of acute respiratory and skeletal muscle weakness in intensive care unit patients treated with massive doses of corticosteroids for rejection prophylaxis or treatment. Compared to the pretreatment condition, approximately 45% of patients showed acute generalized muscle weakness that recovered after approximately 2 months (64). If given a "risk-free" choice, the majority of recipients prefer withdrawal of steroids (65).

Alternative Agents

Azathioprine and mycophenolic acid are adjunctive immunosuppressive agents that inhibit purine synthesis; thereby generally inhibiting immune cells, particularly lymphocytes. These agents have been noted to block de novo purine synthesis by blocking key enzymes. This process is required by T and B lymphocytes for proliferation (66). The role of mycophenolic acid as adjunct therapy in preventing rejection has been established among those with renal and heart transplantation (67). Mycophenolate combined with cyclosporine and prednisone significantly lowers acute rejection frequency in the early post-renal transplantation phase (67). A comparative retrospective analysis of the 5-year results with mycophenolate was performed by Offerman (69). Both the total population and subgroups showed a nonsignificant trend toward better graft survival with mycophenolate, evident at 2 years and persisting for 5 years. Extrapolation indicates that combination therapy with mycophenolate versus azathioprine, results in approximately 10% more patients being alive at 10 years with a functional graft. These agents are also employed in the treatment of acute rejection in such populations (68). Side effects include diarrhea that can mimic Clostridium difficile infection, neutropenia, and pancytopenia. Thymoglobulin was effective as induction therapy in high-risk pancreatic transplant recipients, and resulted in initial reversal of rejection in 74% of patients. Dose adjustments were required in over half the cases and were usually due to leukopenia. Infections occurring subsequent to thymoglobulin were not uncommon and reflected the immunosuppressive burden of the patient population (70).

Monoclonal Antibody Agents

A major thrust of transplantation research is to find more effective and less broadly toxic immunosuppressive agents. One potential way is the use of monoclonal antibodies directed to IL-2 receptors. Several monoclonal antibodies have been introduced that target the IL-2 receptor. Daclizumab and Basiliximab focus on the α-2 chains of the IL-2 receptor (71,72). Pediatric patients appear to be among those that benefit the most from anti–IL-2 receptor therapy (73). Immunosuprophylaxis with daclizumab has been shown to be effective in the prevention of acute rejection in kidney transplant patients. Niemeyer initiated a pilot study in 28 liver transplant patients. Daclizumab was administered intravenously. At 4 years post-transplant, no lymphoproliferative disease was observed (74). Immunosuprophylaxis with a two-dose daclizumab regimen appears safe, effective and well tolerated, and does not lead to increased opportunistic infections. Such agents are utilized in combination with the agents previously described (74–77). A two-dose regimen appears to be as effective as the five-dose regimen in preventing acute rejection and is associated with the lowest acute rejection rates and the highest rate of event-free survival (no rejection or graft loss). However, the benefits
of daclizumab compared with no antibody induction await larger sample size accrual (78).

**Gene Therapy**

One major complication facing organ transplant recipients is the requirement for life-long systemic immunosuppression to prevent rejection, which is associated with an increased incidence of malignancy and susceptibility to opportunistic infections. Presently, exciting investigations into the role of gene therapy for immunosuppression are underway. A further understanding of genetic polymorphisms may also assist in defining those who are at risk for complications such as graft versus host disease (79). Gene therapy has the potential to eliminate problems associated with immunosuppression by allowing the production of immunomodulatory proteins in the donor grafts, resulting in local rather than systemic immunosuppression (80,81). The use of gene therapy may also make xenografts more practical. Alternatively, gene therapy approaches could eliminate the requirement for general immunosuppression by allowing the induction of donor-specific tolerance. Gene therapy interventions may also be able to prevent graft damage owing to non–immune-mediated graft loss or injury and prevent chronic rejection (80,82; Fig. 46-1).

**INFECTION**

Immunosuppression places these patients at extreme risk for opportunistic infections, which has probably remained the most significant cause of morbidity and mortality in the transplant recipient. Since the immunosuppressive regimen includes medications that suppress both T and B cell action, these patients are vulnerable to many types of infections: bacterial, viral, fungal, tuberculosis, and pneumocystis. Several of these infectious agents are not typically seen in immunocompetent individuals. Despite the fact that aggressive pretransplantation infectious evaluation is part of every protocol, infections are common. Moreover, immunosuppressives, especially prednisone, may mask infectious etiologies making the diagnosis of infection difficult in transplant patients (83). Fever and leukocytosis are still the most common signs of infection, although their absence does not rule it out. Systemic infection may present with a low-grade temperature or no fever at all. It is not uncommon for a transplant patient to have two simultaneous opportunistic infections. Again because of the immunosuppression, once a patient begins to show signs of infection, the course to fulminate sepsis and potentially death can be brief and dramatic (84–86). Vigilance and a high index of suspicion is the rule in detecting infections in transplant patients. During the past decade, ever-increasing numbers of patients have undergone renal, pancreatic, small bowel, hepatic, cardiac, or lung transplantation. Significant improvements in patient and allograft survival have been observed in all categories. Unfortunately, despite such improved results, the risks of infection related to immunosuppression continue to be substantial. Dunn performed a review of transplant-related nosocomial infections. These authors noted that suppression of host defenses by exogenous immunosuppressive agents renders patients susceptible to invasion by either resident or environmental bacterial, fungal, viral, and protozoal microbes or parasites (87). Invasion of organisms that typically produce mild infection in nonimmunosuppressed individuals can produce severe, lethal disease in those receiving immunosuppressive agents. Thus in an ideal world, immunosuppression could be decreased or eliminated in such patients.

Posttransplantation infections can be divided into three main etiological categories based on time after surgery: During the first 3 to 4 weeks infections are related to technical and mechanical problems (i.e., line infections, abscess, cholangitis from biliary stenosis, bowel obstruction, and wound infections). In a second phase, the first through sixth months posttransplant, cytomegalovirus (CMV) accounts for two thirds of infections, and beyond 6 months, infections are similar to that seen in a general population (88,89). The typical transplant patient receiving acute inpatient rehabilitation is in the first category, although now with long-term transplant survivors, physiatrists are also treating posttransplant patients with other medical problems, such as hip replacement in a patient with an allograft.

Along with factors related to immunosuppression, those receiving transplant may have prolonged hospitalization and thus secondary infection. Common pathogens include *Aspergillus*, *Staphylococcus*, *Clostridium difficile*, *Pseudomonas aeruginosa*, and *Legionella*. Viral infections are also quite possible and include: CMV, Epstein Barr virus, Herpes Simplex, Varicella Zoster, Hepatitis B and C (90).

Pulmonary infections are the most common cause of morbidity in the lung transplant population. Prompt recognition and treatment is necessary to prevent poor outcomes. An understanding of the temporal relationship between...
immunosuppression and the risk for developing infection can assist the clinician with appropriate treatment. Bacterial pneumonia is common within the first 4 months after transplantation whereas CMV infection becomes more prevalent after the discontinuation of prophylaxis in at-risk patients (91,92). Fungal infections, especially aspergillosis, can be fatal if not treated early and the risk for infection is present throughout the transplant period. Community-acquired viral infections present with upper respiratory symptoms and wheezing that may lead to a persistent decline in lung function. Suspicion of a pulmonary infection in these immunosuppressed individuals should lead to an urgent diagnostic bronchoscopy and empiric antimicrobial therapy (90).

CMV is the most common infection in transplant patients and deserves significant attention. CMV is usually found latent in either recipient or donor and becomes active in a milieu of immune inhibition (91,92). The most virulent CMV infection is seen when a CMV naïve recipient succumbs to CMV carried within the donor allograft. CMV can be widely disseminated and infection can occur anywhere in the body. Fever, leukopenia, and generalized malaise are the most common symptom; although, pneumonitis, hepatitis, cholecystitis, and colitis with occasional GI bleed are not uncommon (92). Posttransplant CMV infection has been associated with posttransplant arteriopathy and can signify invasive infection (93). In many transplant centers, a fever work-up includes CMV culture and antigen detection. In our center, all transplant patients receive prophylaxis with acyclovir (200 mg PO bid), and active treatment is initially with IV ganciclovir (5 mg/kg for 3 to 4 weeks followed by maintenance dose orally for a specified period of time, usually months) (94). Treatment is continued until the patient has three consecutive CMV antigen negative tests. CMV antigen detection is done weekly at many institutions for all patients within the first month of their operation.

Fungal infections are also more common in the immunosuppressed transplant patients than the general population. The clinician has to be vigilant because fungal infections can be devastating. Feltis (95) has noted mycotic aneurysms after transplantation. Fungal infections are associated with high morbidity and mortality, and make up a significant proportion of infectious complications. Unfortunately, the diagnosis is usually made late, and symptoms may be mild and nonspecific, even with dissemination. Mortality associated with disseminated fungal infections is high, while those associated with more local fungal infections is low (96). Although the risk factors for invasive fungal infections in liver transplant patients are well identified, early diagnosis is challenging, and commonly used diagnostic methods lack sensitivity and specificity (96). The incidence of fungal infections following liver transplantation appears to be falling. Future developments should focus on enhancing earlier diagnosis, implementing more effective and less toxic antifungal therapy recipients (97,98). Early fungal infections are related to surgical complications, while the period of 1 to 6 months after transplant reflects opportunistic, relapsed, or residual infections; fungal infections greater than 6 months after transplant and thereafter are usually associated with treatments for chronic rejection or bronchial airway mechanical abnormalities (99). The majority of fungal infections in lung transplant recipients involve Aspergillus species, followed by Candida, Pneumocystis, Cryptococcus, geographically restricted agents, and newly emerging fungal pathogens. Virulent infections such as Aspergillus or Mucormycosis are usually fatal. Aspergillus can be widely disseminated and may not grow in culture despite systemic infection (100).

Of note, persons receiving lung transplants for cystic fibrosis appear to be at particular risk for Aspergillus infection (100). The isolation of Aspergillus fumigatus from respiratory tract specimens in heart transplant recipients appears to be highly predictive of invasive aspergillosis (101). The presence of Aspergillus “fungus balls” lesions in the brain usually means death, and lung involvement requires resection for any chance of survival. Mucormycosis spreads by the nasal sinuses into the brain where their infection portends death; cure, if caught early enough, requires aggressive debridement and resection of facial tissues.

**REJECTION**

Rejection exists as a continuum of responses that are important to understand. Transplant rejection has been classified as hyperacute, acute, subacute, and chronic. Hyperacute rejection occurs when the recipient has antibodies to antigens present on the transplanted organs. This elicits an aggressive immune response resulting in organ failure. This response type has generally been limited by extensive routine cross matching. Acute rejection is a relatively intense inflammatory response to the new organ. The inflammation that occurs has been described as being similar to that seen in a hypersensitivity type of reaction (102). The result of this reaction is the deposition of a large number of inflammatory cells within the graft. The clinical syndrome manifests itself as rubor, edema, fever, and pain. The transplanted organ's function-related laboratories may also become acutely impaired. Despite potent immunosuppression, more chronic forms of rejection are still a common event usually within the first 3 months after transplantation. Symptoms such as fever and leukocytosis are sometimes present although rejection often occurs in the absence of these symptoms. Usually subacute and chronic rejection is heralded by a rise in an allograft laboratory value, such as increased creatinine in persons with renal transplant, increase in bilirubin or liver function enzymes in persons with liver transplantation, and elevation in amylase and lipase among those with pancreas transplants. Tissue biopsy, usually obtained percutaneously, with histologic examination remains the gold standard for the diagnosis of rejection. The incidence of chronic rejection varies by type of solid organ transplantation. The incidence of chronic rejection has been reported as high as 50% in the lung transplantation population and as low as 5% among those receiving liver transplantation (103,104).

Chronic allograft nephropathy (CAN) is, besides death of the recipient, the most common cause of renal transplant loss.
It is characterized by loss of function and replacement of tissue by fibrotic material (103). The pathogenesis is not clear, but seems to be multifactorial and involves events both early and late after transplantation.

A new and exciting concept is that of tolerogenic immunosuppression. Under this concept, individuals develop a partial tolerance to the transplanted organ and thus require less immunosuppression without associated rejection. This concept focuses on the principle of recipient pretreatment and minimum use of posttransplant immunosuppression. Starzl et al. (105) evaluated 82 persons awaiting various solid organ transplants. Posttransplant patients were maintained on sirolimus and subsequently tapered in dose after 4 months of therapy. Immunosuppression-related morbidity was virtually eliminated and transplant graft survival at 1 year was 89% (Table 46-2).

Once the diagnosis of acute rejection is confirmed, treatment is initially with high dose steroids (Solumedrol 1 gm) and a steroid taper that will successfully treat 80% to 85% of allograft rejection. Steroid-resistant rejection can be treated with antibodies, OKT3, or thymoglobulin, which will resolve 95% of all rejections. Resolution is characterized by the normalizing of laboratory values, so that follow-up biopsy is usually not necessary. Solumedrol administration and a steroid taper can be given in the rehabilitation unit without impacting the patient’s level of activity. Antibody treatment will require closer monitoring and is best done upon transfer to a specialized transplant service.

**MALIGNANCY**

Immunosuppressive therapy for organ transplant recipients is complicated by high rates of malignant diseases, one of these being Kaposi sarcoma (KS) (106). A review of 1075 patients who underwent kidney transplantation between 1985 and 2002 noted that a total of 52 malignant diseases were observed in 50 patients (4.7%), during the posttransplantation period (107). KS was identified in 16 of these 50 individuals. One recipient had concomitant lymphoma and KS. The lesions of seven patients were limited to the skin, five cases involved the skin and gastrointestinal tract, and four patients had disseminated disease (108). At the time of writing, nine individuals were still alive: four had normal renal function and five had lost their grafts as a result of chronic rejection. The combination of immunosuppressive drug withdrawal and chemotherapy appears to be very effective in patients with limited disease, but the results in cases of generalized disease have been poor (108). It is believed that the calcineurin inhibitors have an independent cancer promoting effect and only rapamycin has shown simultaneous immunosuppression and antitumor effect (109).

Posttransplant lymphoproliferative disease (PTLD) is a lymphoid neoplasm that has been associated with immunosuppressive therapy (110). The exact mechanism of this malignancy in the posttransplant population is not well understood, however, as predicted and subsequently verified in 1968, an increased incidence of certain de novo malignancies has been observed, particularly with regard to lymphoid neoplasms (111). Jain et al. (112) performed a retrospective review of 4,000 consecutive patients who underwent liver transplantation. The 1-year patient survival for liver transplant patients with PTLD was 85%. The actuarial 20-year survival was estimated at 45%. The overall median time to PTLD presentation was 10 months, and children had an incidence of PTLD that was threefold higher than adults (112). Patient survival was better in children, in patients transplanted in the era of tacrolimus immunosuppression, in patients with polymorphic PTLD, and in those with limited disease.

While the overall incidence of PTLD in pediatric liver transplant recipients has been reported to be 4% to 11%, the long-term risk of PTLD associated with primary tacrolimus therapy was determined by Cacciarelli et al. (113). These authors evaluated 131 pediatric recipients who underwent

<table>
<thead>
<tr>
<th>Type</th>
<th>Symptom</th>
<th>Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hyperacute</td>
<td>Occurs in OR during initial transplantation surgery.</td>
<td>Graft replacement/retransplantation</td>
</tr>
<tr>
<td>Acute</td>
<td>Secondary to preformed antibodies.</td>
<td>Solumedrol, if steroid resistant, thymoglobulin or OKT3</td>
</tr>
<tr>
<td>chronic</td>
<td>T-cell-mediated antigrant inflammatory infiltrate</td>
<td>Organ replacement is the treatment of choice.</td>
</tr>
<tr>
<td></td>
<td>↑ LFT in liver transplants.</td>
<td>Possibly ↑ FK-506 or give steroid bolus.</td>
</tr>
<tr>
<td></td>
<td>↑ Glucose in pancreas transplants.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>↑ Blood urea nitrogen (BUN) with creatinine in kidney transplants.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>↑ Heart/lung biopsy may reveal ↑ inflammation.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>↑ There may be ↑ inflammation.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>↑ Shortness of breath.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>↑ Edema</td>
<td></td>
</tr>
</tbody>
</table>
liver transplantation and received primary tacrolimus therapy. Pre-transplant Epstein-Barr virus (EBV) serologies were negative in 82%, positive in 12% of the patients (113). The cumulative long-term risk of PTLD was found to increase over time. Mortality from PTLD was 12% and a suggestion exists regarding a causal relationship with EBV status. EBV surveillance should occur in early posttransplantation (113).

Recurrent and de novo malignancies are the second leading causes of late death in liver transplant recipients, following age-related cardiovascular complications (106). The greatest incidence of de novo malignancies is seen in cancers associated with chronic viral infections, such as EBV-associated PTLD and skin cancers, including squamous cell carcinoma and KS (110).

**SOLID ORGAN TRANSPLANTATION**

**Liver Transplantation**

The achievement of liver transplantation by Thomas E. Starzl is one of the most remarkable technical surgical accomplishments and remains one of the most complicated operations in surgery. As such, it is metabolically demanding on the patient. Operating in patients with portal hypertension and all of its stigmata, including friable varices and hypocoagulability can be challenging. Because of the multiple vascular anastomoses (IVC, portal vein, hepatic artery), and the extensive dissection in the right upper quadrant, bleeding postoperatively can also be significant. Usually 2 to 4 Jackson-Pratt drains are placed to drain the suprahepatic space where bleeding from the IVC anastomosis can occur, to drain the subhepatic space where the infrahepatic IVC, portal, and hepatic artery anastomosis may bleed, and near the hilum to detect bile leaks from the biliary anastomosis (114). These drains are usually removed by 1 week, although some may stay longer if they are draining copious amounts of ascites postoperatively. T-tube stenting across the biliary anastomosis is also commonly done, and these are removed months after liver transplantation, but can be capped if the bilirubin has normalized to allow internal drainage of bile. Most patients in rehabilitation units will have T-tubes, but not drains.

Vital signs and two to three times weekly laboratory values for complete blood count (CBC), liver function tests (LFTs), and immunosuppressive agent levels are sufficient for routine screening. Fever, leukocytosis, or sudden leukopenia should prompt a fever work-up including CMV titer and culture (115). An elevated total bilirubin is the most sensitive indicator of rejection; however, increases of AST, ALT, and/or gGTP may signify rejection as well as infection (hepatitis), and vascular or biliary complications. Hepatic artery aneurysm and thrombosis are significant complications after transplantation (116,117). Persistence of ascites or the sudden development of ascites in a recent liver transplant patient should prompt an examination of the vasculature of the liver allografts to rule out outflow obstruction or portal vein thrombosis (117). An abdominal ultrasound to assess the hepatic vessels is the first step toward diagnosis. A T-tube cholangiogram should be obtained and if these studies are unrevealing, a biopsy to rule out rejection should be performed (118,119). These studies are usually completed with the patient returned to the primary care of the transplant team. It is obvious that close coordination between the physiatrist and transplant surgeon is critical to the optimal care of the patient, especially in these situations.

**Renal Transplantation**

Unlike liver transplantation, renal transplants are relatively routine and the operation is straightforward. The kidney allograft is placed in the iliac fossa in the retroperitoneal position. The renal artery and vein are anastomosed to the iliac artery and vein, respectively, and the ureter is implanted onto the bladder medially. Lasting only 2 to 4 hours in duration, renal transplantation is not associated with much blood loss. Drains are not usually placed, but in rare situations in which the renal bed is drained, they are removed within 3 to 5 days. Foley catheters are kept for 5 to 7 days to decompress the bladder and protect the ureteral anastomosis.

Graft function is assessed by urine output and creatinine levels. Urine output less than 50 cc/h should prompt an investigation into the cause, keeping in mind the most common cause being hypovolemia. A fluid challenge should be the first step. The goal for urine output in a patient with a renal transplant is approximately 2 L/d. A rise in creatinine signifies renal derangement that could be associated with hydrenephrosis, urinary tract infection (UTI), rejection, or immunosuppressive toxicity. Doppler ultrasound is the first test to obtain to rule out an anatomic abnormality such as hydrenephrosis that may be due to stricture of the ureterovesicular anastomosis. It can also rule out a perinephric collection that compresses the allograft and prevents perfusion (120,121). Finally, the ultrasound can examine the vessels to assess patency and flow, and to rule out thrombosis or stenosis (122,123). A simple urinalysis and culture will help to rule out a UTI. A history of trough levels of immunosuppression may suggest drug toxicity, although tacrolimus or cyclosporin nephropathy can occur with normal levels. A renal biopsy will rule out rejection or drug toxicity (120). First-line therapy for rejection is steroids and antibodies (OKT3, thymoglobulin), if the creatinine does not fall to baseline.

Finally, most renal transplant patients are diabetics or have hypertension, and are vulnerable to cardiac disease (124). Almost all renal patients undergo a cardiac work-up, including a dobutamine echo, but this work-up is not foolproof. Incidences of myocardial ischemia occur with intraoperative graft reperfusion when blood pressure drops, and at postoperative days 3 to 4 when the patient is mobilizing fluid from the interstitium back to the intravascular space (125). These problems are handled by standard medical guidelines.

**Pancreas Transplantation**

Pancreas transplantation is usually offered to type I diabetics with severe diabetic end-organ damage. End-stage renal failure,
severe gastroparesis, neuropathy, retinopathy, and especially hypoglycemic unawareness, are indications for pancreas transplantation. Patients are offered simultaneous kidney and pancreas transplants, pancreas transplant after having a kidney allograft, or pancreas transplant alone (126). The pancreas allograft is typically placed in the iliac fossa, more commonly on the right because of anatomical advantage and the exocrine drainage can be to the small bowel or to the bladder. Recently, there has been a general trend to more enteric drainage due to the complications associated with bladder drainage, including pain, UTI, and bicarbonate loss (127).

There are very strict criteria for the selection of a donor pancreas and recipient, simply because the pancreas is a fragile organ prone to many potential morbidities. For instance, donor or recipient age above 45 is associated with increased morbidity and mortality (127). Isolated pancreas graft survival, which has the worst outcome amongst liver, kidney and pancreas transplants, is only 75% at 1 year and 66% at 5 years, although there has been steady improvement in the last 10 years. Overall morbidity (fistulas, leaks, bleeds, pancreatitis, infection, necrosis, abscess) can be up to 80% (125,128,129). Unlike kidney or liver, there are only a small number of transplant programs in the United States that perform more than 30 pancreas transplants/year.

The postoperative management of a pancreas transplant patient can be very challenging for the surgeon and physiatrist alike. The usual course for the pancreas recipient can be “stormy” in the first month, during which time daily or every other day laboratory tests including electrolytes, CBC, amylase, and lipase should be taken. Fasting blood glucose should be obtained four times per day, and c-peptide levels weekly to assess the pancreas function. These patients can take a regular diet, and fasting glucose should be less than 150, and in many cases less than 100. Recent surgery as well as the heterotopic placement of a pancreas allograft makes these patients prone to ileus, or diarrhea, making nutrition potentially problematic in the early postoperative period.

Venous thrombosis, which occurs in 10% to 15% of all pancreas transplants, usually occurs during the first 2 weeks postoperation, but can occur anytime (124). A sudden and persistent rise in the blood glucose (>200) should prompt an immediate pancreas ultrasound and nuclear flow scan to rule out thrombosis, and immediate notification of the transplant service as this can be potentially dangerous.

Enteric leak at the anastomosis can also occur in 10% of all pancreatic transplants, and can be potentially fatal. Leaks usually occur within the first month of transplant. Fever, abdominal pain, signs of sepsis, and leukocytosis are the usual manifestations of this problem. A computed tomography (CT) scan without IV contrast is usually obtained to rule out leak, and notification of the transplant service should be done as this is a surgical emergency. The presence of steroids and immunosuppression, which can mask peritonitis, makes this diagnosis tricky.

Pancreatitis and its complications such as fistulas, pseudocyst, necrosis, can often occur anytime posttransplantation. All of these conditions can result in the elevation of amylase and lipase levels, which are also indicators of rejection. Again, ultrasound is the first test obtained to evaluate for fluid collections, and to examine vascular flow in the pancreas. Once a collection is ruled out, a pancreas biopsy should be performed, if technically possible, to delineate pancreatitis from graft rejection. Mild pancreatitis can be observed or treated with sandostatin. More severe pancreatitis, in which there is tenderness over the graft, is best treated by making the patient NPO, and administering total peripheral nutrition (TPN) and sandostatin.

Percutaneous biopsy is the only way to conclusively diagnose pancreas rejection, but this can be technically difficult to perform (130–132). Those patients with simultaneous pancreas and kidney transplants from the same donor have the advantage of the kidney as a sentinel organ for rejection. Kidneys are less technically demanding to biopsy, and the creatinine levels serve as another marker of rejection. An increase in amylase and lipase is the best marker of allograft rejection, although the differential diagnosis of hyperlipasemia is not small. Moreover, in small number of cases, there is pancreas rejection without an increase in lipase (133). The pancreas is not as resilient as other types of allografts, so rejection should be treated aggressively.

Mild rejection is usually treated with steroid therapy. Antibody treatment is used for steroid-resistant rejection or for moderate to severe pancreas rejection, depending on surgeon judgment.

**Small Bowel Transplantation**

Small bowel transplantation is reserved for patients with irreversible intestinal failure who no longer can be maintained on total parenteral nutrition. Small bowel transplant can be offered as an isolated transplant, in combination with liver allograft, or as part of a multiviscera transplant, which includes liver, pancreas, stomach, and small bowel. Short-gut syndrome (midgut volvulus, necrotizing enterocolitis), intestinal dysmotility, or enterocyte failure (radiation enteritis, autoimmune enteropathy) are some indications for small bowel transplants (134). Chronic TPN can cause liver disease secondary to cholestasis, adding the need for a liver transplant with the small bowel. Finally, extensive thrombosis of visceral vessels may necessitate a multiviscera transplantation.

The small bowel transplant patient will likely be malnourished, cachetic, and deconditioned. The rehabilitation team can expect intestinal transplant patients to have an ostomy stoma, gastostomy tube, and jejunostomy tube. The gastrostomy tube is usually clamped unless the patient experiences vomiting or bloating at which time the gastrostomy tube can be unclamped for decompression. During the postoperative period, enteral nutrition is required and oral intake is often as tolerated. Occasionally, intestinal absorption remains suboptimal and TPN is still required. Patients require frequent biopsies via colonoscopy as well as careful clinical and tissue observation for rejection (135). Persistent diarrhea because of dumping syndrome is extremely common and can be controlled with Immodium and tincture of opium. Nutrition consultation is usually necessary and helpful. The psychiatric manifestation
and narcotic dependency from chronic abdominal pain, and the use of TPN can hamper optimal rehabilitation (136). For this reason, intestinal transplant patients especially require neuropsychological evaluation and counseling. An appropriate goal during the rehabilitation phase is transition from IV pain medication to orally dosed medication as long as absorption is adequate.

**Lung Transplantation**

Hardy completed the first human lung transplant in 1963; during the next decade several attempts at human lung transplantation were completed, however, survival was brief (137–139). In the early 1980s several reports of long-term successful lung transplants surfaced. The first was a series of combined heart-lung transplants. These had the advantage of insuring adequate bronchial perfusion. Evolution of technique has allowed for single and double lung transplantation while avoiding bronchial airway ischemia. The United States waiting list for lung transplantation now exceeds 3,000.

Those persons with pulmonary fibrosis, chronic obstructive pulmonary disease (COPD), pulmonary hypertension, and chronic bronchitis often present as candidates for transplantation while in their 60s and 70s. Many have comorbidities and should receive a careful screening for malignancies. Among younger patients, those with cystic fibrosis often require transplantation (140). Such patients may have comorbidities associated with brittle diabetes or hepatobiliary dysfunction. While not absolute, general contraindications to lung transplantation are considered to be hepatic disease, severe renal dysfunction, and coronary or left ventricular heart disease not treatable with surgery or concomitant heart-lung transplant. Also of concern are those individuals with a history of malignancy. Those with a propensity for late recurrence of malignancy from any primary cancer should have a longer disease-free period before lung transplantation is considered. Several premorbid conditions have been noted to impact outcome among those receiving lung transplantation (139). Of interest to the rehabilitation community is the notion that premorbid functional status has been shown to be a strong predictor of survival among those receiving lung transplantation. Among those who require pre-transplant invasive ventilation, 1-year posttransplant survival is less than 50% (137). Both morbid obesity and poor nutrition may result in increased morbidity and mortality. Almost one half of the patients with cystic fibrosis and obstructive lung disease presenting for transplantation have metabolic bone dysfunction (40). Those with severe osteopenia and osteoporosis may experience decreased quality of life. The role of living donor transplantation has continued to grow. When selecting donor sites among living donors, the lower lobe is often chosen because of anatomic considerations and functional capacity.

In the immediate postoperative period, attempts are made to limit peak ventilatory pressures to 35 mm Hg or less. Attempts are also made to quickly wean the patient from high FiO2 in order to reduce oxygen toxicity. Complications with the bronchial anastomosis can occur and tracheal dehiscence or airway necrosis will result in extensive pulmonary compromise (141,142). Recent reports regarding the role of sirolimus in bronchial anastomotic complications have raised further concerns. Progressive strictures or bronchomalacia, can be common and is often recognized by direct visualization of a greater than 50% obstruction of the airway. Posttransplant bronchoscopy will focus on the evaluation of anastomotic areas and is often helpful in evaluating for secondary infection (141–144). Ischemia and/or reperfusion injury complicates some 15% of lung transplants and is usually manifested early on (145,146). This syndrome appears to manifest itself in an adult respiratory distress syndrome (ARDS)-like fashion.

During the rehabilitation period, pulmonary artery thrombosis and deep venous thrombosis have been noted to occur in as many as 12% of lung transplant recipients (147). Chest CT imaging may provide a useful diagnostic adjunct (148). The possibility of neoplasm should be considered in the differential of slowly growing infiltrates in this population. Chronic rejection can be manifested by obliterative bronchiolitis (149,150). Those with obliterative bronchiolitis demonstrate significant decrements in FEV-1 and management is focused on progressive immunosuppressive therapies (150). Postoperative delirium has commonly occurred in lung transplant recipients, and appears to be associated with prolonged cardiopulmonary bypass, the administration of cyclosporine, and other nonspecific factors.

Primary bilateral lung transplantation or combined lung and heart transplantation interrupts the vagus nerve and thus the normal stretch receptor input. Despite this interruption, resting spirometry should retain a classical pattern similar to that of normal subjects. Perhaps more important is the fact that cough response to airway irritants is relatively impaired. Those receiving heart-lung or bilateral lung transplants achieve relatively normal total lung capacity, but this may be impaired secondary to sternotomy from the surgical transplantation procedure itself. Among those with single lung transplantation, final pulmonary function becomes rate limited by the function of the remaining diseased lung.

Fink has noticed that exercise capacity can increase up to 10 years post lung transplantation (151). Peripheral mechanisms have been suggested as a source of improvement in exercise capacity (152). Stable lung transplant patients often remain with exercise capacity deficits and much of this is secondary to respiratory and lower limb muscle weakness. Wang (153) has noticed that such persons have a lower proportion of type I fibers and that they have a higher lactate and inosine monophosphate ratios, indicating a greater dependence on anaerobic metabolism. Schweiblmair evaluated 103 consecutive recipients of single-lung, bilateral lung, and heart-lung transplants. Cardiopulmonary exercise testing noted severe exercise intolerance and markedly impaired output parameters (153). Work capacity, oxygen pulse, tidal volume, and peak minute ventilation did not differ in patients following single, double-lung transplantation, or heart-lung transplants (154). These authors noted that despite the persistent limitations in aerobic capacity and work rate seen in many of the recipients, cardiopulmonary performance is reasonably well restored shortly after lung transplantation.
and heart-lung transplantation, further suggesting a peripheral component to the exercise dysfunction (154).

Cardiopulmonary exercise studies of lung transplant recipients have found low maximum oxygen consumptions because of an as yet unexplained mechanism (155). A study of six lung transplant recipients and six age- and sex-matched, healthy control subjects were studied to assess the possibility of a mitochondrial myopathy in liver transplant recipients (155). The transplant group had significantly lower percent-predicted maximum oxygen consumption than the control group. The authors felt that liver transplant recipients have an impaired maximal exercise capacity because of a disorder of peripheral oxygen utilization, perhaps related to a myopathic mechanism. A study of lung transplant recipients noted a significantly lower percent predicted maximum oxygen consumption than the control group and earlier onset of the anerobic threshold (156). Lung transplant recipients also appear to have an impaired maximal exercise capacity because of a disorder of peripheral oxygen utilization, perhaps related to a myopathic component. Hall has suggested that exercise capacity is limited due to disorders of potassium regulation that explain the peripheral muscle weakness (157). Ross has noted persistent ventilation-perfusion (V/Q) mismatches among those with lung transplantation during active exercise (158). Evans has also noted a persistent decrease in VO\textsubscript{2max} after lung transplantation that may be related to abnormalities of skeletal muscle oxidative capacity (159). Such deficits may be longstanding and data suggested that after 6 and 12 months, indices of skeletal and respiratory muscle function and VO\textsubscript{2max} improve further, but still remained lower than normal values (156). These authors concluded that in patients with heart-lung transplantation, skeletal and respiratory muscle function and exercise performance are reduced after surgery, and muscle function may improve with time but still appears impaired at 18 months.

Rehabilitation therapies offer techniques for enhanced coughing and mucociliary clearance, and providing treatment of the musculoskeletal system via exercise appears to enhance peripheral muscle capacity. Kesten (160) has suggested a preoperative rehabilitation program for those awaiting transplantation. Those providing exercise and inpatient rehabilitation to such persons should be aware of the need for close oxygen monitoring, and frequent rest periods until capacity is increased. The rehabilitation environment should also have the capacity to deal with excess pulmonary secretions, as this is a potential cause of secondary complications. Stiebellehner et al. (161) performed a study of an aerobic endurance training program (AET) in comparison to normal daily activities to evaluate exercise capacity in lung transplant recipients. Nine lung transplant recipients were examined and all patients underwent incremental bicycle ergometry, identical exercise tests were performed after 11 weeks of normal daily activities and then after a 6-week AET. Normal daily activities had no effect on exercise performance (161). The AET induced a significant decrease in resting minute ventilation and at an identical, submaximal level of exercise, a significant decrease in minute ventilation. Thus, in a carefully monitored setting an AET improves submaximal and peak exercise performance significantly.

**Heart Transplantation**

Cardiac transplantation is a therapy of choice for end-stage heart disease. Despite major advances with organ preservation and immunosuppression, the physiology of the cardiac allograft is not normal and this decreases exercise tolerance and functional performance. The physiatrist should have a basic understanding of the major physiological changes after heart transplantation, the beneficial effects of exercise and indications for both immediate acute inpatient rehabilitation and long-term rehabilitation follow-up.

The most common indications for heart transplantation surgery are cardiomyopathy, end-stage coronary artery disease, and inability to be weaned from temporary cardiac-assist devices after myocardial infarction or nontransplant cardiac surgery (159). Recipient selection criteria are stringent and approximately one quarter of patients found suitable for transplants die of cardiac disease before a suitable donor organ becomes available. Left ventricular-assist devices and artificial hearts can be used as interim support. The heart is transplanted in an orthotopic position with aorta, pulmonary artery, and the pulmonary vein. Venous return is provided by a single anastomosis joining the retained posterior wall of the recipient’s right atrium to that of the donor organ. More recently, the bicaval anastomosis technique has been used in performing heart transplantation (162). This technique differs from the standard mid-atrial anastomosis by keeping both atria intact and performing anastomoses in both the cava and pulmonary veins. This technique apparently avoids desynchronized atrial contractions and may improve cardiac performance by contributing more blood flow to the ventricles.

**Denervation**

Cardiac transplantation involves removing the diseased heart and leaving an atrial cuff, which results in the complete denervation of the donor heart with loss of both parasympathetic and sympathetic nerve connections. With absent parasympathetic innervation, resting heart rate is elevated between 90 and 110 beats/minute due to the loss of vagal tone (162). The donor heart rate will not respond to vagolytic muscle relaxants, anticholinergics, anticholinesterases, digoxin, nifedipine, or nitroprusside (162). Impairments in chronotropic responsiveness and ventricular function due to surgical interruption of postganglionic sympathetic fibers have been regarded as the major reasons for diminished exercise performance (163). Since the resting heart rate is usually elevated, the resting stroke volume is therefore smaller than normal. The denervated heart requires mechanisms other than neural to maintain normal function during activities of daily living.
The Frank-Starling mechanism remains a major factor to maintain exercise performance in the transplanted heart (162). With this mechanism, increases in venous return allows for an increased preload that subsequently increases the left ventricular end diastolic volume to result in an increase in the stroke volume. Thus, the assurance of adequate preload is especially important after heart transplantation. With vigorous exercise, further increases of cardiac output are mediated by chronotropic and inotropic responses to circulating catecholamines (162). Heart rate continues to rise after cessation of exercise as a result of circulating catecholamines and in recovery, heart rate falls slowly as plasma catecholamines are metabolized (163). This contrasts to the normal innervated heart in which increases of stroke volume and heart rate occur simultaneously rather than sequentially.

While the effects of denervation after heart transplantation are well described, it has become increasingly clear that denervation is not a persistent phenomenon. Bengel et al. (164) using Positron Emission Tomography and a catecholamine analog technique demonstrated partial restoration of myocardial sympathetic innervation that correlated with improvements in the capacity for exercise. In 29 heart transplant recipients, increased sympathetic activity was significantly correlated with the change in global ejection fraction in response to stress. The improvements in chronotropic response to exercise resulted in significantly better exercise performance in those patients who had innervation compared to those who had persistent denervation without recovery of sympathetic tone (165).

Exercise performed 6 months after transplantation surgery can partially restore cardiac sympathetic innervation. Bernardi et al. (166) performed a randomized clinical trial using cycling at 60% to 70% of peak oxygen consumption for 5 days/wk for 6 months versus a control group. Using a novel clinical method of RR-interval changes during cyclical baroreceptor stimulation at different frequencies, it was demonstrated that low frequency stimulation induces cardiovascular changes via sympathetic nerve activity while higher frequency stimulation produces changes only from parasympathetic nerve activity (167).

Results from the randomized trial showed significant gains for trained subjects in peak oxygen consumption, exercise time, and maximal workload compared to controls. Autonomic testing performed at the conclusion of the trial showed only significant increases in the trained group with low frequency baroreceptor stimulation. This response was not present at baseline, indicating pure sympathetic reinnervation of the heart (Table 46-3).

When prescribing an exercise program, it is worthwhile to understand that resting heart rate is increased yet there is a delay in heart rate elevation when beginning vigorous activity and there is also delayed return of resting heart rate after cessation of exercise (162). From a practical perspective, heart transplant patients may require 6 to 10 minutes of steady work to increase heart rate compared to 2 to 3 minutes in subjects without a denervated heart. Isometric exercise in cardiac transplant patients does not increase the cardiac output because of denervation (166). However, there is an expected increase in blood pressure due to an increase in α-adrenergic tone mediated by the central nervous system and not from increased circulating catecholamines.

Exercise training after heart transplantation increases the work capacity without increasing the peak heart rate. Yet, patients undergoing heart transplantation have preoperative inactivity and postoperative deconditioning that impacts exercise tolerance (167). One study estimates that the transplant patient have a 10% to 50% reduction in body mass due to prolonged preoperative physical inactivity combined with high corticosteroid administration (168). Consequently, maximal work output is reduced and maximal oxygen uptake is only two thirds of the normal age-matched population (169).

Although physical exercise has become an important part of the standard therapy for patients who have had an acute myocardial infarction and cardiac surgery, the role after heart transplantation has only recently been defined. In a randomized controlled trial, patients were enrolled approximately 2 weeks after transplantation surgery to an intensive rehabilitation protocol or to a control group that only received

<table>
<thead>
<tr>
<th>Routine Precautions During Initial Rehabilitation</th>
<th>Early Cardiovascular Changes</th>
<th>Later Cardiovascular Changes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sternal incision protection</td>
<td>Elevated resting heart rate</td>
<td>Improvement of exercise parameters with training</td>
</tr>
<tr>
<td>Cardiovascular monitoring using pulse rate, BP, pulse oximetry, and Borg rating of perceived exertion</td>
<td>Stroke volume increases before heart rate increases during exercise</td>
<td>Improvement in sympathetic tone to transplanted organ in some individuals</td>
</tr>
<tr>
<td>Ensure adequate preload</td>
<td>Delayed heart rate cool down after cessation of exercise</td>
<td>Improved left ventricular ejection fraction with training</td>
</tr>
<tr>
<td>Monitor for depression and other psychiatric symptoms, peripheral neuropathy, osteoporotic fractures</td>
<td>Decreased peak heart rate</td>
<td>Possible development of hypertension</td>
</tr>
<tr>
<td>Minimize exercise training during episodes of acute graft rejection</td>
<td>Decreased left ventricular ejection fraction</td>
<td>Possible development of coronary artery disease in transplanted heart</td>
</tr>
</tbody>
</table>

**TABLE 46.3 Summary of Precautions and Exercise Function During Early and Later Exercise Training Following Cardiac Transplantation**

| | Rehabilitation | Early Cardiovascular Changes | Later Cardiovascular Changes |
| | | | |
written guidelines for exercise (170). The exercise program was multifaceted and included supervision by a physical therapist according to each patient's specific needs. Strengthening exercises consisted primarily of closed chain resistive exercises such as bridging, with lifting of the hips while keeping the knees in a flexed, supine position. Other activities included half-squats, toe raises, abdominal flexion and pelvic tilts. Flexibility exercises, which emphasize chest expansion and thoracic mobility included thigh stretches, trunk twists, scapular squeezes and shoulder rolls. Aerobic exercise consisted of walking on a motorized treadmill or paddling on a bicycle ergometer. The goal of this exercise training was at least 30 minutes of continuous exercise at moderate intensity. Patients visited a cardiac rehabilitation clinic one to three times per week but the authors did note that some were unable to attend more than once a week due to transportation difficulty and graft rejection. As noted previously, routine transvenous endomyocardial biopsy is used to diagnose rejection, because other signs and symptoms are often absent and rejection can be detected before graft function deteriorates. Patients who did have rejection were instructed not to exercise until follow-up endomyocardial biopsy showed resolution of rejection and that usually required 2 weeks. Before discharge from the rehabilitation protocol, patients of both groups received written guidelines, primarily focusing on active assisted range of motion exercises, flexibility, and lower extremity strengthening. Patients who participated in the exercise-training program increased their capacity for physical work as compared with controlled patients who did not undergo exercise training. Resistance exercise may also prevent glucocorticoid myopathy (171). Niset and Kavanagh (172,173) have also shown improvements in physiological parameters with a prescribed exercise program. Proximal weakness measured by sitting to standing load and quality of life in these individuals.

The inpatient rehabilitation protocol should avoid strenuous upper limb exercises and teach proper positioning of the upper limbs during transfers or bed maneuvers. Specifically, patients should be instructed to keep their hands forward on the anterior thigh rather than reaching in an abducted arm posture to stabilize the upper torso during sit to stand maneuvers to minimize tension across the sternal incision (i.e., sternal precautions). In addition, patients should routinely have blood pressure, pulse rate, and pulse oximetry monitored during therapy. The Borg rating of perceived exertion can be used to keep the self-report below a rating of 13 (“somewhat hard”) during exercise. Of the twelve patients in the Joshi and Kevorkian (175) series, ten were discharged to community with the remaining two discharged back to acute care because of medical or surgical complications. Patients in this study were also discharged from acute inpatient rehabilitation to phase-II and phase-III cardiac rehabilitation programs that share similarities to outpatient cardiac rehabilitation programs prescribed for patients recovering from myocardial infarction and coronary artery bypass grafting.

Lastly, patients may be seen in acute rehabilitation—patients who are long-standing cardiac transplant recipients, who are admitted for other diagnoses and then require rehabilitative services. It is important to remember that the development of coronary artery disease in the transplanted heart can be observed in 50% of recipients and the presence of ischemic cardiac disease should be considered when prescribing a rehabilitation protocol (176). Heart transplantation is a life-saving technique for those individuals with end-stage cardiac disease. Rehabilitation is extremely important and now is considered to be the standard of care postoperatively. It is important for the rehabilitation physician to understand the concept of the denervated heart and to realize that this state can improve over time via mechanisms that are not fully understood. Isometric and resistive exercise training is an important part of the rehabilitation protocol and has been shown to be safe in heart transplant recipients. Inpatient rehabilitation is useful for a subset of patients who are severely deconditioned or have other medical comorbid illnesses. Precautions should include routine measurement of cardiac parameters and avoidance of pressure at the sternal incision. Cardiac rehabilitation programs appear to be worthwhile not only in the acute phase but also in the more subacute phases of recovery to improve the overall work load and quality of life in these individuals.

REHABILITATION ISSUES

Deconditioning

Patients on a transplant waiting list are severely deconditioned from their underlying disease and associated adverse physiological effects of these diseases. For example, a typical patient with liver failure is a gaunt, cachetic, jaundiced, edematous, ascitic individual. Failure of this vital organ interferes with synthetic function (albumin, acute phase reactants, clotting factors) as well as metabolism of amino acids, glucose and fat,
making this patient malnourished, coagulopathic, edematous, and immunosuppressed. The combination of protein wasting (usually from muscle) and increased weight from edema and ascites is damaging to patient conditioning. Compounding this state is the mechanical impingement to respiration from the massive ascites among those with hepatic dysfunction. Many of these patients also have pulmonary and renal manifestation of liver diseases (hepato-renal syndrome, hepato-pulmonary syndrome) and often require prolonged intensive care stay while waiting for donor organs, further exacerbating an already suboptimal condition.

Similar states of severe deconditioning are also seen in patients with other end-organ failure—it is easy to imagine the result of short-gut syndrome (malnutrition), obliterative lung disease, and heart failure. Even among those with end-stage renal dysfunction who are on dialysis, are wrought with complications, such as anemia, poor healing, hypertension, and prolonged periods of inactivity related to frequent dialysis. In diabetics awaiting pancreas transplants, the insulin supplement does not replace all factors, however undefined, produced by this organ. Thus, these patients are also malnourished, and usually have associated morbidities such as atherosclerotic cardiovascular disease, and poor healing, both of which hinder postoperative recovery.

The very act of surgery itself is a physiologic demand on these patients. Transplantation surgeries are usually complex, time consuming, and associated with much blood loss. The long duration under anesthesia and the paralysis of the patient in a single position, no doubt has physical implications with regard to postoperative deconditioning and rehabilitation. Similarly, cardiopulmonary bypass in lung and heart transplant creates additional stresses on a patient. The inexorable catabolic state seen in all postoperative surgical patients further weakens an already malnourished and protein-depleted patient (see Fig. 46-1).

In order to combat the severe deconditioning described above, initial rehabilitation requires a progressive rehabilitation program that begins with initial remobilization. Standing frames are useful to build endurance in deconditioned patients who exhibit proximal weakness. Tilt tables can be helpful to adapt postural reflexes for patients who have been recumbent for many months while in acute care.

**Neurocognitive Concerns**

Persons receiving transplantation have been noted to experience a variety of neurocognitive dysfunctions. Ghaus et al. (177) performed a study to evaluate such concerns among those with liver transplantation. They reviewed 41 consecutive patients who had 45 procedures associated with liver transplantation. Encephalopathy occurred after 28 procedures (62%) with immediate onset and no significant recovery before death or retransplantation in 11 (24%), slow recovery in eight (18%), and delayed onset (1 to 50 days, average 11) in six (13%) (177). Intermittent confusion and agitation with full recovery followed three, and focal and generalized seizures followed five (11%) procedures with multifocal myoclonus in two and status epilepticus in one. Isolated focal seizures followed two and myoclonus or unclassified seizures, one each. All patients with seizures had encephalopathy. Brain imaging showed atrophy in three instances, intracerebral hemorrhage in two, multiple infarctions in one, and intracerebral and subarachnoid hemorrhage with infarction in one. Cerebrospinal fluid analysis showed increased protein in three subjects. Neurological complications after transplantation were associated with increased mortality (177). Postoperative hypomagnesemia was associated with the development of nervous system complications. In another study of 40 persons who received liver transplantation, Buis (178) noted transient confusion among 48% of those receiving transplantation for alcoholic liver disease but only 6% of those who required transplantation secondary to hepatitis C. Studies of diffusion weighted imaging and recovery imaging (FLAIR) have noted hyperintense lesions. Psychologic consequences such as depression of transplantation are also an important consideration for the rehabilitation team. Such persons often benefit from psychologic evaluation while in rehabilitation. Rothenhausler (179) performed a study of 75 subjects to explore the prevalence of psychiatric disorders among orthotopic liver transplantation recipients, and to investigate how psychiatric morbidity was linked to health-related quality of life (HRQOL). Psychiatric morbidity was assessed using the Structural Clinical Interview for the DSM-III-R. A probable psychiatric diagnosis, according to DSM-III-R, was noted in 22.7% of the sample, 2.7% had full posttraumatic stress disorder (PTSD), 2.7% major depressive disorder (MDD), and 16% partial PTSD. Patients with PTSD symptoms demonstrated lower cognitive performance, higher severity of depressive symptoms, and more unfavorable perception of social support (180). Liver transplant-related PTSD symptomatology was associated with maximal decrements in HRQOL. The duration of intensive care treatment, the number of medical complications, and the occurrence of acute rejection were positively correlated with the risk of PTSD symptoms subsequent to liver transplantation (179). Neuropsychological performance is also an important consideration after transplantation. Temple et al. (180) examined the relationship between cardiac function and cognitive test performance among candidates for heart transplant. Increasing hemodynamic pressure variables, such as pulmonary artery systolic pressure, were associated with decreased cognitive performance on a measure of psychomotor speed and attention. In contrast, cardiac output and cardiac index appeared to not be significantly related to cognitive performance (173). Of note, these authors found that poor performance on cognitive tests among heart transplant candidates appears to be attentionally mediated (180). Activities of daily living are often impacted by cognition disorders after transplantation. Pütze (181) evaluated the ability of patients with heart transplant to perform day-to-day tasks (e.g., medication management, dietary regulation). They studied a series of 75 heart transplant candidates and 38 controls to examine the predictive validity of demographic, neuropsychologic, and cardiac function variables to a performance-based measure of instrumental activities of
Neuropathy and Myopathy

Severe acute polyneuropathy, complicating transplantation is not uncommon. Rezaiguia-Delclaux (182) reported three cases of severe acute motor deficit after orthotopic liver transplantation. In a context of graft dysfunction, these patients developed acute tetraplegia concomitant with early allograft failure. In these three patients, electrophysiological signs of sensorimotor axonal polyneuropathy were found.

El-Sabrout noted that peripheral neurological complications occur frequently in solid organ transplant recipients. Recent evidence supports the role of the immune system in initiating and perpetuating the ongoing neural damage in this entity. Infectious immune may initiate the immune attack, and these authors reported a series of transplant patients with a Gullian-Barre-like syndrome associated with CMV (183). Cyclosporine has now also been associated with a generalized sensorimotor neuropathy (184). An important consideration has been the improvement in function that may occur after transplantation. A study of nerve conduction velocity changes among patients with severe liver syndrome was performed by Lee et al. (185). Twenty-five patients admitted for a liver transplantation were involved in this study. All patients underwent nerve conduction study before liver transplantation and 6 months after liver transplantation. The sensory amplitudes and motor conduction velocities substantially increased after transplantation (185).

Immunosuppressive medications have also been associated with acute myopathy and exercise impairment (186). Critical illness myopathy has been associated with those receiving liver transplant and acute myopathy posttransplant has been reported to produce a tetraplegia-like picture (187). Campellone (188) reviewed the problems associated with acute myopathy among transplant recipients. One hundred consecutive adult patients were prospectively assessed for muscle weakness after orthotopic liver transplant. Electrodagnostic studies and muscle biopsies were performed on consenting affected patients. Seven patients developed acute persistent weakness after orthotopic liver transplant. Electrodagnostic studies were consistent with a necrotizing myopathy. In another study, histopathologic evaluation in five subjects revealed a necrotizing myopathy with loss of myosin thick filaments (187). A higher initial index of illness severity, dialysis requirement, and higher doses of glucocorticoids were associated with the development of myopathy. Patients with myopathy subsequently remained in the intensive care unit longer than unaffected patients. Stephenson et al. performed a study to determine the exercise impairment that occurs in liver transplant recipients. They noted that liver transplant recipients exhibit impaired peak exercise performance, similar to that observed after other solid organ transplants, probably as a result of chronic deconditioning or myopathy related to immunosuppressive medications (189).

QUALITY OF LIFE ISSUES

The initial focus in organ transplantation clinical research was demonstrating acceptable technical and survival outcomes. However, quality of life posttransplantation has now become an important parameter. Bucuvalas (190) performed a study of 77 pediatric liver transplant recipients ages 5 to 18 years, all of whom had had liver transplant at least 6 months previously. These authors evaluated quality of life issues. HRQOL in pediatric liver transplant recipients was lower than that reported for healthy children but similar to that for children with other chronic illness. Age at transplantation and maternal education predicted psychosocial function. HRQOL was decreased in a population of pediatric liver transplant recipients compared with the general population and similar to that for children with chronic illness. Forsberg et al. (191) performed a study of the coping strategies of adult liver transplant recipients. The group showed a homogeneous pattern of change in coping strategies. This prospective study indicated that coping style changed primarily at an individual level during the first year after liver transplantation while changes in coping strategies, were not common for the group. The usual coping strategy during the first posttransplant year was confrontational coping. A German study evaluated the quality of life among those who had participated as living-related liver donors for liver transplant 1 year after surgery (192). Donors viewed living related liver transplant positively. Quality of life after donation did not change. However, donors had a prolonged period of physical rehabilitation, and 41% experienced financial disadvantages (Fig. 46-2).
NOVEL CONCEPTS

Recent advances have allowed for the transplantation of neural and cartilaginous tissues with early success. Bentley (193) has recently described a randomized study demonstrating the potential efficacy of chondrocyte implantation in the knee. Menschlagraft transplantation appears efficacious among younger patients with focal knee dysfunction (194). These tissues are unique in that they appear to be privileged, and thus do not require extensive immunosuppression.

Partial facial transplantation may become an important option for those with severe deforming facial injuries or burns. Partial facial composite transplantation of the lower two thirds is technically feasible (195). These procedures raise ethical and immunological issues.

Kondziolka and Meltzer (196–198) have described neuronal transplantation in humans for chronic motor dysfunction following stroke. These preliminary studies have noted clinical improvements in motor function and integration of the neuronal transplanted when assessed via imaging techniques. Several authors have postulated the potential role that neuronal transplantation may have in the treatment of spinal cord injury (199–201). The role of these implants may be via direct in-growth or perhaps via neurotrophic growth factors. The rehabilitation of persons with neurologic and musculoskeletal disorders may be changed by these novel transplantation concepts in the near future.

Tolerogenic Therapy

The holy grail of transplant medicine has been the idea of tapering or eliminating the need for immunosuppressive therapy. The primary concept is to be able to reduce or eliminate the need for immunosuppressive therapy. Some have felt that the regulation of dendritic cells, which are the inducers of immune reactivity, is an important key (202). In a study by Starzl, 82 patients who had received kidney, liver, pancreas, or intestinal transplant had immunosuppressive pretreatment and subsequent dose spacing. They noted immune activation in graft biopsy samples but function was mostly maintained upon tapering or eliminating immunosuppressive therapy. The primary concept is to be able to reduce or eliminate the need for immunosuppressive therapy (204).

CONCLUSION

Transplant medicine has made tremendous strides during the last 30 years. Persons surviving transplantation are often left with neuromusculoskeletal disorders that require rehabilitation intervention. Many of the associated complications of transplantation are mediated via immunosuppressive therapy. Novel concepts such as tolerogenic therapy may help to moderate such effects. The rehabilitation specialist will continue to play a key role in helping to ameliorate the mobility and cognitive issues associated with solid organ transplant.

REFERENCES


CHAPTER

Introduction

The quest for effective treatment, prevention, and cure of the acquired immunodeficiency syndrome (AIDS) has resulted in one of the most concentrated research efforts in the history of medicine. Since its discovery in the early 1980s, the human immunodeficiency virus (HIV) has become one of the most studied of human pathogens. Although an outright cure for AIDS still seems to be a distant dream, developments have been breathtaking and, as of the time of this writing, the possibility of controlling the progression of HIV infection, preventing the decline in immune function that leads to AIDS, and enabling the reconstitution of the immune system destroyed by the virus, is within grasp. An effective vaccine, however, remains elusive.

HIV is important to the field of rehabilitation for two reasons: its occurrence in traditional rehabilitation patients and it can, in and of itself, result in disability. HIV is among the differential diagnosis of many disorders commonly seen in rehabilitation settings. In patients with known HIV infection, rehabilitation professionals need to be familiar with the many chemotherapeutic agents commonly used for treatment and with their side effects and drug interactions. This last task is particularly daunting given the accelerated process for U.S. Food and Drug Administration (FDA) review and approval of new treatments.

This chapter emphasizes the nature of HIV and the many ways that it can cause disability. Current medical treatment strategies are also discussed. However, the discussion is not intended to be comprehensive, and interested readers are urged to follow the literature closely and keep up to date in this rapidly evolving field. The medical management of HIV disease and its complications should be left to experts. Rehabilitation strategies used to manage functional impairments typical of those found in association with HIV, such as spinal cord injury, brain injury, neuropathies, and rheumatologic disorders, however, are similar to those used in disability resulting from other causes, as discussed elsewhere in this book.

HIV-Related Disability

The full extent of disability resulting from HIV is unknown. One of the earliest studies attempting to evaluate this, the NIH retrospective study, looked only at individuals with AIDS who were involved in research protocols and were referred for various rehabilitation interventions (1). A more comprehensive prospective study of disability in a population of men with HIV was obtained as part of the AIDS Time-oriented Health Outcome Study (ATHOS) (2,3). Of those individuals with AIDS, more than 50% reported difficulty with one or more instrumental ADLs, such as grasping and driving, nearly 30% with basic ADLs, such as bathing and dressing, and more than 40% with basic mobility. Even among symptomatic seropositive individuals without AIDS, nearly 30% reported difficulties with instrumental ADLs, more than 10% with basic ADLs and 15% with basic mobility. With the introduction of highly active antiretroviral therapy (ART, formerly HAART), it has become possible to control and even reverse the progression of HIV. Although functional abilities generally improve with treatment, some degree of functional impairment often remains, and continued functional decline has been reported (4–8). Quality-of-life issues have also become more significant as survival improves (9–11).

Natural History of HIV

Pathogenesis

HIV type 1 infection (HIV-1), the most prevalent form of the virus, results in the destruction of CD4-positive lymphocytes, leading to a decline in their numbers. In acute primary infection, HIV-1 replicates briskly and viral titer rises rapidly (12–14). Within 1 week of onset, $10^5$ to $10^6$ infectious particles per microliter of plasma can be measured (15,16). Fifty to seventy percent of patients experience a clinical syndrome, which may include fever, rash, sore throat, lymphadenopathy, splenomegaly, myalgias, arthritis and, rarely, meningitis.

Within a few weeks to months, seroconversion occurs and HIV-1 antibodies can be measured by the enzyme-linked immunosorbent assay (ELISA) and Western blot tests. Rapid viral replication persists, however, in the bloodstream, lymphoid tissue, and central nervous system (CNS) leading to a high turnover of infected cells (13,17). A steady state of plasma HIV ribonucleic acid (RNA) level is established within 6 to 12 months of seroconversion, and measurements of the viral load are prognostic (18,19). Levels of $10^6$ HIV-1 RNA copies or greater per microliter of blood correlate with rapid
progression over a few years, $10^3$ to $10^5$ with progression in 8 to 10 years and less than $10^3$ with nonprogression (20). Unfortunately, rapid viral turnover and mutation lead to the appearance of resistant strains if viral suppression is not complete (21–23).

Classification

The Centers for Disease Control and Prevention (CDC) developed a revised classification scheme in 1993 based on the development of signs and symptoms and the CD4 lymphocyte count (24). The classification was revised again in 1999, allowing HIV-seropositive individuals without other AIDS-defining conditions to be reported (25). This scheme is summarized in Table 47.1. Group A patients are seropositive but essentially symptom free. Group C patients have one or more AIDS-indicator illnesses. A comprehensive list of AIDS-indicator conditions is beyond the scope of this chapter, but suffice it to say that they include most opportunistic infections and malignancies, HIV dementia, and the AIDS wasting syndrome. Group B patients have early symptoms of immune deficiency in the absence of an AIDS-indicator condition. Patients are further subdivided into categories 1, 2, and 3, corresponding to CD4 counts of 500/mm$^3$ or above, 200 to 499, and less than 200, respectively. Patients falling into either group C or category 3 are classified as having AIDS.

Demographics

The World Health Organization (WHO) estimates that as of December, 2007, 30.8 million adults and 5 million children worldwide are living with HIV-1 infection. Half of those infected are women. HIV continues to run rampant in less developed areas of the world, particularly in sub-Saharan Africa and Southeast Asia, in spite of efforts aimed at prevention (26). About 1.3 million people are infected with HIV in North America. In the United States, there are now about 17,000 deaths annually and in 2005, among persons with HIV, 20% were female, 35% white, 44% black, and 19% Hispanic, reflecting the continuation of a significant shift into minority populations (27).

Documented Routes of Transmission

Documented routes of transmission include sexual contact with infected individuals, percutaneous or mucous membrane exposure to infected blood or body fluids, or transplacental, perinatal, or breast milk transmission from mother to infant. In 2005, of 322,000 HIV-positive men in the United States, 59% were exposed through male-to-male sexual transmission, 20% through intravenous drug use, and 11% through high-risk heterosexual sex. Of the 96,000 women, 65% contracted HIV through high-risk heterosexual exposure and 35% through intravenous drug use (27).

**Occupational Exposure**

The risk of HIV-1 infection is estimated to be 0.3% after percutaneous injury involving infected blood (28) and 0.1% after mucous membrane exposure (29). Universal precautions should be used at all times that contact with blood or body fluids may be anticipated. These include the use of gloves to prevent potential exposure to infectious body fluids or broken skin during procedures such as wound care and electromyography (EMG), and the use of a gown, goggles, and a mask during procedures in which spattering of blood is possible.

Postexposure prophylaxis (PEP) is recommended for percutaneous or mucous membrane exposure to blood or visibly bloody fluid, semen, vaginal secretions, cerebrospinal, synovial, pleural, pericardial, and amniotic fluids. In a case control study, postexposure zidovudine decreased HIV-1 infection by 81% (30). Recently updated guidelines by the New York State Department of Health recommend initiating PEP as soon as possible after it has been determined that an exposure has occurred, ideally within 2 hours and not later than 36 hours. The combination of Combivir with tenofovir is one preferred regimen, but consultation with an HIV specialist is highly recommended in order to choose a regimen based on the source patient’s HIV strain and antiretroviral resistance, if known (31). A national PEP line is maintained by the University of California, San Francisco (http://www.ucsf.edu/hivcntr/PEPline/) 1-888-448-4911. Expert clinicians are available 24 hours a day, to offer advice in the event of an exposure. Nevirapine is not recommended for PEP due to the incidence of severe hepatitis (32). A medical evaluation, counseling, and follow-up by an experienced physician should be provided promptly, but initiation of PEP should not be delayed (33,34) <http://www.hivatis.org/>. Recommendations for pregnant health care workers are similar except that efavirenz and amprenavir should be avoided (35,36).

<table>
<thead>
<tr>
<th>CD4 Cell Count</th>
<th>Group A: Asymptomatic, Acute HIV, or Generalized Lymphadenopathy</th>
<th>Group B: Symptomatic, not A or C</th>
<th>Group C: Indicator Conditions Present</th>
</tr>
</thead>
<tbody>
<tr>
<td>≥500/mm$^3$</td>
<td>A1</td>
<td>B1</td>
<td>C1*</td>
</tr>
<tr>
<td>200–499/mm$^3$</td>
<td>A2</td>
<td>B2</td>
<td>C2*</td>
</tr>
<tr>
<td>&lt;200/mm$^3$</td>
<td>A3*</td>
<td>B3*</td>
<td>C3*</td>
</tr>
</tbody>
</table>

* Indicates the presence of AIDS.
Prevention of Transmission Among the Disabled
One of the most significant things rehabilitation professionals can do to stop the spread of HIV is to educate their patients as to the modes of transmission and means of preventing transmission. High-risk behaviors must be strongly discouraged, and education in safer sex practices should be included as part of all discussions on sexuality.

The most effective protective measure is the use of a barrier device whenever contact with infectious fluids is anticipated. This includes the use of condoms, the dental dam for oral sex, and the use of gloves or a finger cot for digital anal penetration. Safe touching should be substituted for scratching, biting, and abrading activities that might lead to inadvertent fluid exposures. Water-soluble lubricants should be used particularly in insensate areas to reduce the likelihood of inadvertent abrasion. Mutual masturbation may be suggested as an activity that is virtually risk-free. Patients should be reminded to clean and disinfect all enrichment aids (sex toys) using a solution such as 10% chlorine bleach.

When patients experience direct exposure to infectious fluids, prophylaxis should be offered as it is to health care workers. Because about a fourth of children delivered vaginally to untreated infected mothers will also be infected, prophylaxis should also be used to reduce vertical transmission. A prospective trial showed that pregnant women who took zidovudine could decrease the risk of perinatal transmission by 67% (37,38), and several subsequent trials have shown that the risk of mother-to-child transmission can be reduced substantially by perinatal treatment of both mother and neonate (39,40). Prevention of transmission in the adolescent patient demands particular attention (41). The avoidance of HIV transmission among patients who lack the capacity to modify their behavior, such as brain-injured individuals with impaired impulse control, remains a difficult area without clear-cut solutions (42,43).

CLINICAL EVALUATION OF HIV
Serologic Testing
The mainstay of HIV-1 testing involves screening for anti-HIV-1 antibodies with the ELISA, with the Western blot test used for confirmation. Serologic testing can only detect infection after seroconversion has occurred, which may take up to 3 to 6 months after acute infection. When ELISA screening and confirmatory Western blot testing are used, the false-positive rate is negligible. Pre- and post-test counseling are advisable to educate patients about the need for repeat testing, as well as modification of high-risk behaviors and, in the event of a positive test, the need to seek specialized medical care. Home test kits, in which samples of saliva, blood, or urine can be collected by the patient at home or in a doctor’s office and sent directly to testing laboratories, are available. Recently, the CDC has endorsed routine testing for HIV in medical settings with an opt-out clause. This policy change is hoped to help identify patients with HIV who may otherwise not have sought testing or whose physicians may not have thought were at increased risk (44).

Evaluation of the Seropositive Patient
It is particularly important to identify seropositive individuals both because medical therapy for HIV is highly beneficial and in order to identify the individuals at risk for spreading the infection to others. In patients with unexplained sweats, fevers, or weight loss, skin lesions characteristic of opportunistic infections or malignancy, dyspnea unrelated to intrinsic cardiac or pulmonary disease, or altered neurologic function not attributable to a specific lesion, injury, or toxin, HIV infection must be considered in the differential diagnosis. The presence of other sexually transmitted diseases (STDs) also should be considered an indication for HIV testing. Even the presence of diseases seen at higher rates in HIV patients, such as tuberculosis (TB), should be considered an indication for testing.

The physical exam should focus on findings often seen in association with HIV. The skin and oral cavity should be carefully examined for mucocutaneous lesions, particularly those associated with Kaposi’s sarcoma (KS). Lymphadenopathy and hepatosplenomegaly are also important indicators. Because of the propensity of HIV to involve the nervous system, a detailed neurologic examination is essential. Peripheral neuropathies, cognitive disorders, myelopathy, and focal CNS dysfunction may all be seen in association with HIV. Although cognitive dysfunction is not common prior to the development of AIDS, a neuropsychologic evaluation is helpful to establish a baseline.

The laboratory evaluation should include routine chemistries and a complete blood count. A mild increase in transaminases may be seen, along with an elevated erythrocyte sedimentation rate, anemia, and general cytopenias. Serum protein electrophoresis may demonstrate an increase in immunoglobulins. Lymphocyte evaluation may demonstrate a decrease in CD4-positive lymphocytes or a reversal in the CD4/CD8 ratio. Viral load tests such as the polymerase chain reaction (PCR) or branched deoxyribonucleic acid (DNA) are important for establishing prognosis and for evaluating the effectiveness of antiretroviral treatments.

TREATMENT STRATEGIES IN AIDS
Because clinically latent stages of HIV-1 infection involve rapid viral replication, strategies now aim to decrease steady-state viral RNA levels. The benefits of decreasing HIV viral load must be weighed against side effects and the risk of developing resistant strains (45,46). Published guidelines are meant to be interpreted by specialized providers <http://www.hivatis.org> (47,48). In general, the aim is to start antiretroviral medications before complications related to immunosuppression occur and when the patient can achieve maximum adherence. Medications commonly used in HIV along with some of their indications and side effects are summarized in Table 47-2. Numerous important side effects and drug interactions may
<table>
<thead>
<tr>
<th>Drug</th>
<th>Common Usage</th>
<th>Side Effects</th>
<th>Drug Interactions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Antiretrovirals</strong>&lt;br&gt;(nucleoside analog RT inhibitors)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zidovudine (AZT)</td>
<td>Used in combination</td>
<td>Nausea, headache, fatigue, anemia, neutropenia, myopathy</td>
<td>Increased hepatotoxicity of amphotericin B, ganciclovir</td>
</tr>
<tr>
<td>Didanosine (ddI)</td>
<td>Used in combination</td>
<td>Painful neuropathy, pancreatitis, mitochondrial toxicity, lactic acidosis</td>
<td>May decrease absorption of other drugs</td>
</tr>
<tr>
<td>Zalcitabine (ddC)</td>
<td>Rarely used due to toxicity</td>
<td>Painful neuropathy, pancreatitis, oral ulcers, granulocytopenia, rash, fever, hepatitis</td>
<td>Increased toxicity with foscarnet and valproic acid; alcohol may increase risk of pancreatitis</td>
</tr>
<tr>
<td>Lamivudine (3TC)</td>
<td>Used in combination</td>
<td>GI intolerance</td>
<td></td>
</tr>
<tr>
<td>Emtricitabine (FTC)</td>
<td>Used in combination</td>
<td>GI intolerance</td>
<td></td>
</tr>
<tr>
<td>Abacavir</td>
<td>Used in combination</td>
<td>GI intolerance, hyper-sensitivity reaction may be fatal, increased risk of MI</td>
<td></td>
</tr>
<tr>
<td>Stavudine (d4T)</td>
<td>Rarely used due to toxicity</td>
<td>Painful neuropathy, hepatic dysfunction, mitochondrial toxicity, lactic acidosis</td>
<td></td>
</tr>
<tr>
<td><strong>Antiretrovirals</strong>&lt;br&gt;(nucleotide analogue RT inhibitors)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tenofovir</td>
<td>Used in combination</td>
<td>GI side effects</td>
<td></td>
</tr>
<tr>
<td><strong>Antiretrovirals</strong>&lt;br&gt;(nonnucleoside analogue RT inhibitors)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Efavirenz</td>
<td>Used in combination</td>
<td>Rash, neurologic and sleep disturbances, lipid abnormalities</td>
<td>Many drug interactions, may decrease PI levels</td>
</tr>
<tr>
<td>Nevirapine</td>
<td>Used in combination</td>
<td>Rash, hepatitis</td>
<td>May decrease protease inhibitor levels</td>
</tr>
<tr>
<td>Delavirdine</td>
<td>Used in combination</td>
<td>Rash</td>
<td>May decrease didanosine levels</td>
</tr>
<tr>
<td>Etravirine</td>
<td>Works in NNRT-resistant virus</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Antiretrovirals</strong>&lt;br&gt;(protease inhibitors)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Saquinavir</td>
<td>Used in combination</td>
<td>GI distress, elevated LFTs</td>
<td>Decreased bioavailability of Ca channel blockers</td>
</tr>
<tr>
<td>Indinavir</td>
<td>Used in combination</td>
<td>Hepatotoxicity, nephrolithiasis, glucose intolerance</td>
<td>As above, plus increased desipramine, rifabutin, ergotamine, and amiodarone levels; prolonged sedation with midazolam and triazolam; decreased OCP effectiveness; fatal arrhythmias with astemizole</td>
</tr>
<tr>
<td>Ritonavir</td>
<td>Used to boost other PI levels</td>
<td>GI upset, perioral paresthesias, alteration of taste, headache, elevated LFTs, lipid abnormalities</td>
<td>Similar to indinavir, but may be more pronounced, failure of oral contraception, multiple P450 interactions—most potent P450 CYP3A inhibitor decreases some drug levels</td>
</tr>
<tr>
<td>Nelfinavir</td>
<td>Used in combination</td>
<td>Diarrhea, lipid abnormalities</td>
<td></td>
</tr>
<tr>
<td>Amprenavir</td>
<td>Used in combination</td>
<td>GI disturbance</td>
<td></td>
</tr>
<tr>
<td>Fos-amprenavir</td>
<td>Prodrug of amprenavir</td>
<td>GI disturbance</td>
<td></td>
</tr>
<tr>
<td>Lopinavir/ritonavir</td>
<td>Used in combination</td>
<td>GI disturbance, glucose intolerance, lipid abnormalities</td>
<td>See ritonavir</td>
</tr>
<tr>
<td>Atazanavir</td>
<td>Used in combination</td>
<td></td>
<td>Contraindicated with proton pump inhibitors and interacts with H2 blockers.</td>
</tr>
<tr>
<td>Tipranavir</td>
<td>Used in combination</td>
<td>Hyperbilirubinemia</td>
<td></td>
</tr>
<tr>
<td>Darunavir</td>
<td>Used in combination</td>
<td>GI disturbance</td>
<td>Multiple, difficult to predict</td>
</tr>
</tbody>
</table>
### CHAPTER 47
REHABILITATION OF THE INDIVIDUAL WITH HIV

#### Antiretrovirals (novel agents)

<table>
<thead>
<tr>
<th>Medication</th>
<th>Class Description</th>
<th>Adverse Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enfuvirtide</td>
<td>Fusion inhibitor</td>
<td>Injection site reactions</td>
</tr>
<tr>
<td>Maraviroc</td>
<td>CCR5 entry inhibitor</td>
<td>Hypotension</td>
</tr>
<tr>
<td>Raltegravir</td>
<td>Integration inhibitor</td>
<td>Elevated CK</td>
</tr>
</tbody>
</table>

#### Combination agents (names outside the United States may differ)

<table>
<thead>
<tr>
<th>Medication</th>
<th>Class Description</th>
<th>Adverse Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Combivir</td>
<td>Zidovudine + lamivudine</td>
<td>Similar to individual components</td>
</tr>
<tr>
<td>Eptinez</td>
<td>Abacavir + lamivudine</td>
<td>Similar to individual components</td>
</tr>
<tr>
<td>Truvada</td>
<td>Zidovudine + lamivudine + abacavir</td>
<td>Similar to individual components</td>
</tr>
<tr>
<td>Atripla</td>
<td>Tenofovir + emtricitabine + efavirenz</td>
<td>Similar to individual components</td>
</tr>
<tr>
<td>Kaletra</td>
<td>Lopinavir + ritonavir</td>
<td>Similar to individual components</td>
</tr>
</tbody>
</table>

#### Prophylactic agents

<table>
<thead>
<tr>
<th>Medication</th>
<th>Prophylaxis and Treatment</th>
<th>Adverse Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trimethoprim/sulfamethoxazole</td>
<td>PCP prophylaxis and treatment, toxoplasmosis prophylaxis</td>
<td>Rash, fever, neutropenia, thrombocytopenia, hepatitis, headache, meningitis, hyperkalemia</td>
</tr>
<tr>
<td>Dapsone</td>
<td>PCP prophylaxis</td>
<td>Rash, fever, hemolysis in G6PD deficiency</td>
</tr>
<tr>
<td>Pentamidine</td>
<td>PCP prophylaxis and treatment</td>
<td>Aerosol: bronchospasm; IV: nephritis, pancreatitis, hypo- or hyperglycemia, rash, neutropenia</td>
</tr>
<tr>
<td>Rifabutin</td>
<td>MAC prophylaxis</td>
<td>GI, auditory, iritis</td>
</tr>
</tbody>
</table>

#### Common treatments

<table>
<thead>
<tr>
<th>Medication</th>
<th>Condition</th>
<th>Adverse Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acyclovir</td>
<td>Herpes simplex, zoster</td>
<td>Diarrhea, phlebitis (IV)</td>
</tr>
<tr>
<td>Ganciclovir</td>
<td>CMV</td>
<td>Neutropenia, GI</td>
</tr>
<tr>
<td>Foscarnet</td>
<td>CMV</td>
<td>Nephrotoxicity, seizure, nausea, hyponatremia, cakemia and natrema, nephrotoxicity, headache</td>
</tr>
<tr>
<td>Cidofovir</td>
<td>CMV</td>
<td>Nephrotoxicity</td>
</tr>
<tr>
<td>Fluconazole</td>
<td>Antifungal: candida and cryptococcus</td>
<td>Hepatotoxicity, nausea, rash</td>
</tr>
<tr>
<td>Ketoconazole</td>
<td>Antifungal: candida</td>
<td>GI upset, pruritis, headache, rash</td>
</tr>
<tr>
<td>Itraconazole</td>
<td>Antifungal: histoplasmosis, aspergilla</td>
<td>Hepatitis, GI upset, rash headache</td>
</tr>
<tr>
<td>Amphotericin B, Liposomal amphot B</td>
<td>Antifungal: serious fungal infections</td>
<td>Nephrotoxicity, fever, chills, hypomagnesemia, hypocalcemia</td>
</tr>
</tbody>
</table>

#### Notes

CMV, cytomegalovirus; IV, intravenous; LFT, liver function test; MAC, Mycobacterium avium complex; OCP, oral contraceptive preparation; PCP, Pneumocystis carinii pneumonia.
that blocks HIV entry into the cell. Since some strains of HIV positive T-cells. Maraviroc acts as a CCR5-binding antagonist coreceptors that HIV commonly uses to bind and enter CD4-
was FDA approved in the United States. CCR5 is one of the developed resistance to other agents (69).
cutaneously and thus is most often used in patients who have membrane thus inhibiting viral entry. It is injected sub-
formation change required to fuse the virion with the CD4 membrane protein on the HIV virion and prevents the con-
Currently, there is one fusion inhibitor, enfuvirtide, approved for use by the FDA. This polypeptide binds to GP41 trans-
In 1987, the FDA approved zidovudine (AZT), a nucleoside analog RT inhibitor. Subsequently approved were other nucle-
levels by addition of low-dose ritonavir, which is a potent blocking the cleavage of HIV-1 polyproteins into component proteins. Combinations of protease inhibitors and RT inhibitors have been shown to reduce plasma HIV-1 RNA levels below detectable levels and to reduce morbidity and mortality. The protease inhibitors are often “boosted” to higher plasma levels by addition of low-dose ritonavir, which is a potent inhibitor of the cytochrome p450 system.
Protease Inhibitors
As of this writing, ten protease inhibitors have been approved by the FDA for HIV infection: saquinavir, indinavir, ritonavir, nelfinavir, amprenavir, fos-amprenavir, lopinavir, atazanavir, tipranavir, and darunavir (57,64–68). These result in the production of immature, noninfectious HIV-1 particles by blocking the cleavage of HIV-1 polyproteins into component proteins. Combinations of protease inhibitors and RT inhibitors have been shown to reduce plasma HIV-1 RNA levels below detectable levels and to reduce morbidity and mortality. The protease inhibitors are often “boosted” to higher plasma levels by addition of low-dose ritonavir, which is a potent inhibitor of the cytochrome p450 system.
Novel Agents
Currently, there is one fusion inhibitor, enfuvirtide, approved for use by the FDA. This polypeptide binds to GP41 trans-
Immune Stimulants
Cytokines have been used in clinical trials to stimulate production of CD4 cells (71). Interleukin-2 can stimulate CD4 cell production, but it also increases virus production. The use of stimulating cytokines in conjunction with potent antiretroviral combinations is effective but is not often necessary.
COMMON CAUSES OF DEBILITY
Although HIV itself is felt to be responsible for many forms of functional compromise, particularly in the CNS, a large number of disabling syndromes are associated with opportunistic infections and malignancies. The following sections are organized by the types of impairment syndromes that result and, to the extent practical, in order of the timing of their appearance in HIV infection. A summary of these is listed in Table 47-3.
Respiratory Compromise and Disseminated Infection
Pathogenic Fungi
Endemic pathogenic fungi such as histoplasma capsulatum, blastomyces dermatidis, and coccidioides immitis can cause disease in normal hosts and particularly in HIV infection and AIDS. Patients often present with slowly progressive wasting, fevers, weight loss, or ulcerations of skin. Bone marrow dysfunction may occur and, occasionally, disseminated intravascular coagulation and death. The chest x-ray may be normal or show a progressive infiltrate. The serum test for histoplasma polysaccharide is specific and sensitive. Treatment consists of induction with amphotericin B and maintenance with itraconazole (72,73).
Mycobacterium TB
TB caused by mycobacterium TB (mTB) is common and severe, and seen early in the course of AIDS, often at CD4 counts above 200 cells/μL (74,75). Symptoms may be subtle. TB may disseminate to a wide variety of organs including the brain and spine. Guidelines for the diagnosis, treatment, and prophylaxis of infectious TB have been detailed by the CDC and elsewhere (76–79). Of concern, however, are outbreaks of multi–drug-resistant TB (80,81).
**Pneumocystis Jiroveci Pneumonia (PCP)**

*Pneumocystis jiroveci* (formerly *Pneumocystis carinii*) is a fungus-like pathogen that causes PCP, the most common AIDS-defining illness in the United States, accounting for about 18% of cases (82,83). Seen throughout the course of the disease, symptoms may develop gradually and include dyspnea on exertion, shortness of breath, cough, weight loss, fever, and night sweats. The chest x-ray may show a diffuse alveolar infiltrate and serum lactate dehydrogenase (LDH) may be elevated. Diagnosis is most often made by identification of *Pneumocystis* organisms in expectorated sputum or from bronchoscopic specimens.

A room air blood gas with PaO$_2 \leq 70$ mm Hg, indicates severe disease, which is treated with steroids in addition to intravenous trimethoprim/sulfamethoxazole (79,95–100). Primary prophylaxis against PCP in patients with a CD4 count less than 200 cells/μL is recommended. Prophylaxis with trimethoprim-sulfamethoxazole or dapsone is a mainstay of AIDS treatment when CD4 counts are less than 200.

**Mycobacterium Avium Complex**

*Mycobacterium avium* complex (MAC), an atypical mycobacterial infection, is one of the most common opportunistic infections in AIDS patients in the United States, with an incidence of 18% to 43% in untreated patients (84–86). It may be present as a nonpathogenic colonizer or may cause disease in the bone marrow, lungs, liver, colon and rarely, the CNS. It occurs late in the course of the disease and may be indolent, with fevers, night sweats, weight loss and wasting, cough, diarrhea, abdominal pains, lymphadenopathy, and lymph node pain.

Effective drug combinations for treatment of MAC include clarithromycin, with ethambutol as an alternative. Rifabutin, clofazamine, ethambutol, and ciprofloxacin azithromycin and amingoglycosides (aminoglycosides) have also been used (87). Clarithromycin has been shown to be effective in preventing disseminated MAC (88). Primary prophylaxis against MAC with azithromycin is recommended for those with CD4 counts less than 50 to 75 cells/μL (89).

**Gastrointestinal and Swallowing Disorders**

Gastrointestinal disease in AIDS is common and presents with diarrhea, weight loss, biliary disorders, abdominal pain, dysphagia, and oral disease (90–93). A wide range of causes including cytomegalovirus (CMV), herpes simplex virus (HSV), MAC, intestinal parasites, and malignancies such as KS and non-Hodgkin’s lymphoma are seen. Drug reactions are also a frequent cause. Dysphagia is most often caused by *Candida albicans*, HSV, or CMV esophagitis. Other causes include hairy leukoplakia, KS, human papilloma virus infection, and neurogenic causes. The Sicca syndrome, seen often in association with lymphadenopathy, may be treated symptomatically with increased fluids, artificial tears, and careful attention to oral hygiene. Colitis often presents with abdominal pain and bloody or mucous diarrhea and is most often caused by CMV, which responds to treatment with ganciclovir (94). *Clostridium difficile*, *Campylobacter jejuni*, *Entamoeba histolytica*, and *Shigella flexneri* are also seen in colitis in AIDS patients.

**HIV-Related Liver Disease**

Over the past several years, liver disease has risen to prominence in HIV patients. Many ARTs are hepatotoxic and a high percentage of HIV patients are also infected with hepatitis B, C, or both. Moreover, a higher percentage of HIV patients develop chronic infections that may progress more rapidly to cirrhosis than would otherwise be expected. Treatment for hepatitis C with pegylated interferon and ribavirin is commonly used in HIV patients and hepatitis B can be treated with the RT inhibitors lamivudine or emtricitabine, and tenofovir (79,95–100).

**Candidiasis**

Oral candidiasis (thrust) occurs in 43% to 93% of AIDS patients (101). It most often appears as creamy white, yellow or erythematous patches on the palate, tongue, and buccal mucosa and causes angular cheilitis and fissures at the corners of the mouth. With more severe immunodeficiency, candida may extend into the pharynx and esophagus. It presents more often in patients with CD4 counts less than 200 cells/μL, but can occur at higher counts as well. Candidiasis may be treated with oral nystatin suspension, clotrimazole troche, fluconazole or other azoles, or echinocandins.

**Kaposi’s Sarcoma**

KS is a highly vascular malignant lesion containing endothelial cells and spindle-shaped mesenchymal cells and is associated with HHV8 coinfection. It can involve skin, oral mucosa, and visceral organs and often appears as dark blue or purple papules or nodules. Cosmetically, lesions can be treated by laser bleaching and spot radiation. Often, KS will regress with immune reconstitution on ART (102). More severe disease has been treated with antiretrovirals and drugs such as foscarnet (103), localized radiation, immunotherapy with interferon alpha, cytotoxic chemotherapy, and liposomal doxorubicin (104,105). Intralesional human choriocarcinoma gonadotropin (hCG) has also been effective (106). Associated edema may be controlled with the use of compressive garments, and painful lesions in the feet may be managed with the use of soft insoles and by unweighting the affected extremity.

**Intestinal Parasites**

Cryptosporidium parvum is an intracellular protozoan usually seen with CD4 counts below 200. It is responsible for about 1% of AIDS deaths (107). Oral paromomycin or azithromycin may be effective in decreasing diarrhea and stabilizing body weight (108). Microsporidiosis causes similar symptoms and may be treated with albendazole (109). Other intestinal parasites include isospora belli, *E. histolytica*, and giardia lamblia (110,111).
<table>
<thead>
<tr>
<th>Cause of Impairment</th>
<th>When Seen</th>
<th>Features</th>
<th>Prognosis</th>
<th>Management</th>
</tr>
</thead>
<tbody>
<tr>
<td>Respiratory compromise and disseminated infection</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pathogenic fungi</td>
<td>Pre-AIDS and AIDS</td>
<td>Wasting, fevers, etc.</td>
<td>Benign or fulminate</td>
<td>Antifungal agents</td>
</tr>
<tr>
<td>mTB</td>
<td>Early AIDS</td>
<td>Non-specific fever, cough, etc.</td>
<td>Fair, good with ART</td>
<td>Antituberculin therapy</td>
</tr>
<tr>
<td>PCP</td>
<td>Throughout AIDS</td>
<td>Non-specific fever, cough, etc.</td>
<td>Good except in late AIDS</td>
<td>TMPX, dapsone, pentamidine; prophylaxis in all AIDS</td>
</tr>
<tr>
<td>MAC</td>
<td>Late AIDS</td>
<td>Fever, night sweats, weight loss, etc.</td>
<td>Poor but better with ART</td>
<td>Clarithromycin, ethambutol etc.; combination therapy, prophylaxis when CD4 count is &lt;50–75</td>
</tr>
<tr>
<td>GI and swallowing disorders</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neurogenic swallowing disorders</td>
<td>Throughout AIDS, especially late</td>
<td>Related to underlying neurologic disorder</td>
<td>Related to underlying disorder</td>
<td>Video fluoroscopy, modified swallow, parenteral nutrition</td>
</tr>
<tr>
<td>Sicca syndrome</td>
<td></td>
<td>Impaired lacrimation and salivation</td>
<td>Benign course</td>
<td>Symptomatic treatment (artificial tears, oral lubricants, etc.)</td>
</tr>
<tr>
<td>Candidiasis</td>
<td>Pre-AIDS and AIDS</td>
<td>Oral plaques, dysphagia</td>
<td>Good with treatment</td>
<td>Topical antifungal agents; systemic agents in severe cases</td>
</tr>
<tr>
<td>Kaposi's sarcoma</td>
<td>Throughout AIDS</td>
<td>Characteristic lesions</td>
<td>Slow progression</td>
<td>Appropriate chemotherapy</td>
</tr>
<tr>
<td>Intestinal parasites</td>
<td>Late AIDS</td>
<td>Diarrhea</td>
<td>Good if diagnosed</td>
<td>Antiparasitic agents</td>
</tr>
<tr>
<td>Weakness and fatigue</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chronic fatigue</td>
<td>Pre-AIDS and AIDS</td>
<td>Many causes</td>
<td>Fair</td>
<td>Energy conservation and pacing</td>
</tr>
<tr>
<td>Fibromyalgia</td>
<td>Any stage of HIV</td>
<td>Pain and fatigue</td>
<td>Good</td>
<td>Standard measures (NSAIDS, etc.)</td>
</tr>
<tr>
<td>Myopathy (polymyositis)</td>
<td>Any stage of HIV</td>
<td>Proximal weakness</td>
<td>Good</td>
<td>Steroids, exercise as tolerated</td>
</tr>
<tr>
<td>AIDS wasting syndrome</td>
<td>Late AIDS</td>
<td>Proximal weakness</td>
<td>Poor</td>
<td>Supportive care, anabolic steroids, human growth hormone</td>
</tr>
<tr>
<td>Pain syndromes and neuropathies</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arthritis and arthralgias</td>
<td>Increased with decreased CD4 count</td>
<td>Psoriasis associated, Reiter's syndrome, reactive and HIV</td>
<td>Similar to idiopathic varieties</td>
<td>Similar to idiopathic varieties (NSAIDS, injections, joint preservation, etc.)</td>
</tr>
<tr>
<td>Distal symmetric neuropathy</td>
<td>Early or late</td>
<td>Distal numbness, pain, and decreased strength</td>
<td>Responds to symptomatic therapy</td>
<td>Tricyclics, neuroleptics, heat/cold, gait aids, orthotics</td>
</tr>
<tr>
<td>Segmental neuropathies</td>
<td>Related to drugs</td>
<td>Zoster, mononeuritis simplex/multiplex, vasculitis, etc.</td>
<td>Variable</td>
<td>Similar to that in non-HIV patients</td>
</tr>
<tr>
<td>Drug-induced neuropathies, vitamin deficiencies</td>
<td>Throughout AIDS</td>
<td>Didanosine, zidovudine, stavudine, vincristine, dapsone, rifampin, isoniazide, ethambutol</td>
<td>Good</td>
<td>Cessation or alterations of medications, vitamin supplementation, symptomatic management</td>
</tr>
<tr>
<td>Plantar KS</td>
<td></td>
<td>Characteristic lesions</td>
<td>Good</td>
<td>Radiation, custom insoles, etc.</td>
</tr>
<tr>
<td>Cognitive dysfunction</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HIV dementia</td>
<td>Late AIDS, but occasionally seen at presentation</td>
<td>Decreased memory, dysphagia, associated myelopathy</td>
<td>Poor</td>
<td>Antiretrovirals, memory book, adaptive strategies, decreased stimulation, etc.</td>
</tr>
<tr>
<td>Cryptococcal meningitis</td>
<td>Throughout AIDS</td>
<td>Fever, headache, meningismis, etc.</td>
<td>About 50% respond to treatment</td>
<td>Amphotericin B, followed by lifetime fluconazole</td>
</tr>
<tr>
<td>Condition</td>
<td>Stage of HIV/AIDS</td>
<td>Symptoms/Signs</td>
<td>Prognosis</td>
<td>Treatment/Management</td>
</tr>
<tr>
<td>------------------------------------------------</td>
<td>-------------------</td>
<td>-------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Neurosyphilis Encephalitis (herpesviruses)</td>
<td>Throughout AIDS</td>
<td>Nonspecific CNS, Headache, seizure, altered consciousness, focal signs</td>
<td>Fair, frequent relapse, May be self-limited, fair to good with antiviral treatment, Lesion specific</td>
<td>High-dose PCN, Acyclovir for HSV, ganciclovir or foscarnet for CMV, See below</td>
</tr>
<tr>
<td>Focal lesions (see below)</td>
<td>Throughout AIDS</td>
<td>Focal signs</td>
<td>Good</td>
<td>Proper diagnosis and correction of underlying problem</td>
</tr>
<tr>
<td>Metabolic/iatrogenic dementia</td>
<td>Throughout HIV</td>
<td>Nonspecific CNS</td>
<td>Good</td>
<td>Pyrimethamine and folic acid, plus sulfadiazine or clindamycin; prophylaxis when CD4 count is &lt;100</td>
</tr>
<tr>
<td>Focal brain disorders</td>
<td>Throughout AIDS</td>
<td>Focal signs, altered consciousness, ring-enhancing lesions</td>
<td>Good with treatment</td>
<td>Palliative radiation extends both quality and quantity of life</td>
</tr>
<tr>
<td>Cerebral toxoplasmosis</td>
<td>Throughout AIDS</td>
<td>Focal signs</td>
<td>Extreme poor</td>
<td>Anti-retrovirals</td>
</tr>
<tr>
<td>Encephalitis (herpesviruses)</td>
<td>Throughout AIDS</td>
<td>Headache, seizure, altered consciousness, focal signs</td>
<td>May be self-limited, fair to good with antiviral treatment, Lesion specific</td>
<td>Proper diagnosis and correction of underlying problem</td>
</tr>
<tr>
<td>Encephalitis (herpesviruses)</td>
<td>Throughout AIDS</td>
<td>Focal signs</td>
<td>Good</td>
<td>Pyrimethamine and folic acid, plus sulfadiazine or clindamycin; prophylaxis when CD4 count is &lt;100</td>
</tr>
<tr>
<td>PML Lymphoproliferative Disorders</td>
<td>Advanced AIDS</td>
<td>Focal signs, altered consciousness, occasional paraneoplastic signs</td>
<td>Very poor except in rare cases</td>
<td>Palliative radiation extends both quality and quantity of life</td>
</tr>
<tr>
<td>Primary CNS lymphoma</td>
<td>Advanced AIDS</td>
<td>Focal signs, altered consciousness, occasional paraneoplastic signs</td>
<td>Extremely poor</td>
<td>Anti-retrovirals</td>
</tr>
<tr>
<td>Other metastatic disease</td>
<td>Throughout AIDS</td>
<td>Similar to above</td>
<td>Lesion specific</td>
<td>Appropriate to lesion</td>
</tr>
<tr>
<td>CNS abscesses</td>
<td>Throughout AIDS</td>
<td>Similar to above</td>
<td>Lesion specific</td>
<td>Drainage of abscess, antimicrobials, antituberculin agents</td>
</tr>
<tr>
<td>CMV retinitis</td>
<td>Advanced AIDS</td>
<td>Visual field loss</td>
<td>Progressive</td>
<td>Ganciclovir and foscarnet</td>
</tr>
<tr>
<td>Retinitis and myelopathies</td>
<td>Early in HIV</td>
<td>Guillain-Barré, CIDP may be associated with nucleoside analogues</td>
<td>Very good</td>
<td>Same as in idiopathic forms</td>
</tr>
<tr>
<td>Inflammatory demyelinating polyneuropathies</td>
<td>Early in HIV</td>
<td>Flaccid paralysis</td>
<td>Arrested if treatment started quickly</td>
<td>Ganciclovir, foscarnet, rehabilitative interventions as in SCI, As in other forms of SCI</td>
</tr>
<tr>
<td>CMV Polyradiculopathy</td>
<td>Late AIDS</td>
<td>Progressive paraparesis</td>
<td>Poor recovery</td>
<td>As in other forms of SCI</td>
</tr>
<tr>
<td>Vacuolar myelopathy</td>
<td>Late AIDS</td>
<td>Tropical spastic paraparesis (HTLV-1), herpesviruses, meningitis, Pott's disease</td>
<td>Poor</td>
<td>As in other forms of SCI</td>
</tr>
<tr>
<td>Other myelitits</td>
<td>Throughout AIDS</td>
<td>Static, progressive or progressive/plateau</td>
<td>Variable</td>
<td>Corticosteroids, symptomatic management for exacerbations, Rehabilitative Interventions similar to spastic diplegic CP</td>
</tr>
<tr>
<td>Pediatric HIV/LIP</td>
<td>Common throughout AIDS</td>
<td>Nonspecific with variable course</td>
<td>Usually benign</td>
<td></td>
</tr>
<tr>
<td>HIV encephalopathy</td>
<td>Common and progressive with AIDS</td>
<td>Static, progressive or progressive/plateau</td>
<td>Variable</td>
<td></td>
</tr>
</tbody>
</table>

AIDS, acquired immunodeficiency syndrome; CIDP, chronic inflammatory demyelinating polyneuropathy; CMV, cytomegalovirus; CNS, central nervous system; HIV, human immunodeficiency virus; HSV, herpes simplex virus; HTLV, human T-cell lymphotrophic virus; LIP, lymphoid interstitial pneumonitis; MAC, Mycobacterium avium complex; mTB, Mycobacterium tuberculosis; NSAIDS, nonsteroidal anti-inflammatory drugs; PCN, penicillin; PCP, Pneumocystis jiroveci pneumonia; PML, primary multifocal leukoencephalopathy; SCI, spinal cord injury; TMPX, trimethoprim/sulfamethoxazole.
Weakness and Fatigue in HIV

Persons with HIV are often fatigued for a variety of reasons, including fibromyalgia, pulmonary dysfunction, anemia, encephalopathy, endocrine dysfunction, myopathies, cardiomyopathy, psychiatric disorders, and depression. Fibromyalgia is commonly seen, and the response to traditional measures is generally good. Fatigue may also result from the effects of ART. The importance of preventative measures cannot be overemphasized with this population, including the use of light exercise to tolerance, early mobilization after an acute illness, and other measures to prevent common complications of immobility. Psychostimulants and anabolic steroids may also play a role (2,112–129).

HIV-Related Cardiovascular Disease

Large cohort studies have demonstrated an increased risk of cardiovascular disease in HIV patients that is independent of other risk factors. This is manifest in increased rates of myocardial infarction and CVA in patients that may be related in part to chronic inflammation caused by HIV. Lipid abnormalities caused by some of the antiretroviral agents may also play a role (130).

Although the mechanisms involved are not clear, HIV-related cardiomyopathy is common (131). Other causes of cardiac dysfunction include pericardial disease, infectious and noninfectious endocarditis, metastatic KS, and lymphoma. When symptomatic, management is similar to that of congestive heart failure of other causes, with the use of vasodilators and diuretics. Exercise is instituted as tolerated in a similar fashion.

Chronic HIV infection and a number of the antiretroviral agents can also result in fat redistributions that can cause cosmetic, psychologic, and physical disability. Lypodystrophy appears as a combination of lypohypertrophy with abdominal, visceral adipose accumulation, dorsal cervical fat pad formation and gynecomastia and lypoatrophy with loss of subcutaneous fat from the extremities and face. Development of lypodystrophy is associated with increased rates of depression, nonadherence to medications, and metabolic syndrome (132,133).

HIV-Related Metabolic Abnormalities

Some of the antiretroviral agents cause a decrease in insulin sensitivity that in turn may lead to increased rates of diabetes and metabolic syndrome (98–100,133).

Both HIV and some of the antiretroviral agents seem to lead to increased rates of osteopenia and osteoporosis in HIV infected individuals. This can lead to disability from pain syndromes and fractures (134).

An HIV-related nephropathy has been described (135). It is characterized by fluid and electrolyte abnormalities, acid-base imbalance, and proteinuria. Acute renal failure is usually the result of hypovolemia and ischemia or iatrogenic factors from the toxic effects of diagnostic agents and medications.

Myopathies

HIV myopathy may present at any stage of infection, usually as proximal muscle weakness most prominently in the thighs, particularly after exertion (136–138). Myalgia occurs in 25% to 50% of cases (139). Weight loss is common. Creatinine kinase is usually elevated, and EMG demonstrates abundant spontaneous activity with short-duration, low-amplitude polyphasic motor unit action potentials (140). Confirmation can usually be obtained from muscle biopsy, which demonstrates focal necrosis with regeneration. Prednisone has been reported to improve strength (141). The role of exercise is thought to be similar to that in non–HIV-associated myopathies.

Mitochondrial disease and associated myopathies are common in association with ART and may even persist in newborns whose mothers have undergone therapy for HIV (142–144). Nemaline myopathy has also been described and is clinically and electrophysiologically indistinguishable from HIV myopathy. Other myopathies may be toxic, autoimmune, or related to Coxsacki virus or human T-cell lymphotropic virus type I (HTLV-I) coinfection (145). These must be distinguished from the AIDS-wasting syndrome, which presents insidiously, late in the course of disease, with symmetrical proximal muscle weakness. Beneficial effects from anabolic steroids, growth hormone and, particularly, exercise, have been described (117,146–149).

Pain Syndromes and Neuropathies

Arthritic Disorders

Arthralgias arising from a variety of ill-defined causes are common in association with HIV and with antiretroviral treatment (142,150). Arthritis is also relatively frequent, with psoriasis-associated arthritis, Reiter’s syndrome, and reactive arthritis all being described (151–155). Septic arthritis, however, is a fairly rare complication of AIDS (156). The prognosis and management is comparable to that of idiopathic arthritides.

Avascular Necrosis

Numerous reports have surfaced in recent years of avascular necrosis (AVN) of the hips in patients with HIV (157–161). Early speculation that this is associated with the use of protease inhibitors has proven to be false (162). Some reports have demonstrated a positive correlation with steroid use (163,164). Of more concern is that radiologic evidence of AVN has been demonstrated in individuals with HIV who are asymptomatic for hip pain or dysfunction (165). It is not known how many of these if any will develop pain and impairment in mobility as a result of AVN; however, the potential for collapse of the femoral head must be considered. Presumably, those who are symptomatic can be managed similar to those with AVN from other causes.

Other Rheumatologic Disorders

Vasculitis is an unusual finding in HIV, with a reported incidence of 0.4% to 1% (166). Unfortunately, the prognosis is generally poor. Polyarteritis nodosa and an adult Kawasaki-like syndrome have also been described (167,168).

Distal Symmetric Neuropathy

Peripheral neuropathies are frequently seen in HIV, with a distal symmetrical polyneuropathy being demonstrable on EMG in 35% of AIDS cases (169–176). Symptomatic in about 18% of
patients with AIDS, its incidence increases with increasing viral load, and pathologic evidence can be found in nearly all cases on sural nerve biopsy or at autopsy. More common in adults, it has also been reported in children (177,178). It remains essential, however, to rule out treatable causes of neuropathy.

On biopsy, a “dying back” of distal fibers is seen, with loss of both myelinated and nonmyelinated fibers. In some cases, perivascular mononuclear inflammatory infiltrates are seen. Electromyographically, sensory nerve action potentials are reduced in amplitude initially and F waves are abnormal. In time, denervation potentials become evident in distal muscle groups and long-duration polyphasic motor unit action potentials, decreased recruitment, and increased amplitude may be seen (140,179).

The treatment of symptomatic peripheral neuropathies has included everything from traditional medications used in treating painful neuropathies to applied vibration and acupuncture (180–186).

Other Sensorimotor Neuropathies
Drug-induced neuropathies commonly result from the antiretroviral drugs didanosine, zalcitabine, and stavudine. Other drugs commonly used in AIDS patients that can cause neuropathies include vincristine, dapsone, rifampin, isoniazide, and ethambutol (187–189). Mononeuropathy simplex and multiplex may occur early in HIV as a limited peripheral neuropathy that must be distinguished from segmental herpes zoster, also a common cause of painful neuropathy (190). A multifocal neuropathy associated with severe weakness and disability, and often with fever and cachexia, also may be seen in advanced AIDS. EMG is consistent with both axonal loss and demyelination (140). Motor neuron disease has rarely been reported in patients with HIV, and there is little direct evidence of an association between the two (191).

Autonomic Neuropathy
Autonomic neuropathy is commonly found in association with HIV infection, particularly in advanced AIDS (192–196). Patients may experience orthostatic hypotension, syncope, impotence, decreased sweating, impaired gastric motility, and diarrhea. As always, it is important to rule out iatrogenic causes from the use of common medications.

Cognitive Dysfunction in HIV
HIV-associated Dementia
The most common cause of cognitive dysfunction is HIV-associated dementia, also known as the AIDS dementia complex, subacute encephalitis, and HIV encephalopathy (197–200). Although it is usually not seen until advanced disease, it may be present in 3% to 7% of cases at the time of diagnosis and will eventually be found in 15% to 20% of those with advanced disease, making it the leading cause of dementia in young adults (201). Ironically, the incidence in individuals undergoing ART with CD4 counts greater than 200 may actually be increasing (202). Although poorly understood, monocytes and dopaminergic pathways have been implicated (203–208). Alcohol and cocaine may play a role in hastening the progression of dementia (209). Although electrophysiologic abnormalities can be demonstrated on evoked potential testing in a majority of individuals with HIV and abnormalities are evident in functional imaging studies (210), cognitive dysfunction does not usually become evident prior to the development of AIDS, and the existence of subtle cognitive changes in early disease remains controversial (211,212).

Cryptococcal Meningitis
Cryptococcus is an important fungal pathogen in patients with CD4 counts under 200 cells/µL (213). The incidence has declined since the advent of ART (214). Infection most often presents with meningitis, and symptoms include fever, headache, neck stiffness, and memory loss. Patients may show signs of lethargy, confusion, meningismis, or cranial nerve palsies. Signs and symptoms in AIDS patients, however, are often mild (215). Cryptococcal antigen may be detected in the cerebral spinal fluid (CSF) or serum for diagnosis. Initial treatment is with amphotericin B or fluconazole followed by fluconazole maintenance (216–219). Prophylaxis with low-dose fluconazole may be of use in high-risk patients (220).

Other Causes of Cognitive Dysfunction
Treponema pallidum may cause chronic meningitis, neuropathy, or dementia when infecting the CNS of AIDS patients (neurosyphilis) (221,222). CSF pleocytosis is often seen, and a CSF Venereal Disease Research Laboratory (VDRL) test may be positive, but sensitivity is low. Treatment is usually with intravenous penicillin G, but relapses are common.

Varicella zoster, a common cause of pneumonia and skin eruptions in AIDS patients, also results in a variety of neurologic syndromes, including painful nerve palsies, transverse myelitis, ascending myelitis, and leukoencephalopathy (223–225). In contrast to HSV encephalitis, herpes zoster encephalitis rarely causes focal neurologic signs, deep alterations in consciousness, or seizures. The usual duration of neurologic symptoms is around 16 days, and the outcome is generally better than that with the other CNS infections. PCR assay for HZV may aid in diagnosis and management. Acute mononucleosis caused by Epstein-Barr virus is also common, often leading to encephalitis, acute cerebellar syndrome, transverse or ascending myelitis, or acute psychosis (226).

Other causes of cognitive dysfunction include CMV encephalitis, aseptic meningitis, and cerebral vasculitis (166,190,227–229). Cognitive dysfunction also may result from disorders that cause focal CNS syndromes. As always, with any dementia workup, it is essential to rule out iatrogenic and other potentially treatable causes.

Focal Brain Lesions found in HIV
The focal deficits seen in association with AIDS are related directly to the areas of the brain affected and may include blindness, hemiparesis, ataxia, aphasia, dysarthria, and cranial nerve deficits (190,230–234). Movement disorders, including hemichorea-ballismus, segmental myoclonus, postural tremor, parkinsonism, and dystonia, also may be found.
Toxoplasmosis

Cerebral toxoplasmosis is the most common CNS mass lesion in AIDS, occurring in 3% to 40% of cases (190,235–237). Without prophylaxis, 30% of previously toxoplasma-seropositive AIDS patients may relapse when the CD4 count drops below 50. Most patients present with focal or nonfocal neurologic signs produced by space-occupying lesions. Headaches, confusion, lethargy, and seizures are common. Contrast-enhanced head CT or magnetic resonance imaging (MRI) reveals single or multiple ring-enhancing lesions that may be distinguished from lymphoma by biopsy or radiolabeled thallium uptake (233,238–240).

A common diagnostic strategy is to initiate antitoxoplasma therapy and observe for clinical response, which is often dramatic. Treatment regimens include pyrimethamine and folic acid plus either sulfadiazine or clindamycin. Patients with a positive antitoxoplasma immunoglobulin G antibody should take prophylaxis when their CD4 count is less than 100 cells/μL (241). Trimethoprim/sulfa, often used for PCP prophylaxis, is effective, as is clindamycin plus pyrimethamine.

Progressive Multifocal Leukoencephalopathy

Progressive multifocal leukoencephalopathy (PML) is a white matter disease associated with infection by the JC virus (230). It occurs in approximately 4% of AIDS patients. Single or multifocal lesions confined to the white matter without the presence of a mass effect or contrast enhancement may be seen on MRI. Diagnosis can be made based on radiographic appearance or more definitively by biopsy (242). The median survival is of the order of 2 to 4 months, although in approximately 10% of cases a more benign course is seen, and spontaneous remissions may even occur after ART is added (243–245).

CNS Lymphoma

CNS lymphoma is associated with late-stage AIDS and CD4 counts of less than 100 cells/μL. It may present as a mass lesion with confusion, headache, memory loss, or focal neurologic signs. Untreated, it usually progresses to death in 1 month or less. Palliative radiation can temporarily improve symptoms and extend survival, with an increased quality of life in 75% of cases (246). Associated paraneoplastic syndromes are common. Rarely, other malignancies and metastatic lesions may be seen in the brain, particularly from KS (247).

CMV Retinopathy

CMV is one of the most common pathogens in AIDS patients, with more than 90% of patients showing evidence of past exposure and up to 40% of those with advanced disease developing active CMV disease (248). About 85% of CMV disease in AIDS patients presents as retinitis, which usually affects those with CD4 counts less than 50 cells/μL. Productive infection of the retina can cause full thickness necrosis, although patients may be asymptomatic and gradual visual loss may be unnoticed until vision is significantly impaired. Ophthalmologic examination shows characteristic perivascular yellow-white lesions that are frequently associated with retinal hemorrhage (249). These must be distinguished from cotton wool spots, which are white, fluffy, superficial infarcts of the nerve fiber layer that usually spontaneously regress and do not lead to blindness.

The mainstay of treatment has typically involved a combination of ganciclovir, a nucleoside analogue, and foscarnet, a pyrophosphate analogue (249–251). Cidofovir, a newer agent with a long half-life, can suppress CMV up to 3 weeks after a single administration with probenacid (252,253). Intravitreal injection of ganciclovir and cidofovir as well as intraocular implants of timed-release devices with ganciclovir or foscarnet have also been used with success (254–256). Complications are common with all these drugs (see Table 47-2). Although the incidence is decreasing as a result of ART, those infected continue to have a poor prognosis, and drug resistance is common (250,251,257,258). Eventual progression to blindness can occur, making early institution of visual retraining and blind rehabilitation particularly important.

Other Causes of Focal CNS Syndromes

HSV-I may infect the brain in AIDS patients, causing an encephalitis characterized by the acute onset of headache, behavioral changes with focal neurologic signs, seizures, and coma. Although an increased incidence of thromboembolic cerebrovascular events has been reported, it is not clear that the incidence is greater than that of the general population (259–262). This is in contrast to cerebral vasculitis, which does have an association with AIDS. Primary CNS angiitis is a rare and usually fatal disorder that has a predilection for small cortical vessels (229,263,264). Bacterial and fungal abscesses may also cause focal neurologic deficits (265). TB is of particular concern and may require biopsy for diagnosis. Antituberculin therapy is often initiated in those with a positive PPD (Purified Protein Derivative tuberculin test) and a new focal lesion when the diagnosis is uncertain (266).

Radiculopathies and Myelopathies

Inflammatory Demyelinating Neuropathies

HIV is a major cause of inflammatory demyelinating polyneuropathies and must always be considered in the differential (187,267–269). Guillain-Barré syndrome, or acute inflammatory demyelinating polyneuropathy (AIDP), occurs early, often around the time of seroconversion, and is clinically and histopathologically indistinguishable from the idiopathic form. It is managed similarly. In some cases, chronic inflammatory demyelinating polyneuropathy (CIDP) or relapsing polyneuropathy may be seen.

Nerve conduction studies are significant for prolonged sensory and motor latencies with patchy slowing to less than 60% to 70% of normal (140). Marked reduction in motor action potential amplitude or even motor conduction block may be indicative of focal demyelination. F wave and H reflexes are often markedly delayed or absent. The EMG needle exam usually fails to show abnormal spontaneous activity, although denervation potentials with diminished recruitment may sometimes be seen in weak distal muscles.
**CMV Polyradiculopathy**

CMV causes a common and debilitating peripheral neuropathy in which CMV infects the nerve roots and ganglia (268,270). Patients present with back and lower extremity sensory changes and weakness that may rapidly progress to flaccid paralysis (267,271). It can be rapidly fatal unless treatment is started within the first 24 to 48 hours (227,268,270). Treatment is effective in halting progression, although patients are often left with severe deficits that may require rehabilitation.

Findings include loss of lower extremity deep tendon reflexes, saddle anesthesia, and urinary retention, whereas upper extremity strength is preserved. Diffuse spontaneous activity is seen on EMG, with decreased recruitment but near-normal conduction velocities (140). The CSF often has a pleocytosis with neutrophils and may show elevated protein and inclusion cells that may stain positive for CMV antigens (227,267). Improvement upon treatment with ganciclovir has been described (174,267,272,273). Foscarnet also can be effective.

**Vacuolar Myelopathy**

Vacuolar myelopathy is the most common cause of spinal cord dysfunction in HIV, being found in 11% to 22% of AIDS cases and demonstrable in up to 40% of cases at autopsy (190,274,275). It is strongly associated with HIV dementia and shares a virtually identical histopathology. Patients present with progressive paraparesis, ataxia, posterior column sensory loss, spasticity, and neurogenic bowel and bladder.

Spinal cord dysfunction may also result from viral myelitis secondary to varicella zoster virus, HSV, CMV, and HTLV-I (276). Spinal cord disorders may also be seen with tuberculin involvement (Pott's disease), multiple sclerosis, and aseptic, cryptococcal and lymphomatous meningitis (275,277).

**PEDIATRIC HIV**

Pediatric cases of HIV differ significantly from their adult counterparts (278–280). More than 80% of cases of HIV infection in children are the result of perinatal transmission, with transmission from blood products being less frequent. The most effective strategy to minimize vertical transmission is prophylactic treatment of the mother as well as the newborn (39,281). Because of transplacental transmission of antibodies, however, most infants of HIV-positive mothers will initially be seropositive at birth, with signs of infection developing in approximately 80% of those with disease at a median age of 5 months. Significant advances in medical management over the past 5 years, however, have resulted in a substantial decrease in the progression to AIDS and an improvement in survival (282).

** Opportunistic Infections**

In contrast to adults, opportunistic infections of the nervous system are extremely uncommon (283). AIDS-defining criteria in children include recurrent bacterial infections, opportunistic infections, lymphoid interstitial pneumonitis (LIP), malignancies, cardiomyopathies, hepatitis, nephropathy, and encephalopathy. Infants with AIDS often present with hepatosplenomegaly, lymphadenopathy, diarrhea, oral candidiasis, parotid enlargement, delayed developmental milestones, and failure to thrive.

Over 80% of HIV-infected children eventually develop either acute or chronic lung diseases. While most are acute infectious processes, chronic diseases such as LIP may occur in 30% to 50% (284–286). The etiology and natural history of LIP are incompletely understood. Although some patients are asymptomatic with infiltrates on chest x-ray, others experience a slowly progressive course with exacerbations caused by lung infections, and some may progress to chronic respiratory decompensation and hypoxemia. HIV-infected children with LIP are believed to have a better long-term prognosis than those with other pulmonary infections. Corticosteroids are used to suppress the lymphocytic infiltrate.

**HIV Encephalopathy**

The most common neurologic disorder seen in children is HIV encephalopathy (287–290). This may be static, progressive, or progressive with a plateau. Static encephalopathy is characterized by delayed acquisition of developmental skills and is seen in 25% of those with CNS deficits. With progressive encephalopathy, on the other hand, the pace of development becomes progressively slower and may halt altogether, or show signs of regression. The clinical picture is not unlike that of spastic diplegic cerebral palsy, but with progressive deficits.

Other children may demonstrate more subtle neurodevelopmental deficits, including easy fatigability, mild depression, expressive language deficits, and mild regression in school performance. The response to ARTs has been reported to be good. Uncommon neurologic disorders that may be encountered in children include CNS lymphoma, corticospinal tract degeneration, and, rarely, myelopathy, myopathy, and peripheral neuropathy (177).

With the advent of ART, many children born with HIV may live relatively normal lives, although this is not true in much of the developing world, where access to therapy remains limited (280,282,291–293). ART has been shown to reduce viral load in children just as in adults, and CNS pathology has been shown to be reversible with antiretroviral treatment as well. More importantly, neuropsychiatric indicators have been demonstrated to improve significantly with the institution of treatment. As of this writing, a number of children born with HIV are now into their teens; however, the long-term adverse effects of antiretroviral treatment will only become clear over time.

**PSYCHOSOCIAL AND VOCATIONAL ISSUES**

HIV remains one of the most stigmatized diseases of our time. As in oncologic rehabilitation, patients may be faced with issues of death and dying but with appropriate therapy, hope, and long-term survival. Discrimination remains a particular
problem due to the infectious nature of the disease and to public misperceptions and prejudice. The health of the caregiver, if there is one, is of particular concern when developing rehabilitation strategies. As always, rehabilitation professionals must deal with the patient in the context of their surroundings and their external support network.

Before the onset of symptoms, 70% to 90% of people with HIV are still employed (294). Even in the first year after the onset of symptoms, fewer than one-third leave the workforce. People with HIV commonly experience workplace discrimination and, with advancing symptoms, may experience a diminished work capacity. Significant controversy exists regarding persons with HIV in the health care field. Those who are forced to leave the workforce face a potential loss of health insurance. In all other respects, vocational rehabilitation of the HIV patient is similar to that of the patient with cancer or multiple sclerosis, but always keeping in mind the potential for full recovery and long-term health (295,296).

THE FUTURE OF HIV REHABILITATION

The management of HIV is one of the most rapidly changing fields of medicine, with entirely new treatment strategies often becoming available in a period of months rather than years. By the time this is published, it is likely that new medications not discussed here will be in common usage. The complexity of HIV rehabilitation continues to evolve rapidly, and the future of HIV rehabilitation is thus largely unknown. As a cure still does not seem to be imminent; however, it is reasonable to expect a need for rehabilitation interventions for the foreseeable future.

Most HIV strategists currently view HIV as being managed as a chronic disease similar to diabetes mellitus. Although not curable, its progression will be held in check by medical therapies. This comparison to diabetes may be particularly apt in the case of rehabilitation. Although the most serious morbidity and mortality of diabetes is avoidable, even with stringent control of blood glucose, amputation, blindness, and peripheral neuropathy remain common sequelae that often respond to rehabilitation interventions. While new HIV treatments are indeed remarkable in their ability to arrest development of functional impairments and disability, they introduce toxicities of their own that may lead to yet a new set of impairments.

It has now become apparent that long-term viral suppression with antiretrovirals is achievable for most HIV patients, but that it requires an incredible 95% adherence to treat drug-resistant strains. Much of the effort in recent years has been expended in encouraging and enabling this high-level adherence. Adherence is an integral part of HIV treatment programs.

In the last edition of this chapter, we speculated that rehabilitation professionals might face a future practice consisting of otherwise healthy individuals with HIV and a variety of disabling conditions such as Guillain-Barré syndrome, polymyositis, distal symmetric polyneuropathy, HIV dementia, and vacuolar myelopathy. Although the numbers remain limited, many have in fact noted an increase in potentially disabling conditions in otherwise healthy individuals with HIV, leading some to call for more emphasis on rehabilitation (3,298–301). Remarkably, exercise has been shown to play a significant role in improving not only the aerobic capacity but also immune system function in individuals with HIV (302–307). The role of rehabilitation may thus be critical not only in keeping these people functional in their communities, but in maintaining their health as well.

REFERENCES


177. Navia BA. Clinical and biologic features of the AIDS dementia complex [review] [125 refs], *Neuropsychiatry Neuropsychol Behav Neurol*. 1997;5(3):581–592.


279. Mintz M. Neurological and developmental problems in pediatric HIV infection [review] [95 refs]. *J Nutr*. 1996;126(10 suppl):266S–267S.


PART IV

Secondary Conditions and Complications
While physical activity and exercise are generally well-accepted concepts in healthy persons, the consequences of inactivity and bed rest are less well understood by clinicians and the general public. The deleterious effects of immobility and physical inactivity are common and may affect multiple body systems and the stealthy onset of these effects may minimize the awareness of the dangers and the timely response for prevention and treatment.

Bed rest and immobilization were widely used before 1950 in the management of trauma and acute and chronic illness, before the physiologic effects were well understood. It was generally assumed that rest fostered healing of the affected part of the body. What was not appreciated was that immobility and inactivity could be harmful to the unaffected parts of the body. For example, the immobilization of long bones with a rigid cast has a beneficial effect on bone healing after fractures. However, it may also result in undesirable effects, such as joint contracture and atrophy of the healthy muscles and bones.

Clinical studies on enforced bed rest in normal subjects and on astronauts in microgravity conditions (in which their bodies rest from the effects of gravity) have shown significant undesirable effects that may override the therapeutic effects of bed rest in subacute and chronic conditions, impacting complexity and cost of medical treatment as well as functional outcome.

A review of randomized controlled trials on the effects of bed rest and immobilization did not demonstrate improvement or better outcome of primary medical conditions for those on extended periods of bed rest. In many cases, worsening occurred if early mobilization was not provided (1). Persons who are chronically sick, aged, or disabled are particularly susceptible to the adverse effects of immobility. A patient with motor neuron disease and its accompanying limb weakness or spasticity would be expected to develop the same musculoskeletal complications but at a much accelerated rate. The healthy subject on prolonged bed rest may only show some degree of atrophy and weakness; the neurologically impaired subject will likely also lose a significant amount of independent functioning. Many of these complications could be easily prevented or, if they occur, easily treated once they are recognized. Therefore, the prevention of such complications should be one of the basic principles of any rehabilitation management plan.

Negative effects of immobility or inactivity are rarely confined to only one body system (Table 48-1). Immobility reduces the functional reserve of the musculoskeletal system, resulting in weakness, atrophy, and poor endurance. Metabolic activity and oxygen utilization in the muscle are reduced, degrading the functional capacity of the cardiovascular system as well (Fig. 48-1). In addition, postural hypotension and deep venous thrombosis (DVT) are commonly encountered in bedridden patients. Immobilization osteoporosis is yet another complication that has been well documented in the studies of astronauts and individuals exposed to prolonged bed rest. Over time, clinical experience has dictated a move toward earlier mobilization and provision of functional training with a resulting decrease in the length of hospitalization and in the incidence of major morbidity associated with prolonged inactivity (2).

The most deleterious effects of inactivity can be grouped together under the general term "deconditioning," which is defined as reduced functional capacity of musculoskeletal and other body systems. It should be considered a distinct diagnosis from the original condition that has led to a curtailment of normal physical functions (3). This chapter describes widespread effects of immobility and physical inactivity, reviews the evidence-based therapeutic and prophylactic approaches to counteract these complications, and redirects attention to the benefits of physical activity and exercises in maintaining a good health and independence.

**MUSCULOSKELETAL EFFECTS OF IMMOBILITY AND INACTIVITY**

Moving the body and limbs freely in the environment is a principal physical function requiring that the muscles, nerves, bones, and joints be in an optimal physiologic state. Disuse weakness and reduction of free joint motion can initially cause minimal functional limitations that can be easily overlooked or
motion (ROM) may seem trivial; however, neglect of these simple factors can be responsible for prolonging hospital stays, increasing the use of health care resources, and prolonging dependency in mobility and ADL (5).

Three main types of adverse effects from inactivity are encountered in the musculoskeletal system: muscle atrophy and weakness, joint contracture, and immobilization osteoporosis (see Table 48-1).

### INACTIVITY OF SKELETAL MUSCLE

#### Physiological Impairments

**Disuse Atrophy**

Decrease in the size of muscle fibers and reduction of muscle mass are the hallmark of muscle atrophy. In a lower motor neuron lesion, the atrophy is regional and related to the particular nerve or root. Atrophy associated with muscle disease is more pronounced in the proximal muscles. The atrophy of disuse is generalized or localized to the immobilized limb(s) and more prominent in the antigravity muscles. It is a consequence of the limited physical activity and musculoskeletal loading that occurs during immobilization, immobility, bed rest, or exposure to microgravity during space flight. As a general rule, increased muscular activity leads to hypertrophy, whereas limited physical activity leads to disuse atrophy and weakness. The changes in muscle fiber size either in atrophy or in hypertrophy and phenotype changes of muscle fibers are the results of remarkable muscle plasticity and adaptability to external physical demands to generate adequate or peak contraction and endurance.

Disuse atrophy is defined as an alteration of metabolism and muscle cell homeostasis in response to muscle inactivity. Recent studies indicate that muscle protein synthesis as well as whole body protein production is significantly reduced during immobility and is considered the main contributor to muscle atrophy. The rate of muscle wasting during bed rest is slow during the first 2 days but becomes rapid thereafter. By 10 days, it reaches 50% of eventual muscle weight loss. Similarly, muscle protein synthesis is reduced to 50% of the baseline level at 14 days of immobilization and then gradually tapers off to reach a new steady state (6). Reeves et al. (7) studied skeletal muscle changes in healthy men exposed to microgravity condition and found a decrease in calf muscle volume by 29%, resting fascicle length of medial head of gastrocnemius and their pennation angle by 10% and 13%, respectively, and physiological cross-sectional area (PCSA) by 22%. In disuse atrophy, especially if muscle is immobilized in a shortened position, that is flexion, the number of sarcomeres in series is reduced as a result of diminished chronic stretch and adaptation of the muscle to a new resting length. By contrast, immobilization in an elongated muscle resting position or during musculoskeletal growth increases the number of sarcomeres in series (8). In addition, the number of sarcomeres in parallel is reduced contributing to the reduction of muscle fiber PCSA (muscle volume/fascicle length) (9). From data obtained on the

---

**TABLE 48.1 Adverse Effects of Immobility and Inactivity**

<table>
<thead>
<tr>
<th>System(s)</th>
<th>Effect(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Musculoskeletal</td>
<td>Muscle weakness, fatigue, and atrophy</td>
</tr>
<tr>
<td></td>
<td>Muscle and joint contractures</td>
</tr>
<tr>
<td></td>
<td>Muscle stiffness and pain</td>
</tr>
<tr>
<td></td>
<td>Osteoporosis</td>
</tr>
<tr>
<td></td>
<td>Hypocalcemia</td>
</tr>
<tr>
<td>Cardiovascular and pulmonary</td>
<td>Redistribution of body fluids</td>
</tr>
<tr>
<td></td>
<td>Dehydration</td>
</tr>
<tr>
<td></td>
<td>Orthostatic intolerance</td>
</tr>
<tr>
<td></td>
<td>Reduction of cardiopulmonary capacity</td>
</tr>
<tr>
<td></td>
<td>Reduction of VO_{2max}</td>
</tr>
<tr>
<td></td>
<td>Elimination of bronchial secretions</td>
</tr>
<tr>
<td></td>
<td>Hypostatic pneumonia</td>
</tr>
<tr>
<td>Genitourinary and gastrointestinal</td>
<td>Urinary stasis, stones, and urinary infections</td>
</tr>
<tr>
<td></td>
<td>Loss of appetite</td>
</tr>
<tr>
<td></td>
<td>Constipation</td>
</tr>
<tr>
<td>Metabolic and endocrine</td>
<td>Glucose intolerance</td>
</tr>
<tr>
<td></td>
<td>Electrolyte alterations</td>
</tr>
<tr>
<td></td>
<td>Increased parathyroid hormone production</td>
</tr>
<tr>
<td></td>
<td>Other hormone alterations</td>
</tr>
<tr>
<td>Immune system</td>
<td>Impaired wound healing</td>
</tr>
<tr>
<td></td>
<td>Reduction in cellular immunity</td>
</tr>
<tr>
<td></td>
<td>Resistance to infection reduced</td>
</tr>
<tr>
<td></td>
<td>Anti-inflammatory suppression reduced</td>
</tr>
<tr>
<td>Cognitive and behavioral</td>
<td>Sensory deprivation</td>
</tr>
<tr>
<td></td>
<td>Confusion and disorientation</td>
</tr>
<tr>
<td></td>
<td>Anxiety and depression, memory</td>
</tr>
<tr>
<td></td>
<td>Decrease in intellectual capacity</td>
</tr>
<tr>
<td></td>
<td>Impaired balance and coordination</td>
</tr>
<tr>
<td>Cellular/Genetic</td>
<td>Diminished gene expression</td>
</tr>
<tr>
<td></td>
<td>Mitochondrial dysfunction</td>
</tr>
</tbody>
</table>

neglected. However, advanced contractures and disuse weakness can cause a loss of mobility and decrease in activities of daily living (ADL) functions (4).

For the neurologically impaired or multiple trauma victim, considerations such as preserving functional range of

---

**FIGURE 48-1.** Inactivity, immobility, and prolonged bed rest influence total body functioning.
myostatin-immunoreactive protein level increases by 12%. It is
lean body mass declines by an average of 2.2 kg, and plasma
increased during bed rest. After 25 days of bed rest, the total
factor-beta protein that inhibits muscle synthesis and is
important role in disuse atrophy (15). Myostatin is a growth
fiber wound–mediated fibroblast growth factor may play an
factors and prevent muscle fiber atrophy, indicating that myo-
cises during bed rest significantly increase the level of these
tional to the reduction in muscle fiber size. Resistance exer-
injury) are both reduced during bed rest, and this is propor-
tional to the reduction in muscle fiber size. Resistance exer-
cises during bed rest significantly increase the level of these
ners and prevent muscle fiber atrophy, indicating that myo-
ner may lose as much as 12 g/day. Urinary excretion of creatine
nisms, such as loss of appetite and reduced intestinal absorption
degradation may become more prominent than decrease in pro-
tein synthesis. This may be aggravated by gastrointestinal mecha-
nisms, such as loss of appetite and reduced intestinal absorption
protein. Although daily loss of nitrogen for an immobilized
healthy person may reach 2 g/day, a nutritionally depleted per-
son may lose as much as 12 g/day. Urinary excretion of creatine
is minimal under normal conditions, except in pregnancy and
infancy. The excretion of creatine is much greater in starvation,
diabetes, muscular dystrophy, hyperthyroidism, and fever, as
well as during immobility. Prolonged bed rest and weightlessness
causes a significant increase in the excretion of both creatine and
creatinine, the mechanism which is not well understood (18).
The loss of strength associated with disuse atrophy is more prominent in the lower limbs than in the upper limbs. Loss of muscle power during immobilization reaches −20% to −44% in knee flexors and extensors, comparing to an insignificant loss in the upper limbs of −5%. The decrease in peak muscle tension of knee flexors and extensors ranges from −19% to −26%, far more than the reduction in cross-sectional area of these muscles (−7%), indicating that loss of strength and power is proportionally greater than reduction in size of the respective muscles (23). The major contributors to the loss of strength and endurance in persons with disuse atrophy are the reduced number of myofibrils per fiber volume, reduction in size and number of mitochondria, muscle fiber nuclei, and reduction of sarcomeres in parallel (10).

Another aspect of disuse weakness is reduction of maximal instantaneous muscular power. It can be reduced to a greater extent than the peak muscle strength of one repetition maximum in bed rest subjects. The loss of strength has been well documented in immobility; however, the loss of explosive muscle power as measured by the maximal jump with both feet on force plate was only recently studied. After 45 days of bed rest, instantaneous muscular power was reduced 24%, and recovery required one and half months of remobilization (24). These findings indicate that specific training for instantaneous power should be considered during space flight or prolonged bed rest to preserve functions such as standing up, climbing, and performing transfer functions.

Decline in muscle twitch and tetanic tension parallels the decline in muscle strength. In rats, maximal tension obtained by electrical stimulation in the soleus muscle declined significantly after 6 weeks of immobilization (25). Changes in contractile forces of immobilized muscle are the result of diminished levels of myofibrillar proteins and also due to a reduction of sarcoplasmic reticulum calcium ion uptake, but not the rate of its release (26).

### Loss of Endurance

Multiple studies have demonstrated that prolonged inactivity causes a significant and progressive reduction in muscle endurance. Unexercised muscle demonstrates a reduction of adenosine triphosphate (ATP) and glycogen storage sites and rapid depletion of them after resumption of activity. Reduction of muscle protein synthesis and oxidative enzyme function, and premature anaerobic energy production with rapid accumulation of lactic acid, are important factors leading to fatigability and reduced endurance (27).

Metabolic and enzymatic alterations in unexercised muscle result from reduced demand for oxygen and reduction in the blood supply. Initially after bed rest, succinate dehydrogenase enzyme activity per muscle fiber increases, but later, the overall amount is reduced (28). Oxidative enzyme activity and content as well as the number and size of mitochondria are all reduced during immobility (29). Unexercised muscle also shows a decreased ability to utilize fatty acids when compared with trained subjects.

Ferretti et al. studied peripheral and central factors that contribute to the decline of VO_{2max} after 42 days of bed rest. VO_{2max} was found reduced by 16%, cardiac output by 30%, and oxygen delivery by 40%, which parallels the reduction in muscle cross-sectional area of 17%, volume density of mitochondria by 16%, and total mitochondria volume in the magnitude of 28%. Oxidative enzyme activity falls by 11%. These decrements indicate a significant contribution of both peripheral and central factors in causing the decline in VO_{2max} after immobility, confirming a close interrelationship between the muscular and the cardiovascular system (27).

The following sequence of events transpires in the deconditioning process. Prolonged reduction of muscle repetitive contractions below 50% of maximum alters muscle protein synthesis and decreases glycogen and ATP storage, causing a reduction of oxidative enzymes, mitochondrial function, and microvascular circulation, impacting muscle metabolic activity. As a result, the oxygen supply is attenuated, and the extraction of oxygen from blood is diminished, further negatively affecting VO_{2max} and cardiovascular reserve. The loss of muscle mass leads to reduction of muscle strength and endurance, reducing muscle blood flow, red blood cell delivery, oxidative enzyme activity, and oxygen utilization in the muscle precipitating a further loss of musculoskeletal and cardiovascular functional reserve to low or dangerous levels. In this cascade of events, specific muscle gene activation and expression are altered as well. Physical inactivity causes change in muscle fiber type composition and decreases formation of oxidative muscle fiber types 1 and IIa, the main factors in reduction of endurance and fitness (30).

### Functional Impairments of Disused Muscle

#### Mobility and ADL

The progressive weakness and reduced stamina resulting from inactivity negatively impact the ability to perform basic mobility and ADLs. In the lower limbs, type I muscle fibers, which are active during standing and slow ambulation, are especially affected with a rapid reduction in endurance. If the quadriceps muscle is immobilized in an extended position, the deep layer of vastus intermedius, which has predominantly type I and type IIa fibers, shows the greatest histochemical changes in contrast to the rest of the muscle (31). Such accelerated rate of atrophy and weakness was also noted in hip and back extensors, hip abductors, and ankle plantar flexors and dorsiflexors, impacting the ability to walk (5). Disuse weakness and loss of endurance will result in impaired ability to perform ADLs and ambulate safely, reducing personal independence (Fig. 48-2).

#### Muscle Pain and Stiffness

Patients frequently complain of back pain during bed rest. The cause of this pain is still not fully understood. Recent studies in which spine length and degree of movement were measured by miniature ultrasound transmitters have demonstrated that back pain is more prominent when trunk movements are limited in a supine position. It is speculated that localized, prolonged, low-intensity isometric muscle contractions may cause this pain. Back pain can be averted by stretching exercises and walking (32).

Limb muscle pain and stiffness occur after generalized immobility or focal limb immobilization, especially in the presence of limb swelling. However, gentle static or intermittent stretching can alleviate this problem and prevent the reduction in number of sarcomeres in series (33). It has been shown that the position in which a joint is immobilized has a significant
influence on the number of sarcomeres present in a single muscle fiber (33). When immobilized in a shortened position (extensors in a fully extended position or flexors in a fully flexed position), a muscle can lose 40% of its original number of sarcomeres, contributing significantly to weakness and muscle stiffness (34).

Disuse Weakness, Deconditioning, and Cardiovascular Disease

A lack of adequate muscle activity adversely affects the cardiovascular and related systems. As noted in detail above, chronic inactivity impairs and reduces maximal oxygen consumption ($VO_{2max}$), cardiovascular reserve, and fitness (35). Individuals with an inactive lifestyle and low level of fitness are more prone to develop coronary artery disease (CAD) and have greater odds of suffering myocardial infarction and death. A number of epidemiologic studies have demonstrated an inverse dose-response relationship between physical activity and mortality from CAD (36). The mechanism of this interaction is not known, but the current belief is that regular physical activity prevents or suppresses the chronic inflammation associated with atherosclerosis, CAD, and the metabolic syndrome (37–39).

Prevention and Treatment of Muscle Weakness

Principles

- Identify clinical and subclinical changes in strength, endurance, and physical function.
- Determine whether additional conditions exist that may exacerbate atrophy and weakness caused by inactivity, such as acute trauma or chronic disease.
- Ascertain the necessity for continued bed rest or immobility.

Strategies

- Prevent muscle weakness by prescribing progressive resistive exercise, stretching, and aerobic exercise.
- In most cases, use a combination of specific exercises to address all aspects of muscle weakness, including exercise for flexibility, strength, endurance, and fitness.
- Remobilize the patient as quickly as possible; provide progressive mobility training.
- Encourage 30 minutes of walking and leisure activities for at least 3 days a week for the general population.
Flexibility Exercise

Streching to maintain optimal muscle resting length as well as viscoelastic properties is important for maintaining normal muscle function. Several animal studies indicate that passive stretching of striated muscle is associated with muscle hypertrophy, increase in muscle fiber area, increase in number of sarcomeres, and muscle fiber proliferation. In doing so, undifferentiated, quiescent myoblasts residing on the sarcolemma of muscle fibers (satellite cells) are activated and believed to be responsible for this stretch-induced muscle hypertrophy. Passive stretch is also a potent mechanical stimulus to influence gene expression and muscle proliferation. For example, both myogenin mRNA per microgram RNA and muscle cross-sectional area were significantly increased after 3, 6, 14, and 21 days of stretch (overload) in avian latissimus dorsi (43).

Muscle stiffness due to lack of stretch occurs through the reduction of the elongation properties of elastic and collagen fibers and myofibrils. It is also a result of structural changes in the muscle, such as muscle fiber angulations, reduction of sarcomeres in 18 series, and rearrangement of collagen fibers and their cross-links. Even a relative increase in muscle connective tissue may lead to muscle stiffness and reduction of joint motion. Muscles that cross two joints such as the hamstrings, gastrocnemius, and long back extensors are particularly prone to stiffness, even in healthy subjects with limited physical activity. In immobilized and inactive persons, however, this process is accelerated in the absence of activity-induced stretch. Stiffness and subsequent muscle belly shortening of two-jointed muscles can interfere with functional walking. Hip-flexion contracture at 35 degrees as a result of iliopsoas muscle tightness causes a 60% increase in energy consumption per unit distance during ambulation (5). Thus, daily stretching of a muscle for a half hour can prevent the loss of sarcomeres in series of the immobilized muscle and maintain elongational properties of muscle fibers and surrounding connective tissue, maintaining full ROM of joints.

Strengthening (Resistance) Exercises

To prevent disuse weakness and deconditioning, daily physical activity should be encouraged, and progressive resistance exercise (PRE) using isokinetic or isotonic exercises done on a regular basis. A minimum of once a day muscle contraction at 30% to 50% of maximal strength for 3 to 5 minutes, three times a week for a single muscle group, may suffice to prevent muscle loss and weakness in otherwise mobile, active individuals. Strengthening exercise prevents muscle atrophy and improves strength and endurance by mediating the activation of muscle genes, enhancing mitochondrial function, and improving protein synthesis, causing muscle fiber phenotypic changes and muscle angiogenesis (Table 48.2).

Muscle strength can be increased by concentric or eccentric exercises performed isokinetically or isotonically. To improve or maintain strength, there should be 8 to 15 repetitions for each muscle group done twice with brief pauses between sets. There are many studies that indicate that strenuous eccentric exercises may produce muscle damage since CK may increase 50-fold and cytokine and interleukin-6 (IL-6) fivefold. Furthermore, high-intensity eccentric exercises in untrained individuals frequently produce elevation of myofibrillar enzymes in plasma, ultrastructural damage of muscle fibers, infiltration of mononuclear blood cells, and acute inflammation (44). Therefore, it is important that eccentric exercises are started with lower intensity, fewer repetitions, and slower progression of applied resistance or loads.

Muscle weakness can also be prevented by the use of electrical stimulation. For example, applying local stimulation to the quadriceps while in a long leg cast may help preserve muscle bulk and strength and also may shorten rehabilitation time, a factor that may be particularly important in an athlete (45). For astronauts exposed to prolonged periods of microgravity with resulting muscle atrophy, functional electrical stimulation (FES) has prevented atrophy when applied 6 hours/day, with 1 second on and 2 seconds off stimulation at 20% to 30% of maximum tetanic force applied to the two pair of agonist-antagonist muscles in the legs (46).

For chronic disuse atrophy with weakness and stiffness, strengthening and stretching exercises may be required for several months, and, even then, return to normal strength and ROM may not be complete. For example, returning to normal functional capacity and fatigue resistance after 8 weeks of immobilization in persons with ankle fractures required 10 weeks of supervised physical therapy, demonstrating that more time is needed to restore muscle capacity than to develop muscle weakness (47).

Exercises for Endurance and Fitness

Aerobic exercise should be prescribed for physically inactive individuals with history of prolonged bed rest, immobility, or limited physical activity regardless of age and sex and for those with high risk for CAD and metabolic syndrome. Covertino (48) and others have reported that decline in VO2 max is progressive and parallels the duration of bed rest and it occurs independently of age, sex, and the presence of any other disease and it is considered a major factor for the loss of cardiovascular fitness. Daily endurance exercises at 60% to 80% of VO2 max or at the target heart rate are required to maintain or improve aerobic capacity (VO2 max) in persons with deconditioning.

TABLE 48.2 Resistance/Strengthening Exercises Guide

<table>
<thead>
<tr>
<th>Exercise</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Establish one repetition maximum for each muscle group.</td>
<td></td>
</tr>
<tr>
<td>Select initial and subsequent intensity at 50%–80% of the maximum.</td>
<td></td>
</tr>
<tr>
<td>Repetition is performed 8–15 times, twice per session for each muscle group, three times per week or more.</td>
<td></td>
</tr>
<tr>
<td>Progressive increase in resistance (i.e., weights) and time until goal is achieved.</td>
<td></td>
</tr>
<tr>
<td>Include sequentially all muscle groups of the lower and upper limbs, back and abdominal muscles.</td>
<td></td>
</tr>
<tr>
<td>Focus on antagonistic muscles, agonists, antagonists, and functional training.</td>
<td></td>
</tr>
<tr>
<td>Exercise should become a habit and part of a person’s wellness program.</td>
<td></td>
</tr>
</tbody>
</table>
Improve endurance, VO2max, and cardiovascular fitness by performing resistance exercises regularly (35). Daily or three to four times per week, resistance or aerobic exercise targets the muscles of the lower and upper limbs for 2 to 3 months to restore and increase endurance and fitness (49).

Short-time exercises can also prevent, attenuate, or reverse the process of deconditioning and functional decline during the period of bed rest. For example, high-intensity and short-duration isotonic ergometer exercises can maintain work capacity, plasma and red-cell volume, reverse negative water and electrolyte balance, and decrease the quality of sleep and concentration when compared with no-exercise groups during bed rest. However, isokinetic exercises, high-intensity short-duration, can only slightly attenuate the decrease in peak VO2max and the loss of red-cell volume but not loss of plasma volume, cannot reverse negative water balance, and have no effect on quality of sleep and concentration (50). These studies have pointed out that different training protocols are required to reverse or prevent the adverse effects of inactivity.

Regular physical activity and exercises do provide anti-inflammatory effects in conditions with low-grade chronic inflammation such as CAD or type II diabetes resulting in the reduction of inflammatory cytokines and increased production of anti-inflammatory cytokines (51). Regular endurance exercise also activates a number of genes, enhancing their expression and leading to the production of oxidative muscle fibers (type I and type IIA) and phenotypic changes of muscle fibers.

Changes in motor unit recruitment and force of contraction further contribute to the loss of endurance and easy fatigability of immobilized subjects (52). Current international guidelines suggest that 30 minutes/day or 3 to 4 hours/week of aerobic exercise should be enough to improve fitness and prevent cardiovascular deconditioning. One to two hours per week of this activity needs to be of higher intensity (40).

**JOINT CONTRACTURE**

While joint contractures can be associated with a primary joint pathology, they more frequently result from a combination of contributory conditions coupled with a lack of joint motion. Changes resulting in limited joint motion may stem from joint pain, arthropathies, paralysis, capsular or periarticular tissue fibrosis, or muscle belly tendon changes. However, the single most common factor that contributes to the occurrence of fixed contractures is a lack of joint mobilization throughout the full allowable range. For instance, prolonged elbow immobilization in a flexed position will cause reduction of resting muscle length of flexors and capsular or soft-tissue tightness with resultant fixed joint contracture (53).

Many factors, such as limb position, duration of immobilization, and preexisting joint pathology, affect the rate of contracture development. Edema, ischemia, bleeding, and other alterations to the microenvironment of muscle and periarticular tissue can precipitate the development of fibrosis. Advanced age is also a factor since both muscle fiber loss and a relative increase in the proportion of connective tissue occur in the elderly (54). Microvascular changes and relative ischemia as found in diabetes mellitus can predispose to contractures, especially of the hand. Contractures that are precipitated by pathologic changes in the joints or muscles may be classified into three groups (Table 48-3): arthrogenic, myogenic, and soft tissue related. It is important to remember that all tissues surrounding a joint are likely to eventually become affected in joint contracture regardless of the initiating process.

**TABLE 48.3 Anatomical Classification of Contractures**

<table>
<thead>
<tr>
<th>Type</th>
<th>Primary Cause</th>
<th>Secondary Cause</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arthrogenic</td>
<td>Cartilage damage, congenital deformities, infection, trauma, degenerative joint disease</td>
<td>Immobility</td>
</tr>
<tr>
<td></td>
<td>Synovial and fibrofatty tissue proliferation (e.g., inflammation, effusion)</td>
<td>Immobility</td>
</tr>
<tr>
<td></td>
<td>Capsular fibrosis (e.g., trauma, inflammation)</td>
<td>Lack of ROM</td>
</tr>
<tr>
<td></td>
<td>Immobilization as primary cause</td>
<td>Mechanical position</td>
</tr>
<tr>
<td></td>
<td>Penarticular soft tissue (e.g., trauma, inflammation)</td>
<td>Immobility</td>
</tr>
<tr>
<td>Soft and dense tissue</td>
<td>Skin, subcutaneous tissue (e.g., trauma, burns, infection, systemic sclerosis)</td>
<td>Immobility</td>
</tr>
<tr>
<td></td>
<td>Tendons and ligaments (e.g., tendinosis, bursitis, ligamentous tear, and fibrosis)</td>
<td>Immobility</td>
</tr>
<tr>
<td>Myogenic</td>
<td>Traumatic (e.g., bleeding, edema)</td>
<td>Immobility</td>
</tr>
<tr>
<td>Intrinsic</td>
<td>Inflammatory (e.g., myositis, polymyositis)</td>
<td>Fibrosis</td>
</tr>
<tr>
<td>(structural)</td>
<td>Degenerative (e.g., muscular dystrophy)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ischemic (e.g., diabetes, peripheral vascular disease, compartment syndrome)</td>
<td>Immobility</td>
</tr>
<tr>
<td></td>
<td>Spasticity (e.g., strokes, multiple sclerosis, spinal cord injuries, hypertonicity)</td>
<td>Lack of stretch</td>
</tr>
<tr>
<td></td>
<td>Flaccid paralysis (e.g., muscle imbalance)</td>
<td>Faulty joint position</td>
</tr>
<tr>
<td></td>
<td>Mechanical (e.g., faulty position in bed or chair)</td>
<td>Immobility</td>
</tr>
<tr>
<td></td>
<td>Immobilization as primary cause</td>
<td>Lack of stretch</td>
</tr>
<tr>
<td>Extrinsic</td>
<td>Combined arthrogenic, soft-tissue and muscle contractures noted in a single joint</td>
<td></td>
</tr>
<tr>
<td>Mixed</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Mechanical Properties of Connective Tissues

Connective tissue can be subdivided into five major groups: (a) loose connective tissue, (b) dense connective tissue (i.e., ligaments), (c) cartilage, (d) bone, and (e) blood vessels. Loose and dense connective tissues are complex, dynamic structures that are important for structural support, stabilization, and movement. It is not always well appreciated that these are living, changeable tissues that can adapt their structure and composition in response to a change in environment, particularly in response to mechanical stresses. An appreciation of the anatomic design is important to fully understand the mechanical properties of the loose and dense connective tissues and their relationship to passive stretch. Both types of connective tissues are composed of cells (fibroblasts) and intercellular macromolecules surrounded by polysaccharide gel, also called extracellular matrix. The intercellular substances, or collagen, impact the mechanical properties of the tissue, whereas the cells are important for homeostasis, adaptation, and repair functions (55).

Collagen

There are two types of intercellular substances in connective tissues: collagen fibers and proteoglycans. Fibers in tendons, ligaments, joint capsules, and muscle endomysium and perimysium are predominantly of the collagen type, although there is a significant population of elastic fibers in tendons. This is consistent with function in that tendons have great tensile strength and some elasticity that allows a joint complex to move through all stages of muscle contraction and relaxation. Ligaments, on the other hand, are relatively inelastic and are composed primarily of collagen fibers. Collagen is the most abundant protein in the body and accounts for more than 20% of total body mass. At least 12 different collagen types have been well-characterized. More than 30 different genes produce different aggregations of specific polypeptides, resulting in the different collagen types. The terminology used in describing the organization and aggregation of collagen molecules is inconsistent and confusing. All collagen molecules have a unique protein conformation known as the triple helix, a result of three constituent polypeptide chains of the collagen molecule coiled together. The synthesis of these chains from amino acids, known as pro–chains, occurs in the rough endoplasmic reticulum of the fibroblast. The precise amino acid sequence differs between the different types of collagens and accounts for the tissue-specific properties. When the collagen molecules (monomers) are subsequently secreted from the cell, enzymatic cleavage of the end-part of the molecule occurs, and the molecules aggregate in a systematic manner to form fibrils in the extracellular space (56). Collagen fibrils, visible with the electron microscope, are grouped into fibers that are visible with the light microscope. Cross-linking between collagen fibrils is another important structural feature that varies with location and function. The type and location of collagen cross-linking are the key to tensile strength and can be altered, depending on the direction and magnitude of applied mechanical loads. The fibers aggregate into fiber bundles that are grouped together into fascicles. A large number of fascicles form the whole tendon or ligament (57). In striated muscle, collagen fibers form endomysium around the muscle bundles and perimysium around the muscle fascicles. These are covered with thin films of loose or dense connective tissues surrounding collagen fiber bundles of tendon or ligament (endotendon or endoligament), fascicles (peritendon or periligament), and the whole tendon or ligament (epitendon or epiligament). The epitendon and epiligament, as well as endomysium, are thought to be critically important in responding to mechanical loads and injury, and all play a part in the onset of myogenic, arthrogenic, or soft-tissue contractures (58).

In tendons and ligaments, type I collagen predominates, although types III, IV, and VI are also present. Important variations in collagen diameter have been found in association with site, age, activity level, and repair. Investigations in several animal and human models alike have demonstrated that changes in collagen diameter, density, and orientation follow Wolff’s law; that is, connective tissues orient themselves in form and mass to best resist extrinsic forces. This has been established in response to physiologic conditions (e.g., immobilization or exercise) as well as in response to injury. Changes in collagen are mediated by fibroblasts that are sensitive to mechanical stimuli, enzymes (collagenase and tissue inhibitor of metalloproteinases), and growth factors. These factors shift the dynamic equilibrium toward synthesis or degradation, depending on environmental factors (59). If extrinsic factors, such as stretch or weight bearing, are limited, or if a joint is immobilized in a foreshortened position, then collagen fiber density and mass will be readjusted to the new positions or to new loads, reducing ROM of a joint and breaking point of ligaments and tendons (60).

Proteoglycans

Although proteoglycans make up only about 1% of the dry weight of ligaments and tendons, their functions of lubrication, spacing, and gliding are essential (61). Proteoglycans also impact the viscoelastic properties of dense connective tissues. There are several different types of proteoglycans (e.g., hyaluronic acid, chondroitin sulfate, decorin, aggrecan, biglycan) that are specific to site and function. An examination of different regions of a tendon as it traverses around a bony pulley is an excellent example of adaptation of proteoglycans by dense connective tissues. The proximal region of the tendon (at a distance from the bony pulley) is only exposed to tensional forces and contains a scant amount of decorin providing some lubrication to the surrounding collagen fibers. By contrast, the region of tendon that is in contact with bone (and subjected to compression, gliding, and tension) contains approximately 10-fold more proteoglycans, most of which is chondroitin sulfate. In other words, the tissue is more like fibrocartilage in the area of compression to withstand the mechanical forces in that region. Work in animal models has demonstrated that these proteoglycan levels may adjust to environmental factors and be reduced by immobility (61).
Disuse Changes in Connective Tissue

After trauma or inflammation of connective tissue, undifferentiated mesenchymal cells start to migrate to the site of injury and gradually change into mature fibroblasts. The fibroblasts travel along fibrin layers, multiply, and develop collagen-producing organelles (56). These new collagen fibers are either arranged randomly in the loose connective tissue or packed and oriented in the direction of force and stretch in dense connective tissue. Hence, the mechanical property of newly formed connective tissue is the result of the type and amount of collagen produced, as well as bonding and orientation of the collagen fibers and the intensity and duration of loading and stretching (53,57).

The balance between synthesis and degradation is disturbed by physical factors, such as the lack of static stretch that is seen in prolonged immobilization or immobility. Trauma with bleeding into the soft tissue and muscle, inflammation, degeneration, or ischemia could all trigger an increased synthesis of collagen. In these conditions, additional lack of stretch and mobility may cause the collagen fibers to become more tightly packed and randomly arranged (57).

The collagen in muscle connective tissue provides important functions, such as linking muscle cells and tendons, and is a supportive structure that holds muscle fibers and fascicles together. The synthesis of collagen tissue in the muscle is influenced by tension produced by muscular contraction, weight bearing, and stretching imposed by these activities. Hence, collagen synthesis in the muscle is greater during activity and reduced during immobility. Immobilization for 1 week causes 21% and 65% decreases in activity of the enzymes prolyl 4-hydroxylase and galactosylhydroxylase glucosyl transferase in nontrained and trained experimental animals, respectively (57,59). Collagen synthesis is reduced during immobility but proportionally less than muscle protein synthesis, and both can be reversed by resistance exercise and by stretching (19,21). Multiple factors play a role in development of joint contractures. The initial muscle stiffness and tightness in contracture formation is due to the myofibril protein titin, sarcolemma (containing collagen IV), and reduction of sarcomeres in series, as well as changes in the angulations of muscle fibers in respect to their origin and insertion. In the later stages of fixed contracture development, the connective tissues in the muscle, joints, and soft tissue become randomly oriented, tightly packed, and shortened in length, the processes mainly induced by the lack of physiological stretch and physical inactivity (33,62).

Myogenic Contracture

Myogenic contracture is a shortening of resting muscle length that is due to intrinsic or extrinsic causes, limiting full ROM and causing abnormal positioning of the limbs or body. Intrinsic changes are structural and may be associated with inflammatory, degenerative, or traumatic processes. Extrinsic muscle contracture is secondary, resulting from neurologic conditions or mechanical factors. The diagnosis of muscle contracture should be made only after careful physical examination, which should include an evaluation of active and passive ROM. Observing limitation of active ROM alone can lead to an erroneous conclusion since such limitation also could be due to muscle weakness.

Muscular dystrophy is an example of an intrinsic degenerative process in the muscle. The most significant histologic changes in this condition are muscle fiber loss, segmental necrosis, increased numbers of lipocytes, and fibrosis. The replacement of functional muscle fibers with collagen and fatty tissue in concert with chronically shortened resting muscle length results in contracture (63). Direct muscle trauma also can result in fibrosis. After hemorrhage into a muscle, fibrin deposition results initiating muscle fibrosis. After trauma, immobilization can independently increase serum CK activity, local vascular permeability, and swelling, all promoting development of myogenic or soft-tissue contracture (59).

Among the processes that may also cause intrinsic muscle shortening is heterotopic ossification. This is most commonly noted after trauma, joint surgery (especially of the hip), spinal cord injury (SCI), or other central nervous system injuries. The actual initiating factor is unknown. An alteration in local metabolism or blood flow, in connection with the systemic alteration in calcium metabolism that occurs with immobility, may be responsible for initiating this process. Although no truly effective treatment exists, ROM should be aggressively maintained. Surgical resection of the bone may be considered after the bone matures. Surgery for immature heterotopic ossification is often associated with a rebound phenomenon that worsens the extent of previous bone deposition. A single local radiation treatment may be used to limit recurrence. Prophylaxis can be accomplished through the use of disodium etidronate, a diphosphonate compound that prevents the calcification of ground substance, along with nonsteroidal anti-inflammatory drugs and with early mobilization (64).

Extrinsic myogenic contracture is the most common type occurring after multiple injuries and chronic illness, as well as in sedentary individuals. In planning a therapeutic approach, it is useful to identify the associated cause of an extrinsic contracture such as paralysis, spasticity, or biomechanics. If a paralyzed muscle cannot provide adequate resistance to its antagonists across a joint, then the stronger muscle will prevail and become shortened. A common example of this is the shortened triceps surae as seen in persons with chronic peroneal nerve palsy. Stretch applied to the muscle is essential to prevent contracture in these situations; however, strengthening of the weak muscle and proper positioning are also vital.

Similarly, in the presence of spasticity, a dynamic imbalance of muscle control exists across one or more joints. The resting length of spastic muscle is shortened because of increased muscle tone, which encourages faulty joint positioning and precipitating functional or fixed contracture development (see Fig. 48-2). In severe spasticity, it is often clinically difficult to identify the actual onset of structural changes in the spastic muscle. If full ROM is unobtainable even after prolonged stretch and tension, then fixed muscle shortening should be expected.
Mechanical factors can alone cause extrinsic myogenic contracture. Some degree of muscle shortening can be present even in healthy persons with reduced mobility and lack of exercise. In such situations, two joint muscles in the lower extremities are first to be involved since natural stretching during ambulation is reduced. The hamstring and back muscles are among the first to be involved; the iliopsoas, rectus femoris, tensor fascia lata, and gastrocnemius are next most likely to become shortened. In the upper limbs, the internal rotators of the shoulder and elbow flexors show most frequent contractions. The below-knee amputee with prolonged knee flexion while sitting will develop knee-flexion contracture as a result of tight hamstrings and soft tissue behind the knee. On the other hand, the below-knee amputee treated with a rigid postoperative dressing in full knee extension may develop quadriceps muscle tightness that prevents full knee flexion.

Patients with muscular dystrophy provide yet another example of contracture promoted by biomechanical factors. In muscular dystrophy, hip extensors often are very weak, forcing the patient into excessive lumbar lordosis to thrust the center of gravity behind the hip and in front of the knee joints; the patient tends to walk on his toes. Walking on the toes in full ankle plantar flexion prevents natural stretching of the triceps surae from occurring during the stance phase of gait, encouraging muscle shortening. If a clinician does not recognize this sequence of events, he or she might assume that weakness and fibrosis are the only reasons for a plantar-flexion contracture. A surgical lengthening of the Achilles tendon in such a case will not give the expected improvement. The lengthening of the tendon will actually shorten the muscle belly, decreasing plantar-flexor strength and diminishing the ability of these patients to walk on their toes. Because walking on the toes is the only feasible method of ambulation, the result of an ill-advised tendon lengthening may be a wheelchair-dependent patient (63).

**Arthrogenic Contracture**

Pathologic processes involving joint components, such as degeneration of cartilage, congenital incongruency of joint surfaces, or synovial inflammation and effusion, are all accompanied by pain that fosters limited joint mobility, keeping the joint in less painful posture and fostering arthrogenic contracture development. In experimentally induced acute crystalline arthritis, exercise aggravates synovitis, whereas a short period of immobilization helps to reduce inflammation (65). In chronically induced experimental arthritis, however, joints immobilized for several weeks showed much greater destruction of joint cartilage than freely mobilized joints (66). Short-term immobilization is indicated in acute arthritis because of the presence of interleukin-1 in inflammatory synovial fluid. Studies indicate that passive ROM during acute arthritis may increase the release of interleukin-1, promoting interleukin-1 penetration into cartilage and binding to the receptors on the chondrocyte membranes and inhibiting the production of proteoglycans necessary for protection of cartilage (66). However, a study by van den Ende et al. clearly demonstrated that intensive dynamic exercise performed in addition to ordinary physical activity in active rheumatoid arthritis patients did not induce worsening of the disease process, but rather, physical functioning of these patients was significantly improved. The authors concluded that an intensive exercise program is more effective than a conservative exercise program in improving the strength and functional capacity and does not worsen activity of the disease (67).

Cartilage loss and subsequent adhesions and pain cause muscle splinting, leading to loss in ROM and fixed contracture. The fact that pain and muscle splinting, not the loss of cartilage, are more responsible for the development of contracture is illustrated by patients with Charcot joint. These patients lack both pain sensation and proprioception but maintain relatively well-preserved ROM or even hypermobility of the involved joints in the presence of severe destruction of cartilage and joint surfaces.

The joint capsule also can lose its extensibility as a consequence of collagen fiber shortening, inadequate joint stretching, and faulty positioning (68). If the entire capsule is involved, ROM is compromised in all directions of movement. In the later stages of arthrogenic contracture, the periarticular tissue may also undergo histochemical changes (61). If repeated trauma with pain is combined with swelling and prolonged immobilization, periarticular tissue undergoes contracture development. In these situations, proteoglycan content is also decreased, contributing further to joint stiffness and contractures (61).

The shoulder and hip joint capsules often become symptomatic from contracture earlier than other joints. Initiating factors may include bicipital tendonitis, subdeltoid bursitis, rotator cuff damage, spasticity, or poor positioning coupled with immobility. However, posterior knee capsule shortening is a consequence of hamstring contracture or prolonged flexion positioning as seen in patients with wheelchair dependence for mobility (see Fig. 48-2) (68).

**Soft-Tissue Contracture**

Cutaneous, subcutaneous, and loose connective tissue around the joint may also become contracted with immobility. Trauma to soft tissue with bleeding, for example, can initiate fibrosis, which may progress to contracture if joint movement is not allowed. In this situation, collagen fibers proliferate and are laid down in random fashion. In contrast to capsular tightness, soft-tissue shortening usually will limit movement in only one plane or axis. Burned skin is particularly susceptible to contracture. During recovery, burns that cross any joint must be ranged diligently and positioned to oppose the shortening forces of scarred tissue.

Vigorous active and passive ROM exercises, placement of the joint in a functional resting position, and use of compressive garments should be considered to prevent contractures in burn patients. Topical steroid and vitamin E applications have failed to reduce soft-tissue contracture or postoperative scar formation after reconstructive joint surgeries. Here again, adequate stretch and mobilization are important factors in the prevention of fixed contractures.
Ligaments also show biomechanical and biochemical changes during immobility. They become weaker and break easily. In experimental animals, the rate of growth, length of ligament, and elongational characteristics are influenced by tension applied to the ligament, growth hormones, size of underlying bones, and other unknown factors (68). A study in young, skeletal immature rabbits showed that the lateral collateral ligament of the knee could be elongated significantly (140%) when tension was applied for 6 weeks (69). A number of other studies indicate that physiologic stretch and tension are important factors in helping both growing and mature tendons and ligaments to maintain their mechanical properties, to withstand the stresses of weight bearing and mobility (7). Another cause for ligament weakness is decreased collagen synthesis. In addition, the ligament insertion sites on the bone show an increase in osteoclastic activity. During immobility, the fibroblasts of the cruciate ligament may assume a spindle shape with multiple cytoplasmic extensions and demonstrate reduced production of collagen fibrils, an important factor when considering ligament resistance to breakage (68). The disuse atrophy of collateral and cruciate ligaments can be prevented completely with resistive and ROM exercise in non-weight-bearing limbs. However, this exercise could not prevent significant bone loss in the femur and tibia in the same experimental animals (70,71). Another proof for deleterious effects of inactivity came from a study on middle-aged men who lived for 90 days in microgravity conditions. Half of them were placed on exercise regimens, performing fly wheel and isometric exercises every third day. Tensile deformation of Achilles tendon significantly decreased by 58% in nonexercising group. In the exercise group, this decline was 37% indicating that loading in this group was insufficient and did not prevent a decline in tensile property of tendon.

**Management of Contractures**

**Analysis**

A basic approach for initiating treatment for contractures is a careful determination of the predisposing factors, as well as knowledge of what joint components or tissues are actually involved. An observant neuromuscular examination emphasizing active and passive ROM and joint stability is essential. Particular attention should be directed to those muscles crossing two joints and to the application of special maneuvers to detect their shortening. In patients with severe uncontrolled spasticity, it may be necessary to obtain accurate ROM measurements with the use of regional or local nerve anesthesia; this is particularly helpful when surgery to repair a decubitus ulcer is contemplated. Of course, the best treatment is prevention, so a careful analysis of abnormal joint positioning and factors limiting ROM should be undertaken with any patient who is immobilized by disease or by the treatment of disease.

**Indication for Stretching and ROM Exercises**

Once a contracture has developed, the sine qua non for treatment is active and passive ROM exercises combined with a sustained terminal stretch on a daily basis (Table 48-4). For
TABLE 48.4 Basic Principles in the Prevention and Treatment of Contractures

**Prevention**
- In healthy individuals with sedentary lifestyle, elderly
- Flexibility exercises, stretch of two-jointed muscles, yoga, pilates
- In individual with preexisting condition or predisposition:
- Range-of-motion exercises (active or passive) with terminal stretch
- Proper positioning in bed, wheelchair, splinting, casting
- Early mobilization and ambulation (weight bearing)
- CPM (continuous passive motion)
- Resistance exercise to opposing muscles

**Treatment**
- Passive range-of-motion exercises with terminal stretch
- Prolonged stretch using low passive tension and heat (e.g., ultrasound)
- Progressive (e.g., dynamic) splinting, casting
- Treatment of spasticity; pharmacologic, motor point or nerve blocks using phenol, muscle injection of botulinum toxin A or B
- Pain management
- Surgical interventions (e.g., tendon lengthening, osteotomies, joint replacement)

Mild contracture, sustained or intermittent stretching for 5 to 10 minutes daily may be effective. Prolonged stretches of 20 to 30 minutes or more combined with subsequent appropriate joint positioning and splinting are necessary for more severe contractures. This generally is more successful when used in combination with the application of deep heat at the musculotendinous junction or joint capsule. Ultrasound is the most popular heat source for large joints; its properties allow local heating in the presence of metallic implants and rapid increase of tissue temperature to the therapeutic level. Heating of the tissue to 40°C to 43°C will increase the viscous properties of connective tissue and maximize the effect of stretching (see Chapter 63, Therapeutic Physical Agents).

When applying terminal stretch to a joint, the proximal body part should be well stabilized. In many cases, slight distraction of the joint during stretch will prevent joint compression and possible soft-tissue impingement, particularly in the small joints of the hand. The shoulder is a common site of contracture, particularly in adducted and internally rotated positions. In this contracture, the normal downward sliding and rotation of the humeral head on the glenoid fossa do not occur, and forced abduction will, therefore, simply cause painful impingement of the rotator cuff tendon against the acromion. Stretch applied in forward flexion and external rotation will restore some of this motion and should be attempted before abduction.

Sustained stretch lasting 2 hours or more can be obtained by the use of splinting. Serial casting is the application of plaster or polymer bandages with careful padding over bony prominences. The cast is applied immediately after the use of heat and manual stretch to obtain maximal ROM. The cast can be reapplied weekly. In patients with spasticity, chemical denervation (botulinum toxin injection) or local anesthetic injection before casting may improve positioning inside the cast and tolerance for wearing.

Dynamic splinting provides tension in the desired direction with the use of springs or elastic bands. This type of splinting is often used in the hand and arm because it allows a measure of function while providing stretch. Another way to provide a form of sustained stretch is to use a continuous passive mobilization (CPM) device. The use of these devices has become relatively routine for providing postoperative ROM stretching of the knee, and they have been adopted for use on other joints. CPM is recommended for the early mobilization of infected joints, synovectomized knees and hips, knee fractures, ligamentous repairs, some total knee joint arthroplasties, or any incipient myogenic or arthrogenic contractures. Early passive mobilization with CPM has been shown to promote the exchange of joint fluid, reduce the need for pain medication after surgery, and prevent contractures in high-risk patients. During CPM therapy, muscles around the joint remain relaxed, and pain is usually minimal. CPM is typically prescribed for 8 to 12 hours a day for a total of 3 to 5 days after surgery. When used alone, CPM is not effective in the treatment of fixed contractures.

To achieve optimal joint position, it sometimes is necessary to lengthen tendons by surgical means. The benefits and risks of tendon lengthening should be considered carefully. It must be remembered that the muscle belly will remain shortened or become even shorter when the tendon is made operatively longer; therefore, full active ROM may not be restored. Tendon lengthening combined with muscle transfer procedures in spastic, paralytic, or other contractures may give better results because the process attempts to restore equilibrium of the forces around the joint; this method is particularly effective at the ankle, using the tibialis anterior and posterior muscles. Electromyographic analysis of muscle function before tendon transfer will optimize results by ensuring the transferred muscle is innervated. In other situations, such as hip adductor contractures 40 secondary to spasticity, tenotomy may be combined with obturator nerve neurolysis to obtain optimal results. In a chronic, fixed joint contracture that interferes with the patient's basic physical functions, the selection of the appropriate surgical procedure is of great importance (74,75).

**Prevention of Contractures**

Prevention of contracture in a bed-bound patient starts with the selection of an adequate bed and mattress, proper bed positioning, and a bed-mobility training program. The patient should be moved out of bed as soon as the medical condition allows. If bed rest is unavoidable, then bed positioning and bed-mobility training are incorporated into the patient’s nursing management program. To assist a patient in turning side to side or in sitting up, partial side rails for grasp should be a standard part of bed equipment. An overhead trapeze is sometimes useful for patients with impaired bed mobility, allowing them to use their upper extremities to roll from side to side,
scoot up and down, attain a sitting position, and transfer into and out of bed.

For the patient with immobility that is due to paralysis or with compromised extremity function, a variety of assistive devices are used to keep the joints in functional positions. In addition to that, active or passive ROM exercises should be provided on a daily basis. Provision of daily ROM and flexibility exercises is essential for prevention of contracture in high-risk patients as well as in subjects with sedentary lifestyles. Encouraging the use of early and progressive ambulation or ADL for patients in recovery will help in maintaining the function of involved joints. Ambulation and weight bearing provide a physiological stretch and are an important approach for prevention and treatment of contractures.

**DISUSE OSTEOPOROSIS**

Maintenance of skeletal mass depends largely on mechanical loading applied to the bone by muscle pull and the force of gravity. Bone mass will increase with repeated loading and will decrease with the absence of muscle activity or with the elimination of gravity (76,77). Certain populations are more susceptible to the effects of muscle inactivity or reduced weight bearing, such as the aging adult or the person with an SCI (78). Even healthy adults on bed rest lose bone at a rate that exceeds the rate of new bone formation, leading to osteopenia (79). Accelerated bone loss of tibia has been found in subjects on prolonged bed rest. Their losses of peak bone mineral density are similar to those in persons treated with chronic corticosteroids or with estrogen deficiency. Non–weight bearing over several weeks can cause trabecular and endosteal (and later cortical) mineral bone loss in the tibia, which requires 1 to 1.5 years to return to baseline level with normal activities (80). The bone mass begins to decline in the fourth and fifth decades of life, occurring most rapidly in women in the first 5 to 7 years after menopause (81), and the physical inactivity and non–weight bearing are significant additional factors for further loss of bone density. This loss of bone density and frequency of subsequent fractures were reduced in laboratory animals that were exercising during their growing age, indicating that exercise and physical activity during growth provide lifelong benefits to bone structure and fewer fractures in old age (82). However, this early benefit is compromised if a person becomes chronically deconditioned.

After an SCI, a mismatch occurs between bone growth and bone loss. Soon after initial injury, osteoblastic activity diminishes due to paralysis and non–weight bearing, and a rapid loss of bone occurs, resulting in severe osteopenia in the paralyzed region of the body (83). Furthermore, even relatively minor muscle dysfunction can result in bone loss regionally. Persons with rotator cuff ruptures have been shown to have significantly decreased bone mineral density as compared with controls, which is proportional to remaining shoulder function (84). Immobilization of forearms and wrists for a period of 5 weeks also resulted in loss of bone density in both men and women, which was not ameliorated after 5 weeks of remobilization and hand therapy (85).

Immobilization or immobility primarily decreases bone formation, specifically in the zones of high turnover rate (primary spongiosa). When coupled with other risk factors, bone mineral loss could be significantly worsened. Animal studies showed that immobilization, estrogen deficiency, and calcium deficiency each can alone reduce cancellous bone density, but when combined, greater losses in bone volume and density are found. Estrogen and calcium deficiencies result in higher rates of bone resorption in contrast to the decrease in bone formation caused by immobility and non–weight bearing (86).

How much physical inactivity contributes to reduction of bone mass in paralyzed extremities is not entirely known. In patients with hemiplegia and lost mobility, the serum and urine indices of bone resorption did not decline with time from the onset of stroke but rather correlated to the duration and level of immobility. This suggests that disuse osteopenia resulting from combined immobility and paralysis is not self-limiting and that immobility is an important factor (87). A study of identical twins in which one of each pair had SCI demonstrated that the bone mass of the pelvis and legs continues to decline over time regardless of age and gender at onset (88).

Disuse osteopenia is characterized by a loss of calcium and hydroxyproline from the cancellous portion of long bone, epiphyses, metaphyses, and cortical bone near the bone marrow cavity. To what degree increasing bone resorption plays a part in the etiology of disuse osteoporosis needs further research. During 12 weeks of bed rest, bone resorption and osteoclastic activity occur relatively later as compared to the course of reduced bone formation in the process of establishing a new bone equilibrium (89).

It has been documented that with longer duration of immobility, an increasingly long time is required to restore bone density to the preimmobility level. Full body recumbency for 3 weeks in animals resulted in a loss of both trabecular and compact bones, which remained below baseline level even 2 months after free activity was again allowed (76).

**Prevention and Treatment**

The importance of exercise training in preventing inactivity-induced osteopenia should not be underestimated whether for otherwise healthy or disabled individuals. Disuse osteoporosis can be prevented by regular use of isotonic exercises, weight-bearing and functional training. In rats immobilized by tail suspension, there were significant decreases in bone density and bone formation. Abnormalities were especially seen in trabecular bone. Animals that received specific limb exercise in addition to normal remobilization not only recovered bone mass parameters but also had improvement in the trabecular patterns, while nonexercised animals had persisting trabecular bone losses (90).

There is a great deal of evidence that resistance exercise can increase bone mass in humans. Studies have demonstrated that there is a significant correlation between muscle
strength and bone mineral density. For example, the strength of paraspinal muscles correlates with the mineral density of lumbar spine (91). Reduced back extensor muscle strength is associated with a higher incidence of vertebral fractures, thus suggesting that inactivity plays an important role in development of osteoporosis in women. Low mineral density of the spine can be improved with back extensor strengthening exercises (92).

In groups of patients with high risk for significant osteopenia, those either with paralysis or with hormonally induced osteoporosis, exercise should be provided to prevent fracture. Despite possible risk of pathologic fracture, weight-bearing and resistance exercise are particularly important in preventing progression of bone loss. Ambulation, or at least standing on a tilt table or in a standing frame, may retard the loss of calcium in those with SCIs. Additionally, in the elderly, exercise targeted at strengthening limb-girdle muscles and balance will lessen progression of bone loss and the chances of falls and pathological fractures as well (93).

**Principles and Framework**

- Recognize immobility and the lack of exercise or loading as a risk factor acting either alone or in combination with other factors.
- Understand the structural, content and strength, bone changes in response to non–weight bearing and lack of muscle contraction.
- Understand the value of remobilization, weight bearing, and physical activity, including resistance exercises, in prevention and treatment.
- Calcium, vitamin D, and physical activity are essential in maintaining the healthy bones (94).

**Typical Prescription**

- Progressive strengthening resistance exercise training for back extensors, hip extensors and abductors, and shoulder girdle muscles
- Posture, balance training, and long walks
- Avoidance of flexion exercise of lumbar spine or high-impact exercises in flexion position for those with vertebral fractures or advanced osteoporosis
- Controlled axial loading (use of a weighted vest) for reversing vertebral osteopenia

Evidence-based practice for prevention or treatment of osteoporosis and its consequences places regular exercises and calcium and vitamin D intake at the forefront of any management plan (95). In persons with impaired mobility and paralysis, that is, spinal cord injuries, additional interventions need to be used to prevent or treat this secondary osteoporosis. The bisphosphonates inhibit bone resorption and in addition inhibit calcium phosphate crystal formation (96). One bisphosphonate compound, tiludronate, has been used in paraplegic patients with some efficacy in maintaining bone volume without impairing bone formation (97). Salmon calcitonin is another compound that has demonstrated some success in maintaining bone density and may be particularly helpful in patients who cannot be adequately mobilized (98).

**Immobilization Hypercalcemia and Hypercalciuria**

Despite a normal serum calcium level, immobilized patients are markedly hypercalciuric. However, adolescent boys after acute SCI may show a significant hypercalcemia as well. Symptoms of hypercalcemia include anorexia, abdominal pain, nausea, vomiting, constipation, confusion, and, ultimately, coma (99,100). Treatment of immobilization hypercalcemia relies on achieving adequate calcium excretion through hydration with normal or one half normal saline and diuresis with furosemide. Patients with the end-stage renal disease on maintenance hemodialysis can also develop an acute hypercalcemia when placed on bed rest even for 3 days, which is difficult to treat except with remobilization (101). For immobilized healthy persons, urinary calcium excretion (hypercalciuria without hypercalcemia) increases above normal levels on the second and third days of recumbency. Maximum loss occurs during the fourth or fifth week, when urinary calcium excretion is double the level of the first week; on average, calcium loss is 1.5 g/week (94). This decrease in total calcium continues even after resumption of physical activity, and negative calcium balance can last for months and even years (79).

**EFFECTS OF INACTIVITY ON CARDIOPULMONARY FUNCTIONS**

**Hemodynamic Alterations**

Cardiovascular hemodynamic responses to bed rest, upright position, and immobility are related to body fluid distribution, inappropriate vasoconstriction of blood vessels in supine and upon assuming an upright position, and the inadequate pumping effect of the triceps surae during ambulation.

During prolonged recumbency, the resting heart rate increases by 1 beat/minute every 2 days, leading to immobilization tachycardia at rest and abnormal increases in heart rate on submaximal exercise and workloads. After 3 days of strict bed rest, heart rate increases 32% above the pre–bed rest response when resuming submaximal activity, 62% after 7 days, and 89% after 21 days of recumbency (102).

Assuming an upright position also provokes a significant increase in pulse rate, also related to the duration of bed rest, too. A healthy, active person’s heart rate increases 13% on getting up from a supine position (103). Saltin et al. (104) found a pulse rate response on submaximal exercise of 129 beats/minute for active healthy persons, compared with 165 beats/minute at the same level of exercise intensity in persons after bed rest of 3 weeks.

The vertical column of arterial blood exerts significant pressure on the lower part of the arterial tree, provoking noradrenaline release and vasoconstriction upon arising, thus preventing shifting of blood and hypoperfusion of the brain. This
vasoconstriction is delayed or absent in persons on prolonged bed rest. When standing up, there is a shift of venous blood from the thorax and upper limbs to the legs, causing an increase in venous pressure to 80 to 100 mm Hg (102). Even a short bout of leg exercise will reduce the pooling of venous blood, thus normalizing static venous pressure to 25 mm Hg.

Other cardiovascular functions also change with prolonged bed rest. Stroke volume may decrease 15% after 2 weeks of bed rest, a response that may be related in part to blood volume reduction. Although heart rate response to submaximal exercise increases progressively, cardiac output is reduced (103,104). Diminished cardiac output coupled with reduced peripheral oxygen utilization causes a significant decline in maximal oxygen consumption (VO$_{2\text{max}}$). After 20 days of bed rest, VO$_{2\text{max}}$ may decline by 27%. Inactivity also alters the erythrocyte and platelet enzymatic antioxidant defense mechanisms, increases serum triglycerides, and decreases high-density lipoproteins (HDL) and apolipoprotein A1, which may contribute to development of atherosclerosis and increased cardiovascular morbidity (105,106). The alteration of cardiovascular function induced by immobility is commonly referred to as cardiovascular adaptation syndrome (CAS). If a patient with CAD develops CAS, cardiac ischemia may be aggravated. For example, orthostatic hypotension may precipitate the onset of angina in patients with CAD. One way of preventing CAS is to encourage early ambulation and graded activity. In the 1950s, patients with myocardial infarction were kept in bed for 2 to 3 weeks; today, they are encouraged to start walking on the second day. For those patients who, for some reason, cannot be mobilized, electrical stimulation of muscles can partially prevent the negative side effects of prolonged immobility. Electrical stimulation of thigh muscles in patients with refractory congestive heart failure significantly improved muscle strength and cross-sectional area (107).

**Alteration of Body Fluids in Recumbency**

 Normally, 20% of total blood volume is contained within the arterial system, 5% in the capillaries, and 75% in the venous system. Immediately upon lying down, 500 mL of blood shifts to the thorax and cardiac output increases by 24%. Estimated myocardial work is increased by approximately 30%. During lengthy periods of bed rest, there is, however, a progressive decline in blood volume, with the maximum reduction on day 14. This reduction in blood volume is due to a reduced hydrostatic blood pressure and decreased secretion of antidiuretic hormone. Plasma volume decreases more than red-cell mass, leading to increased blood viscosity and, possibly, to thromboembolic phenomena. The loss of plasma volume after 24 hours is 5%, whereas after 6 and 14 days, the loss is 10% and 20%, respectively, of the pre–bed rest level (108). The reduction of plasma volume can be diminished by exercise. Therapeutic isotonic exercises are almost twice as effective as isometric exercises in preventing plasma volume reduction (109).

In addition to plasma volume change, a reduction of plasma proteins is noted after prolonged bed rest. A short period of intensive exercise produces smaller losses of plasma proteins, whereas sustained submaximal exercise actually induces a net gain in plasma protein, which also contributes to the stabilization of the plasma volume depletion (109). Hypovolemia, along with circulatory stasis that is due to bed rest, is an important precipitating factor in thrombogenesis.

**Orthostatic Intolerance**

One of the most dramatic effects of prolonged bed rest is the impaired ability of the cardiovascular system to adjust to the upright position. After several days of bed rest, if a healthy person stands up from a supine position, the shift of 500 mL of blood from the thorax and abdomen into the legs occurs because of impaired venous compliance and response to increased intravenous pressure. Venous return to the heart is reduced due to diminished venous compliance, an increase in venous pooling and intravascular volume depletion with end result of decreased stroke volume and cardiac output, and a significant decrease in the systolic blood pressure response on rising. In normal situations, blood pressure drop is prevented by immediate activation of the adrenergic sympathetic system. Baroreceptors in the right atrium, carotid arteries, and aortic arch trigger adrenergic reflexes releasing norepinephrine. The increase in plasma norepinephrine levels influences the release of renin and angiotensin II, which in turn potentiate the sympathetic vasoconstriction resulting in an immediate increase in pulse rate and restoration of blood pressure (110,111).

Normal orthostatic response after prolonged recumbency is significantly compromised. The circulatory system is unable to maintain a stable blood pressure and, for unknown reasons, is unable to mount an adequate sympathetic vasopressive response. Although plasma renin and aldosterone levels remain normal, blood pressure may fall more than 20 mm Hg upon rising up from supine position. The decrease in venous return triggers a rapid heart rate, which prevents optimal filling of the right ventricle during end diastole. Stroke volume, which depends on diastolic filling, may be insufficient to maintain adequate cerebral perfusion (112). The clinical signs and symptoms of postural hypotension are tingling, burning in the lower extremities, dizziness, light-headedness, fainting, vertigo, increased pulse rate (>20 beats/minute), decreased systolic pressure (>20 mm Hg), and decreased pulse pressure.

In healthy people, adaptation to the upright position may be completely lost after 3 weeks of bed rest. A significant increase in heart rate and decrease in systolic pressure may occur after only several days of recumbency in those with sepsis, major trauma, major medical illness, or advanced age. The process of restoring the normal postural cardiovascular responses can take 20 to 72 days. Older people are slower to restore normal blood pressure and heart rate after remobilization.

As a group, patients with tetraplegia are quite susceptible to orthostatic hypotension. When tilted up, they show a significant decrease in mean arterial pressure and an increase in heart rate. Both sympathetic and plasma renin activities are normal or slightly increased. Two possible mechanisms may
account for orthostatic hypotension in patients with SCI. First, the normal increase in plasma norepinephrine that occurs on tilting is delayed in patients with tetraplegia. Second, the successful use of compressive antigravity suits in treating postural hypotension indicates that venous pooling may play an important role in the occurrence of orthostatic intolerance (113).

Early mobilization is the most effective way to counteract orthostatic intolerance and should include ROM exercises, strengthening exercises in supine and upright positions, and progressive ambulation. Abdominal strengthening and isotonic/isometric exercises involving the legs are optimal for reversing venous stasis and pooling. Elevating leg rests and reclining backs or tilt-in-space wheelchairs are used to assist patients during the reconditioning process. Occasionally, a tilt table may be necessary, with the goal of tolerating 20 minutes at 75 degrees of tilt for SCI patients. Supportive garments such as elastic bandage wraps, full-length elastic stockings, and a variety of abdominal binders are used regularly. Ephedrine, midodrine, and phenylephrine are sympathomimetic agents that help to maintain blood pressure; fludrocortisone (Florinef), a mineralocorticoid, is the next choice of drug to use. Maintaining an adequate salt and fluid intake will prevent any worsening of hypotension secondary to blood volume contraction (110,113). Infusion of saline solution is also indicated when fluid volume and dehydration cannot be corrected by ordinary means (114).

**Impaired Cardiovascular Performance and Fitness**

The efficiency of the cardiopulmonary response to muscle work demand depends on the frequency with which maximal work level is attained and the degree of physical inactivity. Because of close interaction between these two systems, the maximal cardiovascular capacity gradually declines with reduced physical activity (115). This decline of cardiovascular function is accelerated in chronically ill and disabled. The clinical signs of reduced cardiovascular functional fitness are not easily recognized especially if associated with other chronic conditions.

In recumbency and deconditioning, there is a gradual elevation of the systolic blood pressure in response to increased peripheral vascular resistance and increased heart rate. In addition, the absolute systolic ejection time is shortened, and the diastolic filling time reduced, resulting in stroke volume reduction. Work capacity, which is derived from left ventricular pressure and force of ventricular contraction, is also reduced. Overall declines in cardiac output and left ventricular function with prolonged immobility and inactivity have been reported in several studies (102,104).

Furthermore, reduced cardiovascular function and fitness are associated with a twofold increase in the risk of cardiovascular mortality and morbidity (116,117). The cardiovascular risk of inactivity is similar to that of hypertension, diabetes, and high cholesterol, and it is dose related. Increase in physical activity and regular exercises, on the other hand, reduces cardiovascular morbidity and mortality (118). A number of studies have found a dose-response relationship between higher physical activity and lower cardiovascular mortality (117).

Physical activity and exercise (even of moderate intensity) have a beneficial effect on other cardiovascular risk factors such as hypertension, type 2 diabetes, and obesity (118,119). Regular physical activity and appropriate diet can reduce the risk of type 2 diabetes by 58%. This percentage is even higher in people older than 60, who have a nearly one-in-five chance of developing type 2 diabetes. It has also been documented that even low levels of physical activity like walking have a beneficial effect on cardiovascular fitness (120).

**Inactivity is a widespread chronic health problem. At least 25% of the U.S. population is not physically active at all and at risk for development of cardiovascular and musculoskeletal complications (115). In addition to the negative effect on nonexercisers, people who are active can benefit even more by increasing the intensity and duration of their activities (115).**

Many benefits of exercise training, however, can be lost within 2 weeks if physical activity is substantially reduced. These benefits are completely lost within 4 to 8 weeks if physical activity is not resumed.

**Immobilty and DVT**

Immobilty exposes the patient to two factors that are contained in Virchow’s triad and contribute to clot formation: venous stasis and increased blood coagulability. The third factor, injury to the vessel wall, is all that is required to further increase the patient’s risk for thromboembolism. Paralysis and trauma to the lower limbs or pelvis may add to the risk for development of DVT. A direct relationship between the frequency of DVT and the length of bed rest has been observed (121).

In stroke patients, DVT is ten times more common in the involved extremities than in the uninvolved extremities. In nonambulatory stroke patients, DVT is five times more frequent than in patients who can walk more than 50 ft (122). Although the first week of bed rest is the most frequent time for development of DVT, it may occur later during remobilization. When stasis is present, thrombus formation usually starts in the valve cusp of the deep veins. Stasis may contribute to anoxia and damage of the endothelial cells in the valve pocket, thus adding the third factor for initiating the onset of DVT. Whether stasis alone can result in DVT is not fully confirmed. However, several studies suggest that stasis may lead to increased formation of thrombin, which then leads to platelet aggregation and thrombosis (123). DVT mostly commonly forms in the veins in the calf. Usually, such thrombi will attach to the wall of the vein within 1 week; however, 20% of calf thrombi extend to popliteal and thigh veins, and half of these will embolize to the lungs, posing a serious threat to the patient’s life (124).

Venous stasis in the lower extremities is mainly due to decreased pumping activity of the calf muscles and increased orthostatic pressure. Other factors that can contribute to stasis are surgery, age, obesity, and congestive heart failure, all of which can lead to abnormal blood flow mechanics. Also, contributing to the likelihood of DVT occurrence in the patient confined to bed is a hypercoagulable state, produced by decreased blood volume and increased blood viscosity and
associated with many conditions, such as malignancy or blood factor deficiencies.

Clinical detection of DVT begins with the observation of signs and symptoms, including edema, tenderness, hyperemia, venous distention, and Homans’ sign. When DVT is suspected on clinical grounds, one or more of these additional diagnostic studies should be considered:

- **Doppler ultrasound study** may be 95% accurate below the knee DVT.
- **Radionuclide venography** is both sensitive and specific for thrombi above the knee.
- **Contrast venography** remains the standard for DVT diagnosis. However, it is invasive, time-consuming and irritating to the venous lining.

Pulmonary emboli are manifested by a sudden onset of dyspnea, tachypnea, tachycardia, or chest pain and often are associated with a preexisting DVT. Diagnosis rests on arterial blood gases, ventilation/perfusion scanning, and computed tomographic pulmonary angiography (CTPA).

The most common means of preventing thromboembolic complications is to use low-dose subcutaneous injections of heparin 5,000 units twice a day (125). Low molecular weight heparins (LMWHs), however, are more effective in prevention of DVT than subcutaneous injection of heparin after hip and knee surgery. Enoxaparin is a commonly used LMWH in a prophylactic dose of 30 mg subcutaneously every 12 hours or 40 mg once a day starting 12 hours postoperatively. Prevention of DVT in stroke or SCI patients can effectively be accomplished by LMWHs. The treatment of DVT without or with pulmonary emboli is also effective with LMWHs. Enoxaparin in the dose of 1.0 mg/kg subcutaneously every 12 hours or 1.5 mg/kg every 24 hours is used for inpatient treatment of DVT and for outpatient treatment, 1.0 mg/kg every 12 hours, until anticoagulation with oral agents is achieved. However, intravenous heparin is the treatment of choice for femoral or pelvic thrombosis and pulmonary embolism. The usual intravenous dose is load of 5,000 units, mix 25,000 units in 250 mL D5W (100 units/mL) and then infuse it in the rate of 11 mL/hour. The dose is monitored by activated partial thromboplastin time. For chronic treatment of DVT, warfarin (Coumadin) is a standard approach. The dose is monitored by the prothrombin time and the standardized international ratio, which should be in 2.0 to 3.0 range (126).

After the diagnosis of DVT is made and treatment with heparin and warfarin is initiated, ambulation can be permitted on the second or third day if the partial thromboplastin time is within the therapeutic range. Recent literature supports the conclusion that 5 to 7 days of bed rest is not necessary for DVT if a therapeutic level of anticoagulation is achieved before that time (127,128). Other preventive measures include external intermittent leg compression, elastic leg wrappings, active exercise, and early mobilization. After anticoagulation is established, the patient is remobilized and allowed to participate in ADL and mobility functions. Early remobilization will reduce swelling and pain in the calf muscle but will not completely eliminate danger of pulmonary emboli, especially in the patients with a thrombus whose proximal end is unattached to the wall of a vein (127).

**Pulmonary Alterations**

The respiratory complications of immobility are known to be life threatening. Initial pulmonary alterations result from restricted movement of the chest in the supine position and gravity-induced changes in the perfusion of blood through different parts of the lung. When venous and hydrostatic pressures that are due to gravity are increased in different parts of the lung, then perfusion is also increased. The balance between perfusion and ventilation is altered during recumbency (129). A change of position from upright to supine results in a 2% reduction in vital capacity, a 7% reduction of total lung capacity, a 19% reduction in residual volume, and a 30% reduction in functional residual capacity (130). Vital capacity and functional reserve capacity may be reduced by 25% to 50% after prolonged bed rest. Mechanisms responsible for this may include diminished diaphragmatic movement in the supine position, decreased chest excursion, progressive decrease in ROM of costovertebral and costochondral joints, and shallower breathing with a subsequent increase in respiratory rate.

Clearance of secretions is more difficult in a recumbent position. The dependent (i.e., usually posterior) lobes accumulate more secretions, whereas the upper parts (i.e., anterior) become dry, rendering the ciliary lining ineffective for clearing secretions and allowing secretions to pool in the lower bronchial tree. The effectiveness of coughing is impaired because of ciliary malfunction and abdominal muscle weakness. Regional changes in the ventilation-perfusion ratio in dependent areas occur when ventilation is reduced and perfusion is increased. This may lead to significant arteriovenous shunting with lowered arterial oxygenation. Atelectasis and hypostatic pneumonia may be the ultimate result of these alterations.

The intercostal and auxiliary respiratory muscles for deep breathing gradually lose their strength and overall endurance. Treatment or prevention involves early mobilization, frequent respiratory toileting, and frequent position changes. A patient in a recumbent position should be persuaded to perform regular pulmonary toileting and deep breathing and coughing exercises and to maintain adequate hydration. An incentive spirometer, chest percussion, and postural drainage with oropharyngeal suctioning can prevent aspiration and atelectasis. The presence of preexisting pulmonary disease requires the use of bronchodilators.

**GENITOURINARY AND GASTROINTESTINAL SIDE EFFECTS**

**Genitourinary Alteration**

Many compromises occur in the physical and metabolic functions of the urinary tract system. Prolonged bed rest contributes to increased incidence of bladder or renal stones and
urinary tract infections. Hypercalciuria is a frequent finding in persons who are immobilized. Other important factors include an altered ratio of citric acid to calcium and an increased urinary excretion of phosphorus. In the supine position, urine must flow uphill from the renal collecting systems to be drained through the ureters. Patients often find it difficult to initiate voiding while supine, a situation that is not ameliorated by reduced intra-abdominal pressure and weak abdominal muscles. Studies have demonstrated less complete voiding occurs in immobilized animals, leading to urinary retention and infection (131).

Incomplete bladder emptying (e.g., in patients with SCI or diabetes mellitus) places the patient at greater risk for stone formation. The most common types of stones are struvite and carbonate apatite, found in 15% to 30% of immobilized patients. Bladder stones allow bacterial growth and decrease the efficacy of standard antimicrobial treatment. Irritation and trauma to the bladder mucosa by stones can encourage bacterial overgrowth and infection. Urea-splitting bacteria then increase the urine pH, leading to further precipitation of calcium and magnesium (131).

Treatment of these problems lies first in prevention, which includes adequate fluid intake to reduce bacterial colonization, use of the upright position for voiding, and scrupulous avoidance of bladder contamination during instrumentation. Other therapeutic approaches might include acidification of the urine through the use of vitamin C, urinary antiseptics, and, in those populations at highest risk for stone formation, a urease inhibitor. Treatment of stones after they have formed may require surgical removal or the use of ultrasonic lithotripsy. Appropriate antibiotic selection based on urine cultures and sensitivity trials is required to eliminate urinary tract infection. If retention is suspected, postvoiding residual volumes should be measured several times a day with ultrasound scanning devices. After stroke or SCI, removal of the Foley catheter and initiation of voiding trails should coincide with sitting and ambulation training. The adverse effects of immobility will aggravate the function of urinary bladder with poor contractility of the detrusor muscle and deficient sphincter opening and coordination.

Gastrointestinal Alterations
Gastrointestinal alterations induced by immobility are easily overlooked. Loss of appetite, slower rate of absorption, and distaste for protein-rich foods all lead to nutritional hypoproteinemia. Passage of food through the esophagus, stomach, and small bowel is slowed in the supine position. An upright position increases the velocity of the esophageal waves and shortens the relaxation time of the lower esophagus (132). Thus, sleeping on two or three pillows with the upper trunk elevated in bed has therapeutic implications in preventing and treating reflux esophagitis. The transit of food through the stomach is 60% slower, and gastric acidity is higher in the supine position than when a person is upright and moving around (133). Peristalsis and passage of food through the small bowel are reduced with prolonged bed rest. Absorption of food is also reduced. For example, calcium absorption during normal activity is 31% of intake and is decreased to 24% during bed rest, and calcium excretion through the bowels in immobility is increased from an average 797 to 911 mg/day (134). This is also one of the reasons why activity is recommended in management of osteoporosis along with calcium and vitamin D.

Constipation is a common complication that results from the interaction of multiple factors. Immobility causes increased adrenergic activity, which inhibits peristalsis and causes sphincter contraction (134). The loss of plasma volume and dehydration aggravate constipation. In addition, the use of a bedpan for fecal elimination places the patient in a nonphysiological position, and the desire to defecate is reduced by social embarrassment. The end result can be fecal impaction, which requires enemas, manual removal, or, in extreme cases, surgical intervention (135).

Prevention of constipation requires an adequate intake of an appealing, fiber-rich diet, including raw fruits and vegetables, and intake of liberal amounts of fluids and regular exercises. Stool softeners and bulk-forming agents are helpful in maintaining bowel function. The use of narcotic agents should be limited because they slow peristalsis. Limited use of glycerin or peristalsis-stimulating suppositories, in combination with a regularly timed bowel program, will further assist in regulating bowel movement (136).

METABOLISM AND ENDOCRINE SYSTEM ALTERATIONS

Daily human energy needs include basal metabolic activity, thermogenesis of food, and the need for ADL and locomotion. It is unclear whether basal metabolism changes during bed rest; this uncertainty stems from inadequate scientific studies on control of factors that could influence the basal metabolic rate during prolonged bed rest. Lean body mass decreases during bed rest, and an equal gain in body fat maintains constant body weight (137). The reduced lean body mass is associated with decreased metabolic activity of muscle, diminished utilization of oxygen and glucose, increased insulin resistance, and deteriorated musculoskeletal function directly affecting overall metabolic, endocrine, and immune system functions (138).

Electrolyte Balance
Prolonged immobility, especially if associated with post-traumatic electrolyte changes, will alter balance of sodium, sulfur, phosphorus, and potassium. A decrease in total body sodium occurs in tandem with the diuresis seen early during bed rest. However, serum sodium levels do not correlate well with the severity of orthostatic intolerance described earlier. Hyponatremia in the elderly is manifested by lethargy, confusion and disorientation, anorexia, and seizures. Potassium levels progressively decrease during the early weeks of bed rest (136). Immobility alone rarely causes serious electrolyte disturbances, aside from the high calcium levels seen in immobilization hypercalcemia (83). Nevertheless, patients with multiple medical illnesses may be seriously affected by even slight electrolyte abnormalities that are caused by immobility.
Hormonal Disorders
A lack of physical activity can cause altered responsiveness of hormones and enzymes. Although they may be clinically undetected during early immobility, numerous changes have been demonstrated to occur in the endocrine system. Significant carbohydrate intolerance has been noted as early as the third day of immobility, and peripheral glucose uptake may decline 50% after 14 days (138). The duration of immobility correlates proportionally with the degree of carbohydrate intolerance. The glucose intolerance induced by bed rest can be improved by isotonic, but not isometric, exercises of the large muscle groups (139). The reason for this intolerance is not lack of insulin but rather increased resistance to the action of insulin, resulting in hyperglycemia and hyperinsulinemia. Possible explanations include a reduction in number or affinity of insulin receptors or postreceptor changes in the target cells. Inactivity appears to cause a reduction in insulin-binding sites, predominantly on the muscle membrane (140).

There are several other hormonal effects, including an increase in serum parathyroid hormone, which is related to hypercalcemia from immobility, although its precise mechanism is unknown (141). Triiodothyronine (T3) blood levels are also elevated during immobility (142). In addition, a variety of alterations have been reported in androgen levels and spermatogenesis, in growth hormone response to hypoglycemia, in levels of adrenocorticotropic hormone, and in catecholamine secretion from the sympathomedullary system (143,144). Serum corticosteroid levels during bed rest are increased, accompanied with increased excretion of urinary cortisol. Studies on bed rest for periods of 1 month or more have found that adrenocorticotropic hormone levels were three times higher than baseline and required about 20 days of physical activity of returning to normal (144). By contrast, prolonged exercise has been shown to increase plasma hydrocortisone levels and decrease plasma norepinephrine levels.

Free Radicals and Oxidative Stresses
Free radicals are highly reactive compounds that cause oxidation of a number of molecules, inducing inflammation in a number of organs. Free radicals produced in muscles have been implicated as a cause of muscular fatigue as well for a decline in mobility in the elderly population. To counteract their negative effects, the body must produce or ingest antioxidants. Antioxidants are enzymes, vitamins, or other types of compounds that are capable of reducing or preventing the harmful effects of free radicals. Vitamins E, C, and A (i.e., retinol) and nonvitamin compounds like glutathione peroxidase and superoxide dismutase are important antioxidants that are reduced in older adults and, possibly, in chronic immobility.

Whether free radicals interfere with metabolic processes in muscle and its contractility, causing functional limitation in mobility and ADLs, is not entirely known. It has been theorized that increased free radicals in elderly subjects may indeed cause decline in mobility function (145). However, there are no studies to prove that free radicals contribute to disuse atrophy and reduced endurance in subjects on bed rest. On the other hand, there is ample evidence that free radicals cause inflammatory conditions such as in atherosclerosis. For example, oxidized low-density lipoprotein (oxLDL) plays a major role in the development of artherosclerosis and subsequent blood vessel inflammation, which is a leading cause of endothelial dysfunction, thromboembolic events, and ultimately to physical limitations and disability (146).

Regular exercises reduce LDL levels, improve insulin sensitivity, and reduce body weight and subsequent risks for cardiovascular disease. Furthermore, moderate-intensity regular exercise can reduce free radicals and positively affect the balance between antioxidants and free radicals reducing oxidative stress and therefore producing beneficial effects on morbidity and mortality of CAD.

Very high intensity, long duration, or sudden onset of strenuous eccentric exercises can cause muscle damage and trigger release of inflammatory cytokines, that is IL-6, and precursors of C-reactive protein, which are associated with development of chronic low-grade inflammatory conditions. However, regular exercise of gradually progressive intensity does not produce elevation of inflammatory markers such as C-reactive protein and actually can reduce their levels (37,145). Study by Tanasecu et al. further elucidated the difference in types of exercises. They found that aerobic exercises, walking and running or aerobic dancing, were associated with a lower risk of fatal or nonfatal myocardial infarctions, while swimming and bicycling were not (146). The mechanism of how physical activity prevents cardiovascular disease is not fully understood although it has a beneficial effect on its risk factors, for example, on cholesterol, HDL, triglycerides, hypertension, and diabetes mellitus.

IMMUNE SYSTEM ALTERATION
Chronic physical and psychological stresses may alter immune system functions such as a reduction of tumor rejection and increase in tumor growth factors and protection against upper respiratory infection or worsening of some autoimmune conditions. In animal studies, exposure to long periods of physical restraint resulted in a significant decrease in number of lymphocytes, T-lymphocytes response, and mononuclear splenic cells as well as increase in plasma corticosteroid levels (147). Initially, it was believed that these alterations are due to immunosuppressive effects of increased corticosterone levels; however, adrenalectomy did not prevent these changes. On the other hand, improvement in immune system function has been found in subjects after moderate types of exercises or physical activities. For example, randomized studies revealed that daily brisk walking of moderate intensity enhances the host protective immune system function and reduces days of illness in half when compared to nonexercising subjects (41). Long-term highly intensive exercises like marathons actually produce immunosuppression and increased susceptibility to upper respiratory infections in these athletes. These effects are not
ameliórate with vitamin C, glutamine, or zinc, although carbohydrates given before or after running a marathon have been partially successful in reducing this infection through unclear mechanisms. It is currently believed that very intensive long-duration exercise may cause striated muscle damage releasing a preponderance proinflammatory cytokines and suppression of immune function as demonstrated in marathon runners and cyclists (148). In immunosuppressed patients with advanced age and HIV infection, moderate exercise has shown to provide an enhancement in amount of immune system function.

THE NERVOUS SYSTEM AND IMMOBILITY

Sensory deprivation is a silent hazard of prolonged bed rest. Healthy subjects placed on strict bed confinement for 3 hours and required to wear gloves, goggles, and earplugs to reduce sensory input experience hallucinations and disorientation. During prolonged bed rest, exposure to social and chronological cues, such as time of day and movement through space, are reduced causing emotional, cognitive, and intellectual declines (149).

Social isolation alone with preserved mobility can cause emotional liability and anxiety but usually does not cause any intellectual alterations. However, prolonged bed rest and social isolation together produce much greater alterations in mental concentration, orientation to space and time, and other intellectual functions. Restlessness, anxiety, decreased pain tolerance, irritability, hostility, insomnia, and depression may occur during weeks of recumbency and social isolation. Furthermore, judgment, problem-solving and learning ability, psychomotor skills, and memory all may be impaired. Perceptual impairment can be altered even after 7 days of immobility (150). Lack of concentration and motivation, depression, and reduced psychomotor skills may drastically affect the patient’s ability to maintain optimal level of functioning and independence. These behavioral effects of immobility may result in a lack of motivation and diminish the patient’s ability to attain optimal healing and restoration of function. Balance and coordination also are impaired after prolonged immobility, and this effect appears to be due to altered neural control rather than muscle weakness (151).

An important strategy in the prevention and treatment of these complications is to apply appropriate physical and psychosocial stimulation early in the course of illness. Options for the treatment of these effects include group therapy sessions, attention to socialization, encouragement of family interaction, and avocational pursuits during evenings and weekends, as well as participation in regular physical activity and exercise. In a multicenter prospective study of older, noninstitutionalized women tested over 6 to 8 years, greater intensity of physical activity was associated with smaller declines in cognitive function, even when controlling for age, estrogen use, and comorbidities. The authors concluded that physical activity prevents decline in cognitive function in older women (152).

CONCLUSION

In 1862, the English surgeon John Hilton advocated bed rest as a basic physiologic approach in the treatment of human illness. Since that time, bed rest has often been used indiscriminately in the treatment of acute and chronic illnesses. The complications of prolonged bed rest have been increasingly recognized and reported since the mid-1940s. After World War II, clinical investigators, starting with Deitrick et al. (153) in 1948, have shown that prolonged bed rest may cause multiple adverse effects in a number of organs and systems. During the 1960s and 1970s, studies on astronauts greatly advanced our knowledge of the deleterious effects of bed rest and weightlessness, making clinicians more aware of a wide range of adverse effects associated with prolonged bed rest and immobility. As a result, a great number of new epidemiologic and randomized studies in the last three decades have been conducted, which in a great preponderance have demonstrated significant benefits of physical activity and exercises on cardiopulmonary, musculoskeletal, and total body functions.

A new era of advocating exercise and cardiovascular conditioning has begun. Studies have indicated that prolonged bed rest and sedentary lifestyles have negative effects on health and function. These effects are magnified in persons with neurologic disease or in the elderly. The principles long advocated by rehabilitation medicine have contributed significantly to the current philosophy on the use and misuse of immobility.

REFERENCES


CHAPTER 49

Treatment of the Patient with Chronic Pain

We are but dwarfs standing on the shoulders of Giants (1,2).

Pain is purely subjective, difficult to define, and often hard to characterize or interpret. It is currently defined as an unpleasant sensory and emotional response to a stimulus associated with actual or potential tissue damage or described in terms of such damage (3–5). However, pain has never been shown to be a simple function of the amount of physical injury; it is extensively influenced by anxiety, depression, expectation, and other psychological and physiological variables. It is a multifaceted experience, an interweaving of the physical characteristics of the stimulus with the individual’s motivational, affective, and cognitive functions. The pain experience is in part behavior based on an interpretation of the event, influenced by present and past experiences.

Acute pain is a biologic symptom of an apparent nociceptive stimulus, such as tissue damage that is due to disease or trauma that persists only as long as the tissue pathology itself persists. The pain may be highly localized or may radiate. Acute somatic pain may be well localized, aching, and sharp, whereas acute visceral pain may be burning, cramping, and radiating. It is usually associated with upregulated sympathetic activity, tachycardia, hypertension, tachypnea, increased metabolic rate, and hypercoagulability. Acute pain is generally self-limiting, and as the nociceptive stimulus lessens, the pain decreases. Acute pain usually lasts a few days to a few weeks (4). If it is not effectively treated, it may progress to a chronic form.

Chronic pain is a disease process in which the pain is a persistent symptom of an autonomous disorder with neurologic, psychological, and physiologic components. Differing significantly from acute pain, it is defined as pain lasting longer than anticipated (greater than 3 months) within the context of the usual course of an acute disease or injury. The pain may be associated with continued pathology or may persist after recovery from a disease or injury. Due to the complex nature of chronic pain, many other definitions have been proposed. Some include factors such as the persistence of pain despite extraordinary measures in a nonacute setting or pain that is without apparent biological value (6). Operational definitions include aspects such as pain sensation, pain behavior, functional status, emotion, and somatic preoccupation (6). As with acute pain, treatable chronic pain that is due to organic disease is managed by effectively treating the underlying disorder; however, no such identifiable organic disease may be evident. Chronic pain can mimic the qualities of acute pain except that associated signs of autonomic nervous system response may be absent, and the patient may appear exhausted, listless, depressed, and withdrawn. Chronic pain can have exacerbations that are triggered by progression of organic pathology, physiologic stress, or worsening emotional, social, and psychiatric problems. As these problems subside, the pain may improve. Chronic pain may also be highly persistent and reported as severe for years without remission.

Proper management of pain requires an understanding of its complexity and knowledge of the nonneurologic factors that determine its individual expression. The treatment of pain with physical modalities is as ancient as the history of humanity, but the use of interdisciplinary rehabilitation techniques has gained acceptance only within the past few decades.

EPIDEMIOLOGY

Nearly everyone experiences acute pain. Its incidence approximates the cumulative total of all acute diseases, trauma, and surgical procedures. In studies of the general population, patients have identified the head and lower limbs as the most common sites of acute pain and have identified the back as the most common site of chronic pain (7).

Chronic pain is less frequently experienced but is reaching epidemic proportions in the United States. There are more than 36 million individuals with arthritis, 70 million with episodic back pain, 20 million with migraine headaches, and additional millions with pain that is due to gout, myofascial pain syndromes, phantom limb pain, and complex regional pain syndromes (CRPS) (8–11). The pain resulting from cancer afflicts approximately 1 million Americans and 20 million individuals worldwide. Moderate to severe pain occurs in about 40% of patients with intermediate-stage cancer and in 60% to 80% of patients with advanced cancer (12–14). Back pain, as a general condition, episodically affects nearly 75% of the population in most industrialized nations. It is estimated that at least 10% to 15% of the working population of industrialized nations are affected by back pain each year (9,10).
ETIOLOGY

Chronic pain is not merely a physical sensation. In the affective component of chronic pain, most patients show a degree of depression, and many show anger, jealousy, and anxiety. For many individuals, depression is the primary factor in the perception or experience of pain. Fifty to seventy percent of patients with chronic pain have either a primary depression or a depression secondary to their pain syndrome. Chronic pain, with accompanying depression, often leads to extensive periods of reduced productivity as well as inactivity. Prolonged inactivity alters cardiovascular function, impairs musculoskeletal flexibility, and causes abnormal joint function (15–18). Prevention involves the encouragement of patient activity as soon as it is reasonable.

The motivational component of chronic pain is concerned with the vocational, economic, and interpersonal reinforcement contingencies that contribute to the learning of pain behavior and the maintenance of chronic pain. More than 75% of patients with chronic pain display adverse behavioral characteristics, including problems with job or housework, leisure activities, sexual function, and vocational endeavors (19). Sleep disturbances and depression are highly associated with chronic pain. The patient also may have significant functional limitations as a result of multiple previous surgeries with little success and prolonged convalescence, disuse/physical deconditioning syndrome, or opioid medication (20).

Chronic pain’s cognitive component involves how patients think and the part that pain plays in their belief system and view of self. The more the patient perceives pain as a signal, requiring a reduction of activity and protection of the affected part, the more difficult it is for the physician to achieve compliance with exercise, stretching, and other elements of the treatment program. The memories of pain from acute pain episodes may significantly hinder a patient's recovery and contribute to chronic pain syndrome (21–23). Pain is often the result of sensory input, affective state, cognition, motivational, and memory factors, which requires a multidimensional evaluation process, including treatment interventions directed at those components most responsible for the pain experience (24–26).

Pain Pathways

Pain is a central perception of multiple primary sensory modalities. This interpretive function is complex, involving psychological, neuroanatomic, neurochemical, and neurophysiologic factors of both the pain stimulus and the memory of past pain experiences. The peripheral mechanisms for sensing and modulating pain have been extensively studied during the past 30 years. The pathways for pain sensation, from the initial stimulus of the nociceptors to the central nervous system, are summarized in Figure 49-1 (27–31). There appear to be several descending systems that play a role in control of the modification of the ascending pain pathways, which are summarized in Figure 49-2 (28,31–35).

Polymodal nociceptors respond to stimuli that damage tissue. This stimulation results in impulses ascending in the A-delta or C fibers to the marginal layers of the dorsal horn of the spinal cord. The A-delta fibers primarily synapse in laminae I and V, whereas C fibers synapse primarily in lamina II. Deeper regions of the dorsal horn may be polysynaptically involved in the processing of noxious stimuli.

The major ascending nociceptive pathways are the spinothalamic and spinoreticular tracts, which involve both oligosynaptic and polysynaptic neurons. These oligosynaptic pathways are fast conducting, with discrete somatotopic organization resulting in rapid transmission of nociceptive information regarding site, intensity, and duration of stimulus. Further, the oligosynaptic tracts provide somatic information by way of the posterior ventral nuclei of the thalamus to the postcentral cortex. The sensory discriminatory characteristics are delineated from the neospinothalamic portion of the lateral spinothalamic tract and the nonproprioceptive portion of the dorsal columns.
Polysynaptic pathways are slow conducting, with a lack of somatotopic organization resulting in poor localization as well as dull aching and burning sensations. The nociceptive impulses transmitted through this system result in suprasegmental reflex responses related to ventilation, circulation, and endocrine function. Pathways contributing to this slow-conducting system are the paleospinothalamic tract, spinoreticular, spinocollicular, and the dorsal intercornual, as well as the spinomesencephalic tracts. The polysynaptic tracts form the brain stem reticular activating system with projections to the medial and interlaminar nuclei of the thalamus. From these nuclei, diffuse radiation occurs to the cerebral cortex, limbic system, and basal ganglia.

There are multiple levels of processing and convergence of nociceptive information in its ascending transmission to the cerebral cortex. In addition, there appear to be several descending pain control systems that play a role in the control and modification of the ascending pain pathways. The most complete studies have been of the periaqueductal gray (PAG) region of the midbrain. Stimulation of the PAG neurons and the subsequent descending impulses result in release of endogenous opioids at the nucleus raphe magnus (NRM) and nucleus locus ceruleus (NLC). Endogenous opioids activate the serotonergic cells in the NRM and norepineurergic neurons in the NLC. The axons of both these monoaminergic neurons descend in the dorsolateral tract to interneurons, predominantly in laminae I, II, and V. These monoamines activate opioid-secreting interneurons. The morphine-like transmitter released may vary, depending on what type of receptor in the periphery has been activated. Both A-delta and C afferent fibers are inhibited by descending influences in the dorsal horn. The opioid inhibitory interneurons may be influenced by intersegmental and descending pathways, but the intersegmental and segmental mechanisms have not been established. These interneurons may function either by presynaptic inhibition on the terminals with the primary nociceptive afferents preventing the release of substance P or by postsynaptic inhibition on second-order neurons. Cells in the raphe magnus are activated by ascending sensory pathways transmitted to the reticular formation as well as by descending input from cells in the PAG region.

Other descending monoamine systems include locus ceruleus to the dorsal horn interneurons, nucleus reticularis, magnicellularis to the dorsal horn interneurons, and the mesencephalic lateral reticular formation to the dorsal horn interneurons. It has been suggested that monoamines are involved with supraspinal and spinal nociceptive mechanisms. Hormonally based descending pathways have been described but are poorly understood.

**Gate Control Pain Theory**

The gate control theory of pain was developed by Melzack and Wall to account for mechanisms by which other cutaneous stimuli and emotional states alter the level of pain (36). They suggested that within the substantia gelatinosa of the dorsal horn, there are interneurons that presynaptically inhibit transmission of nociceptive information to the ascending tracts. These interneurons are activated by large-diameter afferents and inhibited by small-diameter afferents. In addition, they suggested that the brain exerts descending control on this system, relying on the fact that cognitive factors are known to influence pain behavior.

Several studies have failed to provide support for the gate control theory. It remains significant, although incorrect in detail, in hypothesizing that nociceptive pain undergoes dynamic integration and modulation. The gate control theory of pain has altered the concept of pain as solely an afferent sensory experience, broadening the concept to include the affective and motivational factors involved in the human pain experience (37). The gate control theory has been modified extensively during the past 40 years (38). It still represents the first attempt to describe a pain-modulating system that responds to input by noxious stimuli, innocuous afferent impulses, and descending control.

**Biochemical Pain Theory**

In the peripheral nervous system, the extent and duration of the response to the stimulus may be modified at the biochemical level within the neurons in the pain-conducting pathway;
this neural plasticity is manifest by short-term modulation and long-term modification of the excitability of primary sensory and central neurons. Chronic pain can alter the nociceptive pathways by activation-dependent plasticity (resulting in autosensitization and windup), modulation-dependent plasticity (resulting in heterosensitization and central sensitization), or modification of primary sensory neurons, which results in pain hypersensitivity. NMDA receptors may be involved in synaptic plasticity and postsynaptic sensitization in the substantal gelatinoa and higher brain centers. This modulation of the neurotransmitters that affect the flow of information can induce the stimulation of polysynaptic neuron, or it can pathologically alter the anatomic features of these neurons and their interconnections such that the normal nociceptive stimulus response amplifies is amplified (39).

The biochemical theory of pain has evolved since the discovery of endorphins. The endogenous opioid system consists of three families of opioid peptides: beta-endorphin, enkephalin, and dynorphin/neoendorphin. The beta-endorphins are primarily concentrated in the pituitary and the basal hypothalamus. The other endogenous opioids are distributed extensively in the central nervous system. The dynorphins/neoendorphins and enkephalins are found in the caudate nucleus, amygdala, PAG matter, locus ceruleus, and dorsal horn of the spinal cord. In addition, the enkephalins are found in the NRM and the thalamic periventricular nuclei. The dynorphins/neoendorphins are found in the hypothalamus and substantal nigra (32).

Endogenous opioids are involved in analgesia, as well as multiple other clinical events. There are at least seven different opiate receptors, of which the mu, delta, and kappa appear to be involved in analgesia (40). The others are associated with such functions as respiration, appetite, hallucinations, dysphoria, immune function, temperature regulation, memory, and blood pressure control. Beta-endorphins may function in the modulation of local blood flow and immune function (41). The discovery of multiple opiate receptors and multiple endogenous opioid compounds provides an explanation for the multiple effects of the endogenous opioids. Within the peripheral and central nervous systems, the enkephalins act as neurotransmitters and the beta endorphins act predominately as hormones. Endogenous opioids are only one part of a complex modulatory system involved in the collating, processing, and filtering of information concerning tissue damage (42).

Neural plasticity has been observed in patients with chronic pain. Persistent pain has been shown to increase the strength of the polysynaptic response as well as the anatomic configuration of the pain conduction pathway. This may result in long-term alteration in the connectivity and organization of pain pathways. The consequence of this alteration is a “pain memory” that is evoked with minimal stimulation and results in a maximal nociceptive response (39,43). Of note, neuroplasticity allows for secondary development of hyperalgesia with chronic opioid usage (44). The modulation of pain may occur via pronociceptive or antinociceptive neurotransmitters at multiple neurons along the pain conduction pathway. Pronociceptive neurons include W/W glutamate, aspartate, and substance P. Antinociceptive transmitters include gamma-aminobutyric acid (GABA), serotonin, and acetylcholine (45,46). Other possible neural peptides having analgesia or antinociceptive properties are calcitonin, cholecystokinin, somatostatin, and neurotensin (41).

**Chronic Pain Theory**

The chronic pain theory encompasses many of the physical, motivational, cognitive, and affective components of pain. The anatomic pain pathways are relatively clear and represent a mechanism for nociceptive pain in the animal model. Multiple pain mechanisms exist in the human model because of the complex integration of nociceptive stimuli, conceptual and judgmental factors, sociocultural influences, and the motivational and emotional states of the individual. Pain mechanisms include nociceptive pain, neuropathic pain, central pain, psychogenic, and operant pain. The human perception and reaction to pain are a blending of these mechanisms. The nociceptive pain mechanism is detailed in the pain pathways, as previously described, and represents pain originating from tissue damage, such as pain from cancer, degenerative joint disease, myofascial pain, and trauma. Neuropathic pain can originate from peripheral neurologic disease or injury. Central pain originates from denervation occurring after a cerebrovascular accident, spinal cord injury, or amputation. This pain may be due to a loss of the peripheral modulating influences on the central nervous system resulting in an unmodulated activity of afferent A-delta and C fibers.

Psychogenic pain is the interpretation of emotional distress as aversive and unpleasant sensation and its description in terms of pain language and behavior. The psychological states that are interpreted in this manner include anxiety, neurosis, hysteria, and depression. This mechanism of pain is often overlooked in patients with chronic pain.

Operant or learned pain behavior is often a major factor in chronic pain. Although the initial precipitating event causing the pain may be quite minor, the pain behavior is often long-lasting, owing to reinforcement by environmental influences. Pain behavior may be directly reinforced by family and physician attention or the delivery of medication. Indirect reinforcement (physical or psychological demands) occurs with avoidance of aversive consequences, which would have to be met if there were no pain. Operant pain behavior is also reinforced through a punishment cycle, when an injured party is overprotected and “punished” by outside factors if he or she begins to function more independently.

**Genetics and Pain**

The genetic determinants have been evaluated in humans with clinical pain (47). The variability in pain syndromes has been associated with inherited genetic factors in back pain (48), fibromyalgia (FM) (49), menstrual pain (50), and migraine (51). Family and twin heritability estimates indicate genetic as well as significant environmental factors modulating pain (52,53). Polymorphic pain genes have been associated with
congenital insensitivity to pain (54), drug metabolism due to cytochrome P450 (55), familial hemiplegic migraine (52,56), FM (57), and reflex sympathetic dystrophy (58).

Gene therapy and other advances in molecular medicine may offer a means of enhancing antinociceptive receptors (cannabinoid—1 and 2, acetylcholine—m and n, opioid—µ and κ, adrenergic—α1 or blocking pronociceptive receptors (neurokinin—1-α-amino-3-hydroxy-5-methyl-4-isoxazolepropionate [AMPA], n-methyl-d-aspartate [NMDA]); acting directly on the calcium channels of pain fibers; or acting directly on membrane receptors, protein C-GAMA, or other areas of the central nervous system involved in the transmission of pain (59,60). Definitive applications of gene intervention in pain control remain to be developed.

**Resolution of Pain**

Acute pain is the result of tissue damage in which the initial pain leads to an increase in anxiety, which magnifies the pain experience. The amount of anxiety generated and possibly the level of pain seem to be more influenced by the setting in which the pain develops rather than personality variables. With the healing process comes a reduction or termination of the anxiety and acute pain perception. When acute pain, which functions as a warning signal, fails to respond to treatment with conventional medical therapies, illness behavior and chronic pain develop. The anxiety characteristic of acute pain is replaced by depression with hopelessness, helplessness, and despair. When pain relief fails, physical activities decrease and suffering and depression increase.

Acute pain usually resolves when the source of nociception is removed or cured and resolves quickly with application of appropriate pharmacologic or regional analgesic therapy. The cause of acute pain can be documented by physical examination findings and diagnostic procedures. When indicated, appropriate operative intervention can be performed on the basis of these findings. A short course of analgesic medication usually controls postoperative pain, and a return to full, painless function can be anticipated in a matter of weeks. Acute pain control requires the administration of an efficacious analgesic dosage. Too little analgesia promotes suffering and anxiety, thus defeating the purpose of prescribing medications. Fear of drug addiction contributes to the underutilization of analgesic medications, and physicians tend to undermedicate in terms of frequency and dosage of pain medications (61,62). By prescribing low oral doses of opioids at infrequent intervals, physicians inadvertently force patients to adopt pain behavior in order to obtain adequate opioid analgesia. Pain behavior is characterized by high verbalization of pain, dependency, and the inability to work. Addiction in the acute pain situation is very rare, approximating less than 0.1% (63,64). Unfortu8nately, a significant minority of acute pain patients continues to experience pain, which may progress into a more complex disease entity. Pain, a symptom of physiologic malfunction, now becomes the disease itself. Chronic pain represents a complex interaction of physical, psychological, and social factors in which the pain complaint is a socially acceptable manifestation of the disease. The etiology of chronic pain may be persistent nociceptive input, such as arthritis or terminal cancer; psychological disorders, such as anxiety, depression, and learned behavior; or social factors, such as job loss, divorce, and secondary gain.

The optimal treatment for chronic pain is prevention. Once the disease state of chronic pain commences, reinforcers such as monetary compensation, presence of job-related problems, manipulation of the environment to satisfy unmet needs, and retirement from the competitive work obstruct complete disease resolution. Therapies designed for acute pain are often contraindicated for chronic pain.

Prevention of chronic pain requires identifying contributing factors and resolving them early in the acute stage. Aspects worthy of attention include psychological stress, drug or alcohol abuse, and poor posture or muscle tone, as well as significant psychological and operant pain mechanisms. Physicians should set a reasonable timeframe for the resolution of the acute pain process. Patients should be advised when the pain medication will no longer be needed and that those medications that are no longer effective will not be continued. The patient’s attention should be directed to a gradual return of full activity on a prescribed schedule. Follow-up appointments should be planned at specified intervals so that the patient does not need to justify a visit. Work intolerance and job conflicts should be resolved.

**Pain-Reinforcing Factors**

Chronic pain syndrome is a learned behavior pattern reinforced by multiple factors. These behaviors are frequently found in individuals who are depressed, are inactive, and lack the skills or opportunity to compete in the community. These environmental factors promote pain behavior, regardless of the etiology of the pain, thereby distinguishing the patient with chronic pain from the population at large. Patients often develop a new self-image and see themselves as disabled by their pain. This self-perceived disability justifies their inactivity and manipulation of others and attempts to collect compensation. The typical patient often has been unemployed, has low job satisfaction, or has been on sick leave for long periods of time (65–68). Our data indicate that individuals who have been removed from the labor market because of pain for less than 6 months have a 90% chance of returning to full employment; those removed from the labor force because of pain for more than 1 year have less than a 10% chance of return to full employment (66).

Individuals with chronic pain syndrome receive gains from their pain behavior; hence, they continue this behavior to maintain those positive reinforcers. Physicians reinforce the pain behaviors by lacking knowledge of this chronic disease process, failing to identify the chronic pain behavior and prolonging prescription of inappropriate medications, inactivity, and work limitations. The physician’s failure to acknowledge and direct the patient toward recovery tends to validate the chronic pain syndrome by providing an undiagnosable and untreatable problem. Family members also frequently reinforce the chronic pain behavior. They allow the individual
to become inactive and cater to the patient’s requests and needs over prolonged periods of time. In some instances, patients with chronic pain provide role models for pain or disability behavior for other family members (69,70).

**Socioeconomic Factors in Chronic Pain**

Socioeconomic status likely plays an important role in the prevalence and risks of developing chronic pain. Persons with low socioeconomic status have an increased incidence of a variety of musculoskeletal pains including low back pain, as well as increased incidence of psychological problems. In a large cohort study of over 9,000 adults, lowest adult socioeconomic status (defined as having a “nonskilled” manual profession) was associated with a nearly threefold relative risk of chronic widespread pain, and at least a 1.5-fold relative risk for regional pain problems (71). Conversely, higher socioeconomic status (e.g., professional, managerial, and technical careers) was associated with a higher incidence of forearm pain (71).

**Worker’s Compensation**

In 1911, worker’s compensation laws were enacted in the United States that required employers to assume the cost of occupational disability without regard to fault. These laws have dramatically influenced the recovery from injury. In many instances, they have become counterproductive; financial compensation may discourage return to work, the appeal process may increase disability, an open claim may inhibit return to work, and recovering patients may be unable to return to work. Often, the accident and resulting symptoms represent the patient’s solution to life’s problems (72). The pain literature suggests an enhanced pain experience and reduced treatment efficacy in patients with chronic pain who are receiving financial compensation (73).

**Litigation**

Disability, along with pain and suffering, greatly determines the amount of compensation awarded in worker’s compensation cases. The patient/client’s pain behavior may be reinforced, maximized, and groomed with the hope of a large cash settlement. As a result of this reinforcement, the pain behavior develops into a learned response. The pain also becomes the disability for which the patient/client is seeking compensation. Therefore, a learned behavior becomes a determining factor in the amount of compensation awarded (74,75).

Alteration of the disability laws could decrease the number of acute pain patients who develop the behavioral disease of chronic pain syndrome. Changes that might discourage the development of chronic pain include allowing an injured worker to continue working at a job he or she is physically able to accomplish during the recuperation period, rapid adjudication of disability and compensation claims, and physicians restricting the patient’s use of addicting and depressant medication to less than 1 month. The extensive use of conservative intervention to include physical therapy and stress management early in treatment also could prevent the emergence of a chronic pain syndrome (76,77). For additional information on disability determination and medical-legal aspects, the reader is referred to Chapters 10 and 21.

**Complications of Chronic Pain**

Chronic pain is an elusive disease complicated by iatrogenic, idiopathic, and psychosocial factors. These complications encompass physical, psychological, and environmental issues.

**Physical**

The patient with chronic pain often develops secondary pain loci as a result of inactivity. Decreased range of motion, myofascial pain, and weakness because of disuse also may develop (78).

Medically induced drug addiction and dependence are particularly serious problems for both patients with chronic pain and their physicians (79,80). It has been estimated that 30% to 50% of patients with chronic nonmalignant pain have a significant drug dependency problem (81–85). Substance abuse, dependence, and addiction are negative prognostic factors in outcome studies (86). Treatment of the drug impairment is essential and results in a greater overall improvement in functioning (87).

**Psychological**

Depression is a common complication noted with chronic pain (88). These patients often manipulate family, friends, and coworkers to achieve secondary gain. Chronic pain causes considerable distress to spouses and family members (89), and sexual dysfunctions (90) are common. Twenty-five percent of spouses report clinical depression and more than 35% rate their marriage as maladjusted (91). Insomnia is common and difficult to treat, but appropriate attention to sleep disturbances may help with other symptomatology (92).

**Environmental**

Nearly one-third of the American population has persistent or recurrent chronic pain. One-half to two-thirds of these individuals are partially or totally disabled for varying lengths of time. It is estimated that 80 billion dollars a year (adjusted for the year 2000) are spent on medical needs, lost working days, and compensation (11,93). Data compiled in 1982 suggest that lost wages and social support systems cost the taxpayer $15,000 to $24,000 per chronic pain patient per year (94).

**Anatomy of Pain**

Knowledge of the peripheral anatomy of the human body is essential in evaluating the complex problems found in a patient with pain. The clinical problem is often reduced to
the simple question, “Is the pain in an area supplied by a single nerve root, a single peripheral nerve, or a branch of a peripheral nerve?” A physical examination to evaluate pain, weakness, and their distribution often leads the clinician to better localize the nerves involved in the patient’s report of pain.

**Somatic Innervation**

A difficulty encountered in the diagnosis of pain is secondary to the overlap of cutaneous fields of segmental and peripheral nerves, as well as the overlapping innervation of muscle and bone. Single cutaneous nerves innervate sharply defined regions with little overlap, but these fibers regroup in the peripheral nerve and are again redistributed in the brachial or lumbosacral plexus. This makes it impossible to follow individual fibers from the dorsal roots to the areas innervated by the individual cutaneous nerve. Adjacent cutaneous nerves may be supplied by fibers from more than one spinal nerve (Fig. 49-3) (95). Anatomists and physicians have attempted to define areas of the skin, muscle, and bone that are of the exclusive domain of a single spinal cord root, as well as the areas of overlap. The size of these areas varies from nerve to nerve and from individual to individual. As with all knowledge of the human body, anatomy of sensation and pain rests on a foundation of original work performed by investigators too numerous to list. The figures in this chapter have taken some of the best work in the field and compiled it into a comprehensive review (Figs. 49-4 through 49-17). In delving into the literature in this area of anatomy, one cannot help but feel that he or she is standing on the shoulders of giants without whose work the figures in this chapter would be impossible.

**Dermatomes**

Different techniques were used in determining the site and extent of innervation. Foerster’s (96) data were based on remaining sensibility of the skin innervated by an intact posterior root isolated by severing several nerve roots above and below. In determining the field of unaltered sensation, Foerster demonstrated considerable overlap of contiguous nerve roots. Similar techniques were used by Bing (97), the Armed Forces Institute of Pathology, and Haymaker and Woodhall (98). Keagan and Garrett (99) used hyposensitivity to pin scratch in cases of herniated intervertebral discs, which resulted in the most extensive dermatome map. Keagan and Garrett contended that no sensory overlap exists between dermatomes, which is contrary to most investigators’ experiences. Richter and Woodruff used the electrical skin resistance method (100,101) over sympathectomized areas of the skin. It is noted that dermatomes of almost all cutaneous nerves are beyond the anatomic boundaries noted on gross dissection. Clinical data suggest that areas of sensory deficit extend further proximally than mapped by Foerster, but not as far proximally as suggested by Keagan and Garrett. For this reason, the dermatomes used in the figures are derived from an interpolation of areas determined by personal experience and multiple other authors (96–104).

**Myotomes**

The determination of muscular innervation is the result of the analysis by many observers of traumatic and surgical outcomes evaluated during the postoperative and recovery periods (98,104,105). The results of these analyses are provided in detail in Tables 49-1 and 49-2. The myotomes used in these drawings are derived from the experience of multiple clinicians.

(Text continues on page 1287)


FIGURE 49-8. Dermatome, myotome, and sclerotome distribution for C7. **Dermatome:** midhand and middle finger. **Myotome:** paraspinals, elbow extensors, forearm pronators, and wrist extensors. **Sclerotome:** bones—scapula, humerus, radius, ulna, and middle fingers; joints—facet and wrist; ligaments—longitudinal, ligamentum flavum, and interspinous.

FIGURE 49-9. Dermatome, myotome, and sclerotome distribution for C8. **Dermatome:** medial forearm and medial hand. **Myotome:** paraspinals, elbow extensors, wrist flexors, grip, finger abduction, finger flexion, finger adduction, finger opposition, and finger extension. **Sclerotome:** bones—vertebra and periosteum, ulna, and medial fingers; joints—facet and wrist; ligaments—longitudinal, ligamentum flavum, and interspinous.


 CHAPTER 49  |  TREATMENT OF THE PATIENT WITH CHRONIC PAIN


**FIGURE 49-16.** Dermatome, myotome, and sclerotome distribution for S1. **Dermatome:** posterior lower leg, lateral lower leg, and lateral foot. **Myotome:** hip extensors, hip abductors, knee flexors, ankle evertors, ankle plantar flexors, and toe dorsiflexors. **Sclerotome:** bones—vertebra and periosteum, sacrum, ischium, femur, tibia, and mid-foot; joints—sacroiliac, hip, knee, ankle, and large toe; ligaments—longitudinal, ligamentum flavum, and interspinous.

**FIGURE 49-17.** Dermatome, myotome, and sclerotome distribution for S2. **Dermatome:** posterior upper leg and posterior lower leg. **Myotome:** knee flexors, ankle plantar flexors, toe dorsiflexors, toe abduction, and toe adduction. **Sclerotome:** bones—sacrum, coccyx, distal fibula and lateral foot; joints—sacroiliac, ankle, and toes; ligaments—longitudinal, ligamentum flavum, and interspinous.
Sclerotomes

Areas of segmental innervation of bone (sclerotomes) are closely linked with muscle innervation. Injury to bone, ligament, tendon, fascia, and other mesodermal structures of the body may result in pain referred in a sclerotomal distribution (98, 106). Notable references for peripheral nerve innervation of the skeleton include Haymaker and Woodhall (98), Dejerine (107), Foerster (108), and Brash (109). The peripheral nervous system of the skeleton is closely linked to muscle innervation. Most of the bones of the skeleton receive their innervation

<table>
<thead>
<tr>
<th>Muscle</th>
<th>Peripheral Nerve</th>
<th>Spinal Segment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levator scapulae</td>
<td>Trapezius</td>
<td>C3, C4</td>
</tr>
<tr>
<td>Serratus anterior</td>
<td>Long thoracic nerve</td>
<td>C3, C4, C5, C6, C7</td>
</tr>
<tr>
<td>Rhomboid major/minor</td>
<td>Dorsal scapular</td>
<td>(C4), C5</td>
</tr>
<tr>
<td>Supraspinatus</td>
<td>Suprascapular</td>
<td>C5, C6</td>
</tr>
<tr>
<td>Infraspinatus</td>
<td>Suprascapular</td>
<td>C5, C6</td>
</tr>
<tr>
<td>Latissimus dorsi</td>
<td>Thoracodorsal</td>
<td>C6, C7, C8</td>
</tr>
<tr>
<td>Teres major</td>
<td>Lower subscapular</td>
<td>C5, C6, C7</td>
</tr>
<tr>
<td>Teres minor</td>
<td>Axillary</td>
<td>C5, C6</td>
</tr>
<tr>
<td>Pectoralis major (clavicular)</td>
<td>Lateral pectoral</td>
<td>C5, C6, C7</td>
</tr>
<tr>
<td>Pectoralis major (sternal)</td>
<td>Lateral/medial pectoral</td>
<td>C6, C7, C8, T1</td>
</tr>
<tr>
<td>Deltoid</td>
<td>Axillary</td>
<td>C5, C6</td>
</tr>
<tr>
<td>Coracobrachialis</td>
<td>Musculocutaneous</td>
<td>C6, C7</td>
</tr>
<tr>
<td>Biceps brachii</td>
<td>Musculocutaneous</td>
<td>C5, C6</td>
</tr>
<tr>
<td>Brachialis</td>
<td>Musculocutaneous(^a)</td>
<td>C5, C6</td>
</tr>
<tr>
<td>Triceps</td>
<td>Radial</td>
<td>C6, C7, C8</td>
</tr>
<tr>
<td>Anconeus</td>
<td>Radial</td>
<td>C6, C7, C8</td>
</tr>
<tr>
<td>Brachioradialis</td>
<td>Radial</td>
<td>C5, C6</td>
</tr>
<tr>
<td>Extensor carpi radialis longus</td>
<td>Radial</td>
<td>C6, C7</td>
</tr>
<tr>
<td>Supinator</td>
<td>Radial</td>
<td>C5, C6, (C7)</td>
</tr>
<tr>
<td>Extensor digitorum communis</td>
<td>Radial</td>
<td>C7, C8</td>
</tr>
<tr>
<td>Extensor digiti minimi</td>
<td>Radial</td>
<td>C7, C8</td>
</tr>
<tr>
<td>Extensor carpi ulnaris</td>
<td>Radial</td>
<td>C7, C8</td>
</tr>
<tr>
<td>Abductor pollicis longus</td>
<td>Radial</td>
<td>C7, C8</td>
</tr>
<tr>
<td>Extensor pollicis brevis</td>
<td>Radial</td>
<td>C7, C8</td>
</tr>
<tr>
<td>Extensor pollicis longus</td>
<td>Radial</td>
<td>C7, C8</td>
</tr>
<tr>
<td>Extensor indicis</td>
<td>Radial</td>
<td>C7, C8</td>
</tr>
<tr>
<td>Pronator teres</td>
<td>Median</td>
<td>C6, C7</td>
</tr>
<tr>
<td>Flexor carpi radialis</td>
<td>Median</td>
<td>C6, C7, C8</td>
</tr>
<tr>
<td>Palmaris longus</td>
<td>Median</td>
<td>C7, C8, T1</td>
</tr>
<tr>
<td>Flexor digitorum superficialis</td>
<td>Median</td>
<td>C7, C8, T1</td>
</tr>
<tr>
<td>Flexor digitorum profundus I and II</td>
<td>Median</td>
<td>C7, C8, T1</td>
</tr>
<tr>
<td>Flexor pollicus brevis</td>
<td>Median</td>
<td>—</td>
</tr>
<tr>
<td>Flexor pollicis longus</td>
<td>Median</td>
<td>C7, C8, T1</td>
</tr>
<tr>
<td>Pronator quadratus</td>
<td>Median</td>
<td>C8, T1</td>
</tr>
<tr>
<td>Abductor pollicis brevis</td>
<td>Median</td>
<td>C8, T1</td>
</tr>
<tr>
<td>Opponens pollicis</td>
<td>Median</td>
<td>C8, T1</td>
</tr>
<tr>
<td>Lumbrical I and II</td>
<td>Median</td>
<td>C8, T1</td>
</tr>
<tr>
<td>Flexor carpi ulnaris</td>
<td>Median</td>
<td>C7, C8, T1</td>
</tr>
<tr>
<td>Flexor digitorum profundus III and IV</td>
<td>Ulnar</td>
<td>C7, C8, T1</td>
</tr>
<tr>
<td>Abductor digit minimi</td>
<td>Ulnar</td>
<td>C8, T1</td>
</tr>
<tr>
<td>Palmar/dorsal interossei</td>
<td>Ulnar</td>
<td>C8, T1</td>
</tr>
<tr>
<td>Lumbrical III and IV</td>
<td>Ulnar</td>
<td>C8, T1</td>
</tr>
<tr>
<td>Adductor pollicis</td>
<td>Ulnar</td>
<td>C8, T1</td>
</tr>
</tbody>
</table>

Spinal segments in parentheses signify those that are occasionally present; those in boldface are the major segments supplying the muscle designated. Segments not in boldface are usually present but to a minor degree.

\(^a\)The small branch from the radial nerve also innervates the brachialis muscle.
from nerve twigs of the attached muscles. Some parts of the skeleton, especially the joints, receive branches directly from nerve trunks (99).

**Autonomic Innervation**

The peripheral component of the autonomic nervous system is concerned with innervation of visceral glands, blood vessels, and nonstriated muscle. The relationship to pain has been confined to the visceral components in this section (104, 105, 110–112).

**DIAGNOSTIC AND CLINICAL EVALUATION**

In establishing the etiology of pain, it is essential to consider its location, characteristics, its chronology, the limitations it imposes on the patient, and the results of previous therapy. This is accomplished by a thorough pain evaluation, including a detailed history, comprehensive physical examination, and review of appropriate diagnostic tests. When making the treatment plan, practitioners should keep in mind that many radiographic findings do not correlate with the patient’s clinical presentation, and that most physical exam findings can be nonspecific. Most pain patients present a complex array of physical, motivational, cognitive, and affective manifestations and therefore require detailed psychological and social evaluations. In the intensive effort to integrate these evaluations, it should be remembered that acute and ominous processes such as fractures, infections, and tumors must be ruled out, even if the patient has previously seen other providers.

**History**

A detailed history of the pain report identifies the pain in terms of its origin, radiation, quality, severity, and time intensity attributes, as well as mode of onset, duration, time of occurrence,
and factors that aggravate and relieve it. Previous treatments for pain should be noted, including comments regarding usefulness in the reduction of pain. Medications currently being taken, as well as those used in the past, should be recorded, along with the patient’s perceptions as to the results achieved by each. The details of physical therapy, including types of modalities, exercise, and effectiveness of regimes, should be recorded. An inquiry as to the patient’s attempts at biofeedback, relaxation, and hypnosis is also helpful. Information regarding associated findings including sensory deficits, muscle weakness, and altered body function should be obtained. It should be determined whether compensation is involved and if the patient is working; if not, the employment history should be obtained. The physical examination and related diagnostic studies should be directed toward evaluating the painful site and related regions. This process is useful in acquiring objective data to substantiate the clinical history.

**Physical Examination**

Physical examination begins when the patient is first seen and continues through every contact made with the patient. This provides the opportunity for the physician to evaluate how the pain affects motion and activities. Physical examination always includes examination of related spine and musculoskeletal components, as well as a neurologic evaluation. Painful regions need to be compared with normal areas on the contralateral side of the patient for sensation, temperature, and sensitivity to palpation.

Functional evaluation measures the appropriateness of the patient’s functional capabilities for the level of impairment. Objective, quantitative measurements give a baseline with which to evaluate progress and long-term outcome. Specific tools are available to describe and classify functioning, such as the International Classification of Functioning, Disability, and Health (ICF) (113). The ICF characterizes the interaction between a health condition, personal attributes, and environmental influences. In this model, hierarchical levels categorize health-related states in increasing levels of detail to form a standardized description of the condition. Lists of categories relevant to specific diseases, known as ICF Core Sets, exist for different pain conditions including chronic widespread pain and low back pain (114–118).

**DIAGNOSTIC PROCEDURES**

**Laboratory Testing**

Laboratory findings in acute and chronic pain usually have no features distinct from those found with a primary disease. Drug screening tests of the blood and urine may provide valuable information as to the variety and type of pain medications being ingested. Serum drug level testing provides data to determine the bioavailability of medications being taken by the patient.

**Imaging**

Radiographic procedures are extensively used in the evaluation of pain. Spinal radiography has minimal value in the evaluation of most low back pain conditions because of the equal prevalence of abnormalities in symptomatic and asymptomatic populations, that is, low specificity and poor predictive value (119,120). Specific indications for routine radiological evaluation for patients with back pain have been developed (121) (Table 49-3). In the specific diagnosis of low back pain with sciatica due to herniated nucleus pulposus, plain radiography has been shown to have minimal value (122,123).

Radiography, computed tomography (CT), and magnetic resonance imaging (MRI) demonstrate anatomic or structural disorders, which account for a low percentage of functional abnormalities. However, CT, myelography, and MRI are well established in the diagnosis of disc herniation and provide significant false-positive results in subjects with no history of back pain. Asymptomatic individuals are diagnosed as having abnormalities (24% to 50%) on myelography (124), discography (125), CT (126), and MRI (127). The clinical usage of diagnostic tests requires a careful correlation between clinical signs, symptoms, and test results. For additional information on imaging studies, the reader is referred to Chapters 6, 7, and 8.

**Psychological Evaluation**

The assessment of psychological issues in the overall evaluation and treatment of patients with chronic pain is an important component of any pain management program (128). Psychological evaluation often involves the use of questionnaires, inventories, and the clinical interview. Psychophysical methods of pain assessment (129) (tourniquet test, cold pressor...
test, and visual analog scales) often augment the psychological evaluation.

The McGill Pain Questionnaire (130) is an often-used instrument designed to measure three dimensions of the pain experience: sensory, affective, and evaluative. The Minnesota Multiphasic Personality Inventory (MMPI or MMPI-2) has been used in the United States perhaps more than any other psychological instrument in the assessment of personality factors contributing to the experience of chronic pain (88). The typical MMPI profile for a patient with chronic pain denotes increased levels of hypochondriasis, depression, and hysteria. Numerous MMPI typologies based on clustering algorithms appear to have clinical and demographic correlates for the patient with chronic pain. However, these clustering algorithms have not been shown to be useful in determining successful outcomes of multidisciplinary pain management programs. Numerous MMPI typologies based on clustering algorithms have been proposed for the patient with chronic pain (131). The clinical interview helps to identify the affective, motivational, cognitive, and personality components of the patient with chronic pain. The emphasis of this evaluation is on the patient’s behavioral response to pain, adjustments to impairment/disability, primary/secondary gain, and motivation. The psychological evaluation will help clarify affective, cognitive, and motivational components of the chronic pain. In conjunction with pentothal (132–134) and amytal challenges (134–136), the psychological evaluation is helpful in differentiating patients with significant somatic/peripheral pain from mood disorders, somatoform disorders, factitious disorders, and conscience exaggerators (137).

**Electrodiagnosis**

Electrodiagnosis is an objective neurophysiologic extension of the physical examination. It typically includes the determination of nerve conduction velocities and needle electromyographic (EMG) studies of individual muscles. In addition, somatosensory evoked potentials have expanded the armamentarium of the electromyographer in evaluating the peripheral and central nervous system (138). For additional information on electrodiagnostic evaluation, the reader is referred to Chapters 4, 28, 29, and 30. EMG and evoked potential studies demonstrate the pathophysiologic changes that are associated with or due to structural abnormalities. The documentation of existing pathology is vitally important to the comprehensive management of complex painful conditions. Careful clinical correlation is essential when interpreting all tests related to chronic pain (139).

**Anesthetic Procedures**

The response of the body and mind to pain results in sympathetic nervous system and psychogenic responses. Although clinical examination and appropriate investigations can help in delineating the contribution of the various mechanisms in an individual patient, many patients do not exhibit clear-cut mechanisms of pain. In these patients, differential diagnostic blocks may be valuable. Clinicians using differential blocks are frequently impressed by their usefulness in pointing to mechanisms not suspected when considering the results of previous clinical examinations and laboratory work-up (140–142). For further information on anesthetic procedures, the reader is referred to Chapter 67.

**CLINICAL MANAGEMENT**

The primary goals of treating a patient with pain are alleviating the pain and enhancing the patient’s quality of life and functional capabilities. The management of acute pain is based on pharmacologic, psychological, medical, and surgical innovations or advancements within the past century. And yet, the management of chronic pain has been recognized as a major health care problem only in the past 25 years. It remains unclear why some people become patients with chronic pain and others resolve their acute pain without significant difficulty.

**Multidisciplinary Approach**

The chronic pain problem is clearly multifaceted. No single physician has the resources to care comprehensively for the complex psychological, social, legal, medical, and physical problems involved in chronic pain. Therefore, the multidisciplinary team approach is necessary. Using a multidisciplinary approach does not mean the patient is referred from one specialist to another, because this tends to result in conflicting and overlapping treatment and a loss of hope of treatment in the patient. Ideally, the team should work together to provide a unified explanation of the illness and a comprehensive treatment program. The multidisciplinary pain service has the advantage of offering a variety of coherent treatment approaches to the patient. This type of program recognizes that a multifaceted problem requires a multifaceted approach, as well as continuity of care in which the patient is an active participant (143).

The core group for the multidisciplinary treatment service should include a physiatrist, a pain medicine physician, a neurologist, a physiotherapist or occupational therapist, and a psychiatrist or clinical psychologist. This group may vary considerably according to local needs, resources, and available expertise. However, the team must have the knowledge to manage the psychological and social problems with optimal medical and anesthetic treatments. They must also have a thorough understanding of physical treatments and the rehabilitation process.

The multidisciplinary pain approach begins with a complete clinical evaluation. Comprehensive medical and psychosocial evaluations with particular emphasis on functional capabilities and behavioral responses to pain are essential. The somatic, affective, cognitive, and emotional components of the chronic pain experience are explored. All previous medical records are needed to avoid repeating appropriately performed studies and unsuccessful treatment approaches. This comprehensive clinical evaluation also includes functional capabilities to determine impairment level. The psychosocial evaluation focuses on the behavioral response to pain, adjustments to the physical impairment, and degree of motivation (144). The MMPI-2 and other written tests are often used for generalized screening.
The multidisciplinary team functions at several levels within the treatment process. It attempts to identify and resolve documentable organic problems when present and to improve the patient's ability to cope with the pain through medication, psychological intervention, and patient education. In addition, considerable effort is devoted to improving the patient's functional outcome as measured by increased activity time, improved activities of daily living, increased distance walked, and increased tolerance for specific homemaking or vocational activities (145–147). To accomplish these objectives, the multidisciplinary team must use many skills. In many cases, the patient with chronic pain is so entrenched in pain behavior that a behavior modification approach is essential. These patients are often characterized by low levels of activities of daily living, high demand for medication accompanied by physical and psychological dependency, high verbalization of pain, and the inability to work.

Pain Treatment Centers
The organization and operation of the multidisciplinary pain clinic has been discussed by multiple authors (148,149). Many behavior modification programs use the Fordyce model (69,88). This approach uses the general principles of interruption of the pain behavior reinforcement cycle, reward of healthy behavior, appropriate goals that the patient must achieve, measurement of improvement by functional assessment, as well as pain level, and psychosocial adjustment. Particular emphasis is placed on detoxification and medication reduction, pain reduction, increased activity, and modification of pain behavior.

If the patient is on pain medications and determined to be physically or psychologically dependent, he or she must be detoxified. This is routinely accomplished by establishing the equivalent dosage of each medication type (e.g., opioids, benzodiazepines, barbiturates, alcohol). Opioid medications are replaced with methadone, and long-acting barbiturates are replaced with phenobarbital or pentobarbital. Medication equivalents are placed in orange juice or in capsulated form and decreased at a rate of 5% to 10% per day. The medication is then given on an around-the-clock basis at fixed intervals. Gradual reduction of the pain ingredients occurs without significant side effects of withdrawal. The patient is not aware of the timing of the decrease but has been informed of the concept before starting the program. Nonsteroidal anti-inflammatory drugs (NSAIDs) and tricyclic or selective serotonin reuptake inhibitor (SSRI) antidepressants are routinely integrated as long-term medications. The pain management program is designed to reduce rather than eliminate pain, while increasing the patient's functional capabilities.

The patient with chronic pain usually exhibits a decreased activity level, which results in a disuse syndrome. The exercise programs are based on the initial specific and general exercise that the individual can perform. The exercise regimen is progressive, with the goals rising along with the patient's ability. Rewards for accomplishing tasks are a mainstay of this program with no reinforcement given for pain behavior. The achievable goals provide success and confidence and allow for frequent reinforcement when they are met. Cooperation by all staff members is essential; they must consistently ignore reports of pain and encourage improved function. Psychological intervention is used as indicated. The chronic pain behavior modification programs report short-term success rates in medication reduction, increased activity, and more productive behavior patterns (150,151). Statistics suggest improvement of 60% to 80% in patients with chronic pain without major psychosocial components, 30% to 50% in patients with significant psychosocial components, and approximately 20% in patients with major psychiatric components or secondary gains (148,152–156).

Multidisciplinary chronic pain treatment is a focused, unified approach to the chronic pain syndrome. In the United States, pain treatment centers differ widely in organization and emphasis. They are generally multidisciplinary centers that use some combination of anesthesiologists, clinical psychologists, dentists, neurologists, orthopedists, pharmacists, physiatrists, and psychiatrists. The goals of these centers are to diminish, if not eliminate, chronic pain; increase the patient's functional capabilities to allow for a more active life; and decrease the patient's dependence on drugs for pain control (Table 49-4).

Physical Modalities
Physical modalities are valuable adjuncts in the successful management of acute and chronic pain. Therapeutic heat and cryotherapy are time-honored interventions in the treatment of painful musculoskeletal conditions. Clinical evaluations of transcutaneous electrical nerve stimulation (TENS), acupuncture, and cold laser have questioned the efficacy of these methods in alleviating discomfort (157–161). For additional information on physical agents, the reader is referred to Chapter 63.

Pain arising from the musculoskeletal system is often caused by muscle spasm (127). Heat and cold applications are primarily directed at reducing spasmodic muscle shortening. The shortened muscle may be a result of direct muscular trauma or underlying primary neurologic or skeletal disease. Investigators have studied the muscle spindle and its firing rate in relation to thermal changes (162,163). Direct and indirect effects on the muscle spindle are detected from both heat and cold applications (164). The return of the muscle to its normal resting length is also believed to promote the reduction and resolution of pain (165), but precisely how muscle spasm is relieved is not completely understood.

Physical modalities should be used with caution and to a limited extent. Passive treatment programs such as hot packs, massage, and ultrasound may be appropriate for a short period of time; however, an active treatment program should be implemented as early as possible. The patient should be transitioned to a home program involving exercise, stretching, and self-applied modalities as early as possible. Additional outcome studies have demonstrated the effectiveness of therapeutic exercise for chronic back pain patients. Unfortunately, there remains limited consensus regarding the most effective exercise programs for patients with low back pain (166–169).
#TABLE 49.4 Step 1: Nonopioid Analgesics for Mild to Moderate Pain

<table>
<thead>
<tr>
<th>Drug</th>
<th>Plasma Half-life (h)</th>
<th>Peak Effect (h)</th>
<th>Duration of Analgesia (h)</th>
<th>Usual Dose (mg)</th>
<th>Maximum Recommended Dose (mg/d)</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aspirin</td>
<td>4–16</td>
<td>1–2</td>
<td>4–6</td>
<td>650 q 4–6 h</td>
<td>6,000</td>
<td>Standard of comparison for non-opioids; irreversible effect on platelet aggregation</td>
</tr>
<tr>
<td>Acetaminophen</td>
<td>1–4</td>
<td>3</td>
<td>1–4</td>
<td>650 q 4–6 h</td>
<td>4,000</td>
<td>No effect on gastric mucosa, platelet aggregation, or anti-inflammatory response. Hepatotoxic &gt;4 g/d</td>
</tr>
<tr>
<td>Diclofenac</td>
<td>6–8</td>
<td>2–4</td>
<td>10–12</td>
<td>50–75 q 12 h</td>
<td>150</td>
<td>Decreased GI toxicity</td>
</tr>
<tr>
<td>Diflunisal</td>
<td>8–12</td>
<td>2–3</td>
<td>8–12</td>
<td>500 q 12 h</td>
<td>1,500</td>
<td>No effect on platelet aggregation with &lt;1 g/d, decreased GI toxicity</td>
</tr>
<tr>
<td>Indomethacin</td>
<td>4–5</td>
<td>1–2</td>
<td>4</td>
<td>25–50 q 8 h</td>
<td>200</td>
<td>Not routinely used because of greater incidence of GI toxicity and CNS side effects</td>
</tr>
<tr>
<td>Ibuprofen</td>
<td>2–4</td>
<td>1–2</td>
<td>4–6</td>
<td>400–800 q 8 h</td>
<td>2,400</td>
<td>Rapid onset of action. Side effects include nephrotoxicity, tinnitus, CNS, and cardiovascular problems</td>
</tr>
<tr>
<td>Flurbiprofen</td>
<td>5–6</td>
<td>2–4</td>
<td>4–6</td>
<td>50 q 4–6 h</td>
<td>300</td>
<td>No effect on platelet aggregation</td>
</tr>
<tr>
<td>Naproxen</td>
<td>12–17</td>
<td>2–3</td>
<td>8–12</td>
<td>250–500 q 12 h</td>
<td>1,500</td>
<td>Increased GI toxicity &lt;20%</td>
</tr>
<tr>
<td>Prioxicam</td>
<td>45–50</td>
<td>3–5</td>
<td>12</td>
<td>20 q 24 h</td>
<td>40</td>
<td>Not recommended with liver or kidney dysfunction. Higher incidence of side effects at a dosage of 40 mg qd over 3 wk</td>
</tr>
<tr>
<td>Salsalate</td>
<td>4–16</td>
<td>1–2</td>
<td>4–12</td>
<td>1,000 q 8 h</td>
<td>4,000</td>
<td>No effect on platelet aggregation. Low GI toxicity</td>
</tr>
<tr>
<td>Sulindac</td>
<td>14–16</td>
<td>3–4</td>
<td>8–12</td>
<td>150 q 12 h</td>
<td>400</td>
<td>Decreased renal toxicity</td>
</tr>
<tr>
<td>Nabumetone</td>
<td>20–30</td>
<td>1–2</td>
<td>8–12</td>
<td>500 q 12 h</td>
<td>2,000</td>
<td>Near placebo level, GI side effects (4%)</td>
</tr>
<tr>
<td>Choline magnesium trisalicylate</td>
<td>4–6</td>
<td>3–5</td>
<td>4–12</td>
<td>1,000–1,500 q 12 h</td>
<td>4,000</td>
<td>No effect on platelet aggregation. Low GI toxicity</td>
</tr>
<tr>
<td><strong>Cox II inhibitors</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Celebrex (celecoxib)</td>
<td>11–13</td>
<td>2–9</td>
<td>8–12</td>
<td>200 q 12 h</td>
<td>800</td>
<td>Low GI toxicity (7%)</td>
</tr>
<tr>
<td><strong>Atypical analgesic</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Tramadol</strong></td>
<td>4.5–7.5</td>
<td>1–4</td>
<td>6–8</td>
<td>50 q 6 h</td>
<td>800</td>
<td>Atypical central acting analgesic. Risk of seizure, low GI toxicity</td>
</tr>
</tbody>
</table>

CNS, central nervous system; GI, gastrointestinal; NSAIDs, nonsteroidal anti-inflammatory drugs; q 12 h, every 12 h (etc.).

The authors will not assume liability for this table. No medication should be given until the complete prescribing recommendations, drug use indications, and potential side effects that are listed in the package insert with the product or contained in a drug reference manual are reviewed and understood thoroughly by the physician and patient. Drug dosages may need to be modified for the elderly.

*Dosage should be adjusted in elderly patients, patients on multiple medications, and patients with renal insufficiency or hepatic failure. Doses may be increased at weekly intervals if pain relief is inadequate and dosage is tolerated. Doses and intervals titrated to effect.

*The patient should be evaluated routinely for hepatic toxicity. Patients receiving NSAIDs also should be evaluated for renal toxicity and fecal blood loss that is due to gastrointestinal irritation. It is recommended that patients who develop visual symptoms during treatment undergo ophthalmic evaluation. Gastrointestinal disturbance may be reduced if taken with milk, on a full stomach, or with antacids.

*Combined action as an opioid agonist and monoaminergic drug with potential abuse potential. Nausea and vomiting side effects at same rate as opioid medications. Classified by World Health Organization as weak opioid, however classified in United States as nonopioid medication. Serious potential consequences of overdose are central nervous system depression, respiratory depression and death. 100 mg PO tramadol approximately equivalent 30 mg PO morphine. GI disturbance reduced with misoprostal 200 mg q 6 h. Misoprostal or a proton pump inhibitor is recommended in patients who are at increased risk for upper gastrointestinal adverse events. Max dose 1,600 mg/d. Do not give to pregnant women.
Cryotherapy

Therapeutic cold has four distinct applications in medicine: stop or slow bleeding, induce hyperthermia, decrease spasticity, and relieve pain (170,171). It may be applied as a solid, liquid, or gas. Ice is a common, solid form of direct cold therapy and is usually rubbed in a circular fashion over the localized painful area. Immersing a body part in water combined with ice chips may also provide pain relief. Cold applied directly to an injured area in acute musculoskeletal trauma serves to reduce hemorrhage and vasodilation, blunts the local inflammatory response, decreases edema production, and reduces pain perception. The so-called PRICE (protection, rest, ice, compression, and elevation) method is commonly prescribed for acute sports-related injuries (172).

In addition to acute musculoskeletal injuries, cryotherapy has been shown to benefit chronic painful conditions. Pain may be alleviated by direct or indirect mechanisms (173). Direct cold application serves to decrease temperature in the affected area. Reduced pain sensation is presumed to result through an indirect effect on the nerve fibers and sensory end organs. Additionally, the lowered temperature reduces muscle spindle firing rate and decreases the painful muscle tone (163,173).

The direct application of ice massage has been shown to be therapeutically efficacious in several clinical trials. Grant demonstrated the beneficial effects of direct ice application on a large population of young individuals sustaining acute and chronic musculoskeletal trauma (174). Pegg and colleagues demonstrated that chronic, inflammatory joint disease improved with cold application with regard to pain, stiffness, and range of motion (175). Patients with low back pain also responded well to controlled clinical trials using cryotherapy (176,177).

The application of cold as a vapocoolant spray was popularized by Travell in treating myofascial pain syndromes (178). A counterirritant effect is presumed to provide the mechanism of muscle spasm relief and pain alleviation. The combination of vapocoolant spray, stretching, and trigger point injection has been reported to provide significant pain relief in myofascial pain syndromes (179,180). The mechanism of this pain relief remains unclear.

Adverse effects also have been reported with the use of cold (173). The major adverse reaction is hypersensitivity. Patients with Raynaud’s and peripheral vascular disease should not have their limbs exposed to cold temperatures because this produces vasoconstriction. Additionally, patients who display a marked cold pressor response are poor candidates for cryotherapy. Care should be taken not to apply ice for a sustained period of time over the medial elbow or the head of the fibula as this may result in cold-induced injury to the ulnar and peroneal nerves, respectively. When applying an ice bag to a large joint or limb, a patient should be advised to place a towel between the ice and their skin and to limit their exposure to 20 minutes.

Heat Therapy

Heat application is a common form of pain treatment. It is generally accepted that therapeutic heat is best tolerated in the subacute and chronic phases of a disease process. The physiologic responses produced by heat are increased collagen extensibility, increased blood flow and metabolic rate, and inflammation resolution. Decreased joint stiffness, muscle spasm, and pain are also beneficial effects of heat.

Therapeutic heat is believed to have direct and indirect effects on the muscle spindle. Local elevated temperatures have been shown to directly decrease the spindle sensitivity (181), and superficial heating of the skin has been demonstrated to indirectly reduce spindle excitability (164). This mechanism is similar to that proposed for decreased muscle spindle activity with the application of cold. It is also believed that the pain threshold may be raised by the direct and indirect actions of heat.

Pain associated with numerous conditions has been successfully treated with therapeutic heat application (182). Musculoskeletal contractures respond well to deep heat used in association with prolonged stretch (183). Joint stiffness associated with chronic inflammatory diseases, particularly those affecting the limbs, responds to superficial heating with decreased pain and increased range of motion and function. Subacute and chronic bursitis, tenosynovitis, and epicondylitis also may respond to heat with decreased pain and symptom resolution (177).

Contraindications to superficial heat include sensory impairments, circulatory insufficiency, malignancy, and infection. Edema is considered a relative contraindication with the benefits of the treatment weighed against the potential disadvantages. Contraindications for the use of ultrasound include use over the eye, malignancy, impaired circulation, pregnancy, impaired sensation, and infection. Contraindications for other deep-heating modalities include pregnancy, ischemic tissue, pain or sensory deficit, areas of high fluid volume, and any metal implants.

Transcutaneous Electrical Nerve Stimulation Therapy

The use of electrical currents dates back to the Greeks, who applied torpedo fish to individuals suffering from pain (184). Therapeutic electricity eventually fell into disrepute until Melzack and Wall proposed the gate control theory of pain (36). This gate control model provided the theoretical basis for the use of electrical current in pain control. They found that the preferential activation of large afferent fibers using TENS would inhibit the transmission of painful impulses. The exact physiologic basis by which TENS produces pain control is unknown.

TENS also has been used extensively to manage chronic pain. The results have been less promising and more variable than those in acute pain trials (185,186). Rigorous controlled trials suggest that TENS is no better than placebo in the management of chronic low back pain (3,158,187). Patients with CRPS, phantom limb pain, and peripheral nerve injury have demonstrated pain control with TENS (188–190). The most efficacious use of TENS appears to be in the control of selected acute pain conditions, such as postoperative incisional pain.

There are relatively few side effects related to TENS. The major problem is one of skin irritation related to the conducting paste or tape used to secure the electrodes. The use of TENS in patients with demand cardiac pacemakers is inadvisable (190).
Acupuncture

Acupuncture (originating from the Latin *acus* or “sharp point” and *punctura*, “to puncture”) is an ancient Chinese therapy practiced for more than 2,500 years to cure disease or relieve pain. Thin, solid metal needles are inserted into specific body sites and slowly twisted manually or stimulated electrically. Various sensations may be produced, ranging from a dull ache or warmth to that of a pinprick. The Chinese believe that acupuncture achieves its beneficial results by restoring the balance between *yin* (blood) and *yang* (spirit), which flow in 14 channels or meridians containing 361 acupuncture sites.

Researchers have considered acupuncture to be a form of neuromodulation. Two theories have been proposed for its use in pain control. First, acupuncture may stimulate large sensory afferent fibers and suppress pain perception as explained by the gate control theory of pain. Second, the needle insertion may act as a noxious stimulus and induce endogenous production of opioid-like substances to effect pain control (190).

It has been demonstrated that there is a significant overlap between traditional acupuncture sites, myofascial trigger points, and muscular motor points (191). The sensation induced by the application of an acupuncture needle is very similar to the dull ache often experienced by the patient when a trigger point is injected. The insertion of a needle, regardless of the substance injected, appears to produce the beneficial pain relief and is termed the needle effect. The injection of trigger points may share not only similar areas of needle insertion, but also associated mechanisms of pain control.

Acupuncture has been used in a wide variety of painful conditions (192,193). The insertion of a needle is considered an invasive procedure, and many states require a physician to perform or supervise the treatment. Uniform agreement does not yet exist as to the preferred time necessary for an adequate trial of acupuncture. There is limited evidence to suggest that acupuncture is more effective than placebo or sham acupuncture for chronic low back pain in the short term only. There is no evidence that acupuncture is more effective than other treatments such as NSAIDs. While acupuncture may be a useful adjunctive treatment in pain management, more high-quality studies are needed to confirm its efficacy and guide its use (159,160,194,195).

Reported side effects of acupuncture include localized hyperemia occurring frequently following needle insertion. Infrequently, syncopal episodes, hematoma formation, and pneumothorax also have been reported. Caution is suggested in the use of electrical acupuncture with patients who have cardiac pacemakers (196).

Therapeutic Exercise

During acute injuries to the musculoskeletal system, a muscle may shorten as a protective reaction to pain. Treatment typically consists of immobilization combined with compression and cryotherapy. As the pain subsides, mobility is restored gradually. If normal range of motion is not achieved, the muscle may become chronically shortened and result in additional pain.

Prolonged muscle shortening will add to the painful condition by producing contracted soft-tissue structures. In the chronic phase of pain, the optimal treatment methodology combines graded stretching movements, strengthening exercises, heat or cold, and massage. The patient is also educated regarding proper body mechanics and the need to continue the prescribed therapeutic exercise regimen outside of formal therapy sessions.

Therapeutic exercise, prescribed to correct a specific abnormal condition, is often used to treat chronic painful conditions. The primary goal is to aid the patient in achieving pain control. This may be accomplished through the restoration of normal muscle tone, length, strength, and optimal joint range of motion (197). Finally, the patient is urged to continue a home program after formal therapy sessions have ceased (121).

Therapeutic exercise consists of passive movements, active-assistive exercises, active exercises, stretching, and relaxing exercises. Each may be used alone or in combination to achieve the desired effect (153,198–202). In the case of low back pain, there has been disagreement as to which exercise program is most effective in the treatment of chronic, low back pain. One school of thought is that therapeutic stretching exercises, similar to the YMCA program developed by Kraus, are effective in the management of chronic low back pain (3,121,167,169). A recent meta-analysis of randomized, controlled trials in patients with low back pain showed a positive effect in function both immediately and at follow up. While this study highlighted the variety of programs that are incorporated under the umbrella of “exercise,” the majority (12 out of 16) incorporated an element of strengthening (202). There is an emerging school of thought that the maintenance of lumbar spinal stability over time can limit the pain associated with degenerative disc disease and spondylolisthesis.

This stability is thought to be secondary to three components: the bones and ligaments of the spine, the muscles, and the nerves that coordinate muscle activation (203). The muscles are the major stabilizers, specifically the multifidi, the transversus abdominis, the pelvic floor muscles, and the diaphragm (204). Without the assistance of the muscles and their nerves, a cadaver spine with bones and ligaments intact will buckle and collapse under a load of 20 lb (205).

The multifidi are rich in muscle spindles; therefore, they act to both stabilize individual segments of the spine and to provide proprioceptive feedback of spinal movement. There is some evidence that these muscles may be atrophied in patients with low back pain (206).

The other muscles, including the transversus abdominis, form a muscular cylinder that increases intra-abdominal pressure also effectively stabilizing the spine. The transversus abdominis is one of the first muscles to be activated to stabilize the spine; however, studies have suggested that its activation may be delayed in patients with back pain (207,208).

Several small studies have shown a promising effect of lumbar stabilization programs targeted at strengthening the above, deep muscles in treating a number of painful and potentially destabilizing spinal conditions. Of note, O’Sullivan and colleagues reported a significant improvement in pain and function of a group of patients with spondylolisthesis treated with lumbar stabilization compared
to patients treated with general exercise (209). In a study by Cairns (210), 95 patients with recurrent back pain were randomly assigned to treatment with “conventional physical therapy” involving general active exercise or a program including specific strengthening of the spinal stabilizers. Both groups improved, but there was no statistically significant difference in function between the two groups on follow up.

Passive forms of therapeutic exercise (mobilization, manipulation, and massage) have not been shown to be superior to other routine treatments (general practitioner care, analgesics, home exercises, or back school) in the management of neck and back pain (211–215). For additional information on therapeutic exercise, the reader is referred to Chapter 61.

**Behavioral Treatment Modalities**

Among the treatment goals of pain management are the decrease in ill behavior (reduced drug use and visits to physicians) and the increase in well behavior (increased physical activities, mobility, and return to gainful employment). This may be accomplished by blocking noxious sensory input, decreasing tension and depression, rearranging reinforcement contingencies, or assisting in the learning of new behaviors (216). Biofeedback, cognitive behavior modification, operant approaches, hypnosis, operant pain hypnosis, and relaxation training can assist in the treatment of chronic pain behavior (217–222,222a).

Biofeedback has found some use in the treatment of chronic pain (223–225). Typically, biofeedback teaches muscle relaxation (through surface EMG) or temperature control. The instrumentation is reported to be somewhat useful, although clinical experience suggests that relaxation without instrumentation is of equal value. Through biofeedback training, the patient learns self-regulation of pain.

Cognitive behavior modification (220) helps the patient learn self-coping statements and problem-solving cognitions (226) in order to alter the cognitive structures (schemata, beliefs) and cognitive process (automatic thoughts, images, and internal dialogue) associated with the pain experience. Cognitive strategies of imaginative inattention, imaginative transformation of pain, focused attention, and somatization in a dissociation manner have been found helpful.

The operant approach involves the identification of behaviors to be produced, increased, maintained, or eliminated. Reinforcement is then regulated to achieve the desired outcome. Activity and walking programs are followed to the prescribed level, not to discomfort. All medication is prescribed by schedule. Family and friends are instructed to avoid reinforcing all pain behavior.

Relaxation methods (227,228) to reduce tension may include deep muscle relaxation, deep diaphragmatic breathing, meditation, yoga, and autogenic training. Patients also may be taught self-hypnosis (223). Hypnosis (229–233) has the advantage of providing pain relief without unpleasant side effects, no reduction in normal function, and no development of tolerance. Hypnotic strategies can suggest analgesia or anesthesia, substitute another feeling for pain, move the pain perception to a smaller or less vulnerable area, alter the meaning of pain, increase tolerance to pain, or, in some individuals, dissociate the perception of body from the patient’s awareness.

**Pharmacologic Intervention**

Two groups of patients are encountered in the clinical management of chronic pain. The first group includes individuals with recurrent, overwhelming acute exacerbations of pain due to flares of chronic medical illness, such as rheumatoid arthritis, cancer, and burn injuries. Primary therapy is usually directed at the underlying cause of pain. When the therapy is successful, the pain treatment is successful, and the patient can return to their baseline level of pain (if any). Treatment commonly includes NSAIDs. Opioid analgesics are used on time-limited regimes to minimize the development of psychological dependence. The second group of patients consists of individuals who have chronic pain without apparent organic etiology. Pharmacologic intervention is the most common means of treatment for all forms of chronic pain. Pharmacological substances may be divided into three categories: nonopioids, opioids, and adjuvant drugs.

**Nonopioid Analgesics**

Nonopioid medications are primarily NSAIDs that include aspirin, acetaminophen, and simple NSAIDs (see Table 49-4). These drugs are peripherally active analgesics that do not inhibit nociception or alter the perception of the pain input. They are best considered remitent agents that alter the pathologic processes that generate pain. Aspirin and other NSAIDs reduce pain by interfering with prostaglandin sensitization of nociceptors and inhibiting the synthesis of prostaglandins. Additional NSAID actions include inhibition of tissue reaction to bradykinin, suppressed release of histamine, and decreased vascular permeability. This improves the environment of the nociceptor, increasing pain control by decreasing sensitivity. With the exception of acetaminophen, the NSAIDs also possess anti-inflammatory effects that reduce local heat, swelling, and stiffness. These drugs are used to treat patients with acute and chronic pain of low to moderate severity.

NSAIDs are often chosen over opioids because they have fewer side effects, including no constipation, very little sedation, no psychological or physical dependence, and no development of tolerance. The mainstays of nonopioid analgesic therapy are aspirin and acetaminophen. Acetaminophen is an excellent alternative to aspirin in patients who are unable to tolerate other NSAID medication. All NSAIDs have ceiling effects, but the ceilings for some of these drugs are higher than that of aspirin (234).

Although aspirin, acetaminophen, and other NSAID compounds are available over the counter, they all have potential side effects. The most common complication, involving the gastrointestinal tract, is seen in 5% to 10% of patients. These drugs produce in varying degrees gastrointestinal, hematologic, renal, and hepatic toxicities. The side effects from prolonged use of all NSAIDs are similar to those occurring with aspirin, but there are significant differences in the potential for side effects. Aspirin, in the nonsalicylated salicylate form, remains the preferred drug for the patient with chronic pain.
The newer NSAIDs have not been proven to offer any major advantage over aspirin, although their cost is much greater (121,167,235–238). Notable exceptions to this are patients unable to tolerate acetylsalicylic acid or who demonstrate difficulty with compliance. For additional information on NSAID medications, the reader is referred to Chapter 66.

**Opioid Analgesics**

Opioid medication is useful and appropriate in the treatment of acute, recurrent, or cancer pain. Acute and recurrent pain are usually best managed by diagnosis and treatment of the underlying cause of the pain. The opioid medication is used as an adjunct to provide relief during a period of temporary, excruciating pain. Opioid analogues are also preferred for relief of intractable pain that is due to cancer. The greatest obstacle to treating postoperative pain, pain related to cancer, acute pain, and recurrent pain that is due to chronic disease is the excessive concern of the physician regarding addiction. Repeated studies report that a high percentage of clinicians tend to overestimate their patient’s potential for addiction and, consequently, undermedicate them (83–85,239). Ironically, undermedication may increase the potential for addiction as a result of the operant conditioning, anxiety, and dependent behavior created by inadequate pain relief. Psychological dependence has not been a major problem in patients with acute pain or cancer pain who receive appropriately dosed opioid analgesics for moderate to severe pain (240,241).

Opioid medications should be avoided in the treatment of patients with chronic pain without active organic etiology. There is limited justification for the use of opioid drugs in a patient without an organic etiology for chronic pain (242). Long-term use of opioid drugs with these patients often produces behavioral complications that are more difficult to manage than the initial pain problem. The hazards of tolerance, physical dependence, and psychological dependence present major problems in the long-term management of this patient group. Deficits of cognition and motor function, as well as the masking of psychological disorders, are common. Immune suppression, hormonal axis suppression, and hyperalgesia are also possible (44). The indiscriminate use of opioid medications in an attempt to control chronic pain only enhances chronic pain behavior. For patients with pain secondary to organic etiology, tolerance and dependence are expected with long-term use of opioids and do not equal addiction (Table 49-5).

<table>
<thead>
<tr>
<th>Drug</th>
<th>Aspirin Oral</th>
<th>Morphine IM</th>
<th>Morphine Oral</th>
<th>Action Time</th>
<th>Usual Dose Limits</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Codeine</td>
<td>32</td>
<td>120</td>
<td>200</td>
<td>2–3</td>
<td>32–65 q 3–4 h</td>
<td>Weak, short acting; as dose increases, nausea, vomiting, and constipation occur more frequently.</td>
</tr>
<tr>
<td>Hydrocodone</td>
<td>2.5</td>
<td>15</td>
<td>30</td>
<td>2–4</td>
<td>5–10 q 3–4 h</td>
<td>Only available combined with NSAIDs.</td>
</tr>
<tr>
<td>Meperidine</td>
<td>50</td>
<td>90</td>
<td>300</td>
<td>2–4</td>
<td>50–100 q 3–4 h</td>
<td>Not recommended for the treatment of cancer pain.</td>
</tr>
<tr>
<td>Pentazocine</td>
<td>30</td>
<td>60</td>
<td>180</td>
<td>2–3</td>
<td>50–100 q 4 h</td>
<td>Not recommended for the treatment of cancer pain.</td>
</tr>
<tr>
<td>Propoxyphene</td>
<td>100–200</td>
<td>300</td>
<td>300</td>
<td>2–3</td>
<td>100–200 q 4 h</td>
<td>Not recommended for the treatment of cancer pain.</td>
</tr>
</tbody>
</table>

IM, intramuscularly; NSAIDs, nonsteroidal anti-inflammatory drugs; q 4 h, every 4 h (etc.).

The authors will not assume liability for this table. No medication should be given until the complete prescribing recommendations, drug use indications, and potential side effects are listed in the package insert with the product or contained in a drug reference manual are reviewed and understood thoroughly by the physician and patient. Drug dosages may need to be modified for the elderly.

Varies with route of administration.

Dose providing analgesic equivalent to 650 mg aspirin orally.

Dose providing analgesic equivalent to a single dose of 10 mg morphine IM, intravenous, or subcutaneous.

Relative potency of IM morphine to oral morphine is 1:6 with acute dosing and is 1:2–3 with chronic dosing.

Most common side effects of opioid drugs include constipation, nausea, and sedation. Less common side effects include urinary retention, bladder spasm, respiratory depression, and intermittent vomiting. Rare side effects include psychotic symptoms, pruritis, orthostatic hypotension, bronchoconstriction, and biliary colic. The dosage should be adjusted in elderly patients and patients with impaired ventilation, increased intracranial pressure, liver failure, or bronchial asthma.
It is often necessary to educate the patient and family regarding the appropriate use of opioids for pain that is due to organic etiology. For additional information on opioids, the reader is referred to Chapter 66.

Morphine is the opiate drug prototype and commonly prescribed by many clinicians (8,237). Pain relief is often obtained by titrating the dose to the patient's needs. At equianalgesic doses, there is no significant pharmacologic evidence to suggest the efficacy of one opioid over another (243,244), but there are significant differences in their action times, equianalgesic dose, and parenteral/oral ratio (Table 49-6). Inappropriate drug dosing often occurs because of a lack of knowledge or attention to equianalgesic doses, resulting in inadequate pain relief.

All clinically useful opioids produce similar side effects in equianalgesic doses. The undesirable side effects of opioids include respiratory depression, unwanted sedation, mental clouding, inability to concentrate, lethargy, impairment of mental and physical performance, constipation, nausea and vomiting, tolerance, physical dependence, and psychological dependence. These side effects compromise a patient's main goal of maintaining a normal lifestyle. When introducing opioid medications for the first time or changing a patient's chronic dosage, it is imperative to counsel them about precautions regarding increased sedation, driving impairment, risk of falls, and risk of respiratory depression over the next days to weeks until the patient becomes acclimated.

Oral administration of medication is preferred in the treatment of all pain. A time-contingent, round-the-clock schedule for pain medications is superior to an as-needed schedule. This form of administration minimizes alterations in plasma levels and provides optimal pain control. The schedule should be based on such variables as potency, duration of the analgesic effect, and efficacy of the analgesic medication. A regularly scheduled dosage optimizes the reduction of pain by minimizing the peaks and valleys of pain intensity. Generally, it is better to begin the initial dose of medication too high rather than too low.

<table>
<thead>
<tr>
<th>Drug</th>
<th>Equianalgesic Dosage (mg)a</th>
<th>Action Timeb (h)</th>
<th>Duration of Analgesic</th>
<th>Usual Dose Limits (mg)d</th>
<th>Commentsa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hydromorphone</td>
<td>3</td>
<td>2–3</td>
<td>1/2–1½</td>
<td>2–4</td>
<td>4–8 q 4–6 h</td>
</tr>
<tr>
<td>Levopranol</td>
<td>2</td>
<td>12–16</td>
<td>1–2</td>
<td>4–5</td>
<td>2–4 q 4–6 h</td>
</tr>
<tr>
<td>Methadoneh</td>
<td>10</td>
<td>15–57</td>
<td>1–2</td>
<td>6–8</td>
<td>15 q 8 h</td>
</tr>
<tr>
<td>Fentanyl transdermal patch</td>
<td>—</td>
<td>16–24</td>
<td>12–24</td>
<td>58–72</td>
<td>100 μg/h patch q 48 h</td>
</tr>
<tr>
<td>Methadoneh</td>
<td>10</td>
<td>30</td>
<td>2–4</td>
<td>½–1½</td>
<td>10 q 4 h</td>
</tr>
<tr>
<td>Sustained-released morphine</td>
<td>—</td>
<td>2–4</td>
<td>—</td>
<td>8–12</td>
<td>60 mg q 12 h</td>
</tr>
<tr>
<td>Dicyclomine</td>
<td>1</td>
<td>0.5</td>
<td>2–3</td>
<td>30–90</td>
<td>10 q 4 h</td>
</tr>
</tbody>
</table>

IM, intramuscularly; IV, intravenously; q 4 h, every 4 h (etc.); q 3 d, every 3 d.

The authors will not assume liability for this table. No medication should be given until the complete prescribing recommendations, drug use indications, and potential side effects that are listed in the package insert with the product or contained in a drug reference manual are reviewed and understood thoroughly by the physician and patient. Drug dosages may need to be modified for the elderly.

aDose providing analgesic equivalent to a single dose of 10 mg morphine IM, IV, or subcutaneous.

bVaries with route of administration.

cRelative potency of IM morphine to oral morphine is 1:6 with acute dosing and is 1:2–3 with chronic dosing.

dDosage varies considerably, titrate to control pain.

Step 3: Opioid Analgesics for Severe Pain

- Fast acting
- Increased sedation with repeated doses
- Avoid in patients with significant respiratory, hepatic, cardiac, or renal failure
- Usually not recommended for opioid naive patients
- Standard of comparison for opioids
- Rectal suppository

10 mg rectal equianalgesic to 10 mg morphine IM.
too low. Starting suboptimally and titrating upward results in the patient's experiencing pain and anxiety because of a lack of adequate analgesic. The as-needed, or ‘prn,’ schedule does not have a place in the control of chronic pain. Such a schedule results in operant conditioning, craving, a sense of dependence, and anxiety about the drug wearing off. The administration of opioids before the recurrence of pain sensation decreases the development of central nervous system sensitization. In chronic pain management, the drugs with longer duration of action are usually preferred. There is considerable patient-to-patient variation with respect to effective analgesic dosage; therefore, it is important to individualize the medication regime for each patient (245,246).

Titration of opioids is generally accomplished by increasing or decreasing the next dose by one fourth or one half of the previous dose. During titration, patients are often provided with medication for breakthrough pain. Rapidly acting opioids of approximately 10% of the individual's baseline opioid equianalgesic dose may be given every 2 to 3 hours as needed (247,248).

Adjuvant Analgesic Medications

The adjuvant analgesic drugs produce or potentiate analgesia by mechanisms not directly mediated through the opiate receptor system (Table 49-7). This group includes a wide variety of compounds with no proven specific analgesic properties: antidepressants, anticonvulsants, and antispasmodics. The use of these drugs is often based on anecdotal data, clinical surveys, or limited drug trials.

SSRIs are often used to manage the psychological factors common to chronic pain syndromes. These medications have not been shown to have direct analgesic properties, and their mechanism of action is unclear. SSRIs such as citalopram, fluoxetine, fluvoxamine, paroxetine, delafaxine, and sertraline have few serious side effects but have been noted to cause insomnia, memory impairment, and weight changes. SSRIs are used for multiple pain conditions (249–252). SNRIs, such as duloxetine and venlafaxine, are also now used in treatment of neuropathic pain and depression (253,254).

Tricyclic antidepressants, such as amitriptyline, doxepin, and imipramine, have been used in the treatment of chronic pain syndromes. One of the primary mechanisms of tricyclic compounds is to block the reuptake of the neurotransmitter serotonin in the central nervous system. This enhances pain inhibition by way of the dorsolateral pathway (32,255). In addition, amitriptyline is a potent sedative drug, which may be used as a sleeping medication in patients with chronic pain. The combination of antidepressant effect, enhanced cortical serotonergic mechanism, and improved sleep contributes to these medications being one of the most commonly used group of psychotrophic agents in pain management (256–258).

Gabapentin, pregabalin, and other anti-epileptic drugs are used in the management of pain syndromes affecting the central nervous system. Anticonvulsants such as pregabalin, gabapentin, lamotrigine, levetiracetam, oxcarbazepine, tiagabine, topiramate, and zonisamide have the advantage of a much-improved side-effect profile compared with carbamezipine, phenotoin, and valproic acid. Their applications include trigeminal neuralgia, postherpetic neuralgia, causalgia, phantom pain syndromes, painful diabetic neuropathy, and FM (259–262). Although the mechanism is unclear, they appear to have a stabilizing effect on excitatory cell membranes, which decreases afferent and deafferent second-order neuron activity (263–267).

Alternative anticonvulsants used for pain control are valproic acid and clonazepam. They increase the effectiveness of GABA-induced inhibition in the pre- and postsynaptic systems. These drugs appear to be most effective in the treatment of neuralgias and neuropathies. Antispasmodics such as baclofen are presumed to act by inhibiting gamma transaminases and their reuptake at gamma receptor sites. Valproic acid and clonazepam act in a similar fashion (268–270).

In addition to these medications, a number of other adjuvant medications have been used. These include mehtotrimethazine, chlorpromazine, and fluphenazine. Butyrophenones are anecdotally in the management of pain disorders, with haloperidol being the most often reported. Antihistamines, amphetamines, steroids, and cannabinoids are also reported anecdotally. Steroids such as prednisone and dexamethasone are thought to interfere with prostaglandin sensitization of nociceptors. Serotonin antagonists such as ergot alkaloids, the β-blocking agents such as propranolol, and the antihistamines such as hydroxyzine all function by antagonizing transmitters that directly activate nociceptors. These medications have been used extensively in the treatment of migraine and cluster headaches. Lithium and calcium-blocking agents have been proposed as drugs to interfere with the release of transmitters involved in the pain process (41).

Benzodiazepines and barbiturates are two groups that have little or no place as adjuvant drugs in chronic pain management. Long-term use of these medications may result in psychological and physical dependence as well as in interference with cognition and motor function. Historically, benzodiazepines, because of their claimed muscle relaxant properties, are often prescribed to patients with pain. However, their role as muscle relaxants is questionable in clinical studies (271–273). In addition to the adverse effects of dependency, it has been suggested that these medications adversely affect the serotonin system. These medications are depressants that, with long-term use, lower pain tolerance, increase hostility, and tend to induce clinical depression as well as psychological and physical dependence (274,275). Because of their sedative effects, depressants often act as potent reinforcers of pain in drug-seeking behaviors. Chronic use of these medications may result in physical and mental incapacitation, emotional instability, and the inability to deal with initial physiologic or psychological problems. Benzodiazepines deplete serotonin, alter sleep patterns, and increase pain perception. It is recommended that these two groups of medications should not be part of the long-term management of chronic pain. The only possible indication for these sedatives, or for anxiolytic agents, is for the short-term (<1 month) treatment of a self-limited crisis unrelated to the particular pain problem, or as an adjunct when detoxifying a patient from opioid medication.
### TABLE 49.7 Adjuvant Analgesic Medications

<table>
<thead>
<tr>
<th>Drug Class</th>
<th>Drug</th>
<th>Indication</th>
<th>Usual Dose&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Maximum Dose (mg/d)</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Antidepressants</td>
<td>Amitriptyline</td>
<td>For neuropathic deafferentation, and somatic pain complicated by insomnia or depression.</td>
<td>25–75 mg qhs or divided dose</td>
<td>300 mg/d</td>
<td>Tertiary amines—sedative, if morning somnolence is a problem, give dose earlier in the evening</td>
</tr>
<tr>
<td></td>
<td>Doxepin</td>
<td></td>
<td>25 mg tid-qid</td>
<td>300 mg/d</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Imipramine</td>
<td></td>
<td>50–200 mg qd</td>
<td>375 mg/d</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Venlafaxine</td>
<td></td>
<td>37.5 mg bid</td>
<td>60 mg/d</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Duloxetine</td>
<td></td>
<td>20 mg qam</td>
<td>80 mg/d</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Citalopram</td>
<td></td>
<td>20 mg qam</td>
<td>200 mg/d</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Venlafaxine</td>
<td></td>
<td>20 mg qam</td>
<td>60 mg bid</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Desipramine</td>
<td></td>
<td>50 mg qam</td>
<td>3,600 mg/d</td>
<td>Secondary amines—less sedative</td>
</tr>
<tr>
<td></td>
<td>Venlafaxine</td>
<td></td>
<td>300 mg qid</td>
<td>1,800 mg/d</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Desipramine</td>
<td></td>
<td>200 mg bid</td>
<td>200 mg/d</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Venlafaxine</td>
<td></td>
<td>50 mg bid</td>
<td>30 mg bid</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Venlafaxine</td>
<td></td>
<td>200 mg bid</td>
<td>20 mg qam</td>
<td></td>
</tr>
<tr>
<td>Anticonvulsants</td>
<td>Pregabalin</td>
<td>Intermittent lancinating and continuous neuropathic pain</td>
<td>50 mg tid</td>
<td>450 mg/d</td>
<td>Few side effects, few drug-drug interactions</td>
</tr>
<tr>
<td></td>
<td>Gabapentin</td>
<td></td>
<td>300 mg qid</td>
<td>3,600 mg/d</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Topiramate</td>
<td></td>
<td>50 mg bid</td>
<td>1,800 mg/d</td>
<td>Risk of leucopenia</td>
</tr>
<tr>
<td></td>
<td>Lamotrigine</td>
<td></td>
<td>100 mg tid</td>
<td>600 mg/d</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Carbamazepine</td>
<td></td>
<td>50 mg bid</td>
<td>400 mg/d</td>
<td></td>
</tr>
<tr>
<td>Cardiovascular</td>
<td>Lidocaine</td>
<td>Burning neuropathic pain</td>
<td>50–100 mg IV</td>
<td>300 mg/h</td>
<td>2–50 mg/min q 5 min</td>
</tr>
<tr>
<td></td>
<td>Mexitetine</td>
<td>150 mg tid</td>
<td>1,200 mg/d</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tocaide</td>
<td>400 mg tid</td>
<td>2,000 mg/d</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neuroleptics</td>
<td>Fluphenazine</td>
<td>Refractory pain</td>
<td>2.5 mg tid</td>
<td>40 mg/d</td>
<td>Used with pain complicated by nausea or delirium</td>
</tr>
<tr>
<td>Corticosteroids</td>
<td>Dexamethasone</td>
<td>Refractory bone and nerve pain</td>
<td>16 mg/d</td>
<td>96 mg/d</td>
<td>Used for malignant lesions, multiple side effects</td>
</tr>
<tr>
<td></td>
<td>Prednisone</td>
<td></td>
<td>10 mg/d</td>
<td>60 mg/d</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pemidronate</td>
<td>Metastatic bone pain</td>
<td>90 mg IV</td>
<td>90 mg q 4 wk</td>
<td>Inhibit bone resorption</td>
</tr>
<tr>
<td></td>
<td>Alendronate</td>
<td></td>
<td>10 mg q d</td>
<td>70 mg q wk</td>
<td></td>
</tr>
<tr>
<td>Biphosphates</td>
<td>Pamidronate</td>
<td>Bone pain</td>
<td>10 µCi</td>
<td></td>
<td>Possible thrombocytopenia</td>
</tr>
<tr>
<td></td>
<td>Alendronate</td>
<td></td>
<td>5 mg bid</td>
<td>40 mg/d</td>
<td>Reduce sedative effects of opioids</td>
</tr>
<tr>
<td></td>
<td>Methylenidate</td>
<td>Somatic and visceral pain</td>
<td>10–20 mg tid</td>
<td>60 mg/d</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Alendronate</td>
<td></td>
<td>37.5 mg qam</td>
<td>112 mg/d</td>
<td></td>
</tr>
<tr>
<td>Neurostimulants</td>
<td>Dexaemaphetamine</td>
<td></td>
<td>65 mg/d</td>
<td>200 mg/d</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Methylphenidate</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pemoline</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Caffeine</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tizanidine</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Clonidine</td>
<td>Peripheral neuropathy</td>
<td>q 12 h</td>
<td></td>
<td>Transdermal patch</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3–4 times/d</td>
<td></td>
<td>Causes burning sensation</td>
</tr>
<tr>
<td>Topical agents</td>
<td>Lidocaine</td>
<td>Neuropathic pain</td>
<td>25–50 mg</td>
<td>150 mg/d</td>
<td>Epidural injection (significant hypotension may occur)</td>
</tr>
<tr>
<td></td>
<td>Capsaicin</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>α&lt;sub&gt;2&lt;/sub&gt;-adrenergic agonist</td>
<td>Clonidine</td>
<td>Neuropathic pain</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NDMA antagonists</td>
<td>Ketamine</td>
<td>Neuropathic pain</td>
<td>12.5–25 mg IV</td>
<td>150 mg IV test</td>
<td>Possible hallucinations and nightmares</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>50 mg PO</td>
<td>250 mg/d</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Intrasinal</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>GABA agonist</td>
</tr>
<tr>
<td>Miscellaneous</td>
<td>Calcitonin</td>
<td>Bone pain</td>
<td>200 IU qd</td>
<td>80 mg/d</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Baclofen</td>
<td>Muscle spasm</td>
<td>5 mg tid</td>
<td>80 mg/d</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>8 mg tid</td>
<td>40 mg/d</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Diazepam</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tizanidine</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup> Dosage should be adjusted in elderly patients, patients on multiple medications, and patients with renal insufficiency or hepatic failure. Doses may be increased at varied intervals if pain relief is inadequate and dosage is tolerated. Doses and intervals titrated to effect.

bid, twice a day; IU, international unit; PO, by mouth; q 12 h, every 12 h; q 4 wk, every 4 wk; qam, every morning; qd, every day; qhs, at bedtime; qid, four times a day; tid, three times a day.

The authors will not assume liability for this table. No medication should be given until the complete prescribing recommendations, drug use indications, and potential side effects that are listed in the package insert with the product or contained in a drug reference manual are reviewed and understood thoroughly by the physician and patient. Drug dosages may need to be modified for the elderly.
Nerve block by itself, however, is not effective in relieving most common procedures in the management of chronic pain. Blocking the nerve with a local anesthetic agent is one of the Anesthetic Procedures

Use of Opioids for the Treatment of Pain: Consensus Documents from Formulation or Routes of Administration, and All Reports of Nonopioid Treatments, Insistent on Rapid-onset concerns are the use of analgesic medications for other than analgesic purposes, and use of nonprescription psychoactive drugs in addition to prescribed medications. Other behaviors that may raise concerns are the use of analgesic medications for other than analgesic effects (such as sedation), an increase in energy, a decrease in anxiety, intoxication, noncompliance with recommended nonopioid treatments or evaluations, inconsistent on rapid-onset formulation or routes of administration, and all reports of no relief whatsoever from any nonopioid treatments (276).

Anesthetic Procedures

Blocking the nerve with a local anesthetic agent is one of the most common procedures in the management of chronic pain. Nerve block by itself, however, is not effective in relieving pain completely for a long period in the majority of patients. Therefore, nerve blocks should be considered as only one of the therapeutic modalities used in the multidisciplinary pain clinic. Other factors, such as psychological problems and associated muscle tightness and weakness, should be treated by using other appropriate modalities. Nerve blocks are helpful in many patients by interrupting the pain process. It is the experience of many clinicians that when pain is temporarily interrupted by a local anesthetic block, often the patient’s pain is permanently relieved. Nerve blocks are also useful in delineating the pain mechanisms and in blocking the pain when the patients are required to take part in physical therapy to mobilize the muscles and joints. Patients who have a nerve block followed by appropriate physical therapy usually have excellent results (277). The best results have been shown in patients requiring manipulation and mobilization of the knee and other joints. Continuous nerve block catheters can be placed perioperatively or in the outpatient setting to assist with pain control and physical therapy. There are various nerve block techniques used in pain clinics; the most common and useful include epidural use of a local anesthetic or opioid and peripheral nerve block (140, 277). For more information on injection techniques, the reader is referred to Chapters 67 and 68.

Administration of a local anesthetic or opioid can provide prolonged relief by placement of a catheter in the epidural space, which can be left in place for several days to a few weeks. A local anesthetic or opioid can be administered intermittently or by continuous infusion, providing somatic and sympathetic block and analgesia for physical therapy. This is the most commonly applied technique in patients who have low back and lower extremity pain while they undergo physical therapy and mobilization (277).

Peripheral nerve block, such as suprascapular nerve block, is very useful in patients who have shoulder discomfort, frozen shoulder, or shoulder pain of other etiology. Patients can tolerate stretching and physical therapy after a suprascapular block with local anesthetic. In the upper extremities, brachial plexus blocks, especially continuous blocks performed by axillary or supraclavicular route, are of great value in patients requiring physical therapy. Other peripheral blocks, such as the lateral femoral cutaneous nerve block for patients with meralgia paresthetica, the femoral nerve block for patients with thigh and knee pain, and sciatic nerve and intercostal blocks, have been extremely useful in managing patients with chronic pain.

“Depo” types of steroids, injected into the epidural space, have been extremely useful in relieving nerve root irritation and inflammation in patients who have a herniated disk. They also have been effective in patients who have nerve root irritation secondary to radiation or a malignancy. Steroid preparations commonly used are methylprednisolone (40 to 80 mg), betamethasone (3 to 6 mg), dexamethasone (4 to 8 mg), and triamcinolone (25 to 50 mg). Although steroid preparations may be used to produce anti-inflammatory action, many of the steroid preparations contain various preservatives, such as benzyl alcohol, which may produce serious side effects, including paralysis. Agents with preservatives that are designated for

| TABLE 49.8 Definitions Related to the Use of Opioids |

| Addiction (psychological dependence) — Addiction is a primary chronic neurobiological disease with genetic psychosocial and environmental factors influencing its development and manifestations. It is characterized by behaviors that include one or more of the following: impaired control over drug use, compulsive use, continued use despite armed craving. |
| Physical dependence — Physical dependence is an adaptation that is manifested by a drug class-specific withdrawal syndrome that can be produced by abrupt termination, rapid dose reduction, decreasing blood level of the drug, and/or administration of an antagonist. |
| Tolerance — Tolerance is a state of adaptation in which exposure to a drug induces changes that result in a diminution of one or more of the drug’s effects over time. |
| Diversion — Diversion is the illegal use or inappropriate use of medications. Diversion of controlled substances should be the concern of every health professional, but efforts to stop diversion should not interfere with prescribing opioids for pain management. Attention to patterns of prescription requests and prescribing of opioids is part of an ongoing relationship between a patient and health care provider and can decrease the risk of diversion. |

Many of the medications used in pharmacological intervention in pain may produce dependence, and some may lead to addiction in susceptible individuals. These medications include angiotensin, cannabinoids, muscle relaxants, opioids, sedatives, stimulants, and steroids. The clinical implications in management of physical dependence, tolerance, addiction, and diversion of medications are discrete and different phenomena (Table 49-8). Behavior suggestive of addiction may include inability to take medication according to agreed-upon schedule, taking multiple doses together, frequent reports of lost or stolen prescriptions, doctor shopping, isolation from family and friends, and use of nonprescription psychoactive drugs in addition to prescribed medications. Other behaviors that may raise concerns are the use of analgesic medications for other than analgesic effects (such as sedation), an increase in energy, a decrease in anxiety, intoxication, noncompliance with recommended nonopioid treatments or evaluations, inconsistent on rapid-onset formulation or routes of administration, and all reports of no relief whatsoever from any nonopioid treatments (276).
articular injections should not be used as epidural agents. Only preservative-free steroid preparations have been extensively used in the epidural space without producing significant neurologic damage (278). However, any particulate-based steroid can cause embolic spinal cord infarction if injected into the spinal arterial supply.

The steroids are injected into the epidural space close to the involved root. These injections can be performed at any level, including cervical, thoracic, or lumbar. The steroid preparations stay in the epidural space for 2 to 3 weeks. Repeat injections, if necessary, are given a minimum of 2 to 4 weeks apart. Many clinicians give a series of three injections regardless of successful response after the first injection. The preferred procedure is to administer one injection and wait 2 weeks in order to assess the patient's response. If the patient is significantly pain free, no further injections are administered. If there is minimal or limited response, then further injections are administered. If the patient does not respond to two or three epidural steroid injections, or the patient receives only short-term pain relief, then steroid injections are discontinued. Frequent injections of epidural steroid can produce problems related to chronic steroid administration as well as a remote possibility of infection.

Many clinicians combine the steroid injections with the administration of a local anesthetic agent. When this is done, the patient needs to be observed because the sympathetic block may produce postural hypotension, and the patient may become unable to ambulate due to motor block. No single technique has been proven to be more successful than another. Many clinicians use the caudal approach, using large volumes of local anesthetic agent or saline mixed with the steroid. This gives a 60% success rate as opposed to an 80% or higher success rate with a lumbar epidural technique, because caudally administered agents may not reach the site of pathology at L4 to L5 or L5 to S1 levels in significant concentration, owing to the leakage through the sacral foramina. Epidural steroid injection is also a useful technique in patients who have neural irritation. In addition, steroids may be injected into the subarachnoid space, especially in patients who have had multiple surgeries and in whom the epidural space has been obliterated. These patients, especially those with arachnoiditis, show significant pain relief. Epidural steroids can produce an initial increase in pain for 8 to 24 hours.

**Clinical Research in Chronic Pain Treatment**

Multiple reports have been published providing guidelines for clinical research in pain (3,121,200,279,280). There is continued need for more rigorous scientific research on most aspects of chronic pain, including natural history, etiology, diagnosis, therapy, and the cost effectiveness of treatment. Determining the efficacy of pain treatment from the medical literature is difficult because studies fail to meet basic scientific standards for internal validity and applicability of the results. Much research has been based on a collection of anecdotal or highly selected cases from which the true outcome of treatment is impossible to extrapolate.

Some critics continue to suggest that researchers often neglect factors in chronic pain treatment. These factors include the high selection factors of patients in pain clinics, medical exclusion criteria, financial exclusion criteria, patient refusal to accept treatment, and patient attrition (281). Outcome studies that include the brief duration of follow-up periods, vague criteria used for establishing the success of the therapeutic intervention, and other factors cast doubt on the applicability of the results to the general population. In examining the significance of intervention in outcome studies, it is important to consider the limiting factors imposed by the methodology.

The diversity of techniques in pain research has resulted in limited consensus regarding the appropriate diagnosis and treatment of many chronic pain syndromes. Unorthodox therapies continue to propagate, with few standard treatments shown to be more effective than no medical intervention or placebo. For progress to be made in the management of chronic pain, reasonable research design must be incorporated into the evaluation of treatment efficacy. Contemporary studies should assess outcomes measured within recognized domains of clinical significance, including pain, physical functioning, emotional functioning, participant ratings of improvement and satisfaction with treatment, symptoms and adverse events, and participant disposition (282,283).

**COMMON PAIN SYNDROMES: DIAGNOSIS AND TREATMENT**

**Myofascial Pain**

Myofascial pain syndromes are commonly seen when evaluating and treating patients for chronic pain. Trigger points are characterized by pain originating from small circumscribed areas of local hyperirritability and myofascial structures, resulting in local and referred pain (178). The pain is aggravated by stretching the affected area, cooling, and compression, often giving rise to a characteristic pattern of referred pain (179,180). Although the exact pathophysiology of the trigger point phenomenon has not been identified, myofascial pain syndromes appear to be initiated by trauma, tension, inflammation, and other unidentified factors. The trigger point acts as a source of chronic nociception. The resultant muscle dysfunction and altered mechanics lead to the referred pain and associated phenomenon.

Trigger points may occur in any muscle or muscle group of the body. They are commonly found in muscle groups that are routinely overstressed or those that do not undergo full contraction and relaxation cycles. In the upper body, the groups of muscles involved commonly include the trapezius, levator scapulae, and infraspinatus. In the lower body, they include the gluteal, tensor fasciae latae, quadratus lumborum, and gastrocnemius muscles.

Trigger points are best located by deep palpation of the affected muscle so as to reproduce the patient's pain symptom both locally and in a referred zone. When they are present,
passive or active, stretching of the affected muscle routinely increases the pain. The muscle in the immediate vicinity of the trigger point is often described as rropy, tense, or having a palpable band. Compared with equivalent pressure in palpation to normal muscle, the trigger point region displays isolated bands, increased tenderness, and referred pain.

The most reliable method of treating trigger points consists of routine, regular stretching to restore the normal resting length of the muscle. Methods to interrupt the pain cycle include injection or needle stimulation of the hypersensitive trigger points (282–287), coolant sprays (178), relaxation therapy, and pressure techniques (285). Botulinum toxin injection has been utilized to provide pain relief for weeks to months by selectively weakening the painful muscles (288). After interrupting the pain cycle, the treatment is directed at restoring the normal resting muscle length with a regular routine stretching program of the involved muscle groups. This may be accomplished with physical modalities including heat and electrical stimulation, or cold, along with the correction of poor body mechanics (289). Psychological intervention may be necessary if long-standing stress and tension are the underlying cause of the problem. A long-term home modality and stretching program is essential in the management of patients with myofascial pain. Attention to body mechanics, stress, and daily routines may significantly alter their functional capabilities. For more information on myofascial pain, the reader is referred to Chapter 36.

**Cumulative Trauma Disorders**

Repetitive overuse is a type of disorder identified as a causative factor of many industrial injuries. This is a common problem seen in the rehabilitative industrial practice referred to as repetitive motion disorders, occupational overuse injuries, or repetitive strain injuries. Cumulative trauma disorders now account for greater than 50% of all workers’ compensation claims (290). Upper-limb cumulative trauma disorders are far more common than lower-limb trauma disorders.

Etiology of cumulative trauma disorders appears to be repetitive motion and stress, resulting in microtrauma primarily to the muscle and tendon. Less often, this microtrauma involves the ligaments, joints, cartilage, bones, and musculoskeletal structures. Presenting symptoms may include tendinitis, muscle strain, ligamentous injury, bursitis, myofascial pain, compression neuropathies, and intravertebral disc disease.

Treatment of cumulative trauma disorders often requires aggressive intervention and adaptation of the workplace (290). If a patient’s problem fails to resolve, careful screening as to the aspects of repetitive motion needs to be assessed, and careful discussion with the patient’s supervisor/employer needs to occur in order to determine an adaptive environment to bring about satisfactory problem resolution. An altered work environment should be designed for the worker to perform different or rotating work tasks, using different muscle groups that in turn will minimize cumulative trauma disorders. For more information on cumulative trauma disorders, the reader is referred to Chapter 36.

**Pain from Peripheral Neuropathy**

Pain is a common feature of peripheral neuropathy that is due to diabetes, amyloidosis, alcoholism, polyarteritis, Guillain-Barré syndrome, brachial neuritis, acquired immune deficiency syndrome (AIDS), porphyria, and riboflavin deficiency. This pain may be of either a constant or an intermittent nature and is often described as burning, aching, or lancinating. It may occur with or without signs of sensory loss, muscle weakness, atrophy, or reflex loss (291, 292).

Evidence-based recommendations for the pharmacological management of pain from peripheral neuropathy include anticonvulsants (pregabalin, gabapentin), tricyclic antidepressants (TCAs such as amitriptyline, desipramine), dual-action selective serotonin and norepinephrine reuptake inhibitors (SSNRIs such as venlafaxine, duloxetine), and topical lidocaine as first-line medications in alleviating the pain from peripheral neuropathy (253, 254, 259–261, 293–296). These medications may be used in combination, as each may have a different mechanism of action. It is often suggested that rehabilitative interventions include exercise, desensitization training, TENS, and medications. NSAIDs are often used in association with the above adjunctive agents and simple physical measures (202, 297).

Pain resulting from alcoholic neuropathy typically resolves after the correction of nutritional deficiencies. Pain resulting from polyarteritis may resolve after corticosteroid treatment, whereas pain secondary to cryoglobulinemia may resolve with plasmapheresis. Pain from brachial neuritis and other self-limiting conditions resolves spontaneously in weeks to months. Pain secondary to diabetic neuropathies rarely resolves completely. Drugs that induce painful peripheral neuropathies include isoniazid and hydralazine, which cause a decrease in tissue levels of pyridoxal phosphate, as well as nitrofurantoin, which has neurotoxic effects.

The patient with chronic pain due to peripheral neuropathy has many of the associated problems of chronic pain, including depression, inactivity, disuse syndrome, and significant alteration of lifestyle. Psychological intervention and physical therapy are important aspects of treatment in these patients. Medication, conventional physical therapy modalities, TENS, and general conditioning programs are often helpful, as are psychological programs and other nonpharmacologic methods of pain control.

**Complex Regional Pain Syndrome**

A contributing factor in many chronic pain syndromes is CRPS resulting from overactivity of the sympathetic nervous system. This is often reported as continuous burning pain in an extremity after trauma. CRPS I or reflex sympathetic dystrophy refers to cases of minor injury or no injury, with resulting overactivity of the sympathetic nervous system. CRPS II or causalgia refers to partial injury to a major nerve, followed by the symptoms of sympathetic system overactivity. Examples include shoulder-hand syndrome, post-traumatic edema, Sudeck’s atrophy, and various other syndromes in which sympathetic overactivity seems to be the primary etiologic factor.
Dysfunction of the sympathetic nervous system is by no means the sole component of CRPS (298,299).

The etiology of the sympathalgia is not distinct. The most common aspect of a sympathalgia is burning pain. Associated with the hyperpathia are hypersensitivity to touch and relief of pain with an appropriate sympathetic nerve block. The patient also may show evidence of overactivity of the sympathetic nervous system, including hyperhidrosis and vasoconstriction, resulting in the cooling of the extremity. When accompanied by disuse, this may produce trophic changes, including shiny thin skin, loss of hair, and demineralization of bone (300).

Diagnosis of CRPS is difficult. The patient’s pain is usually diffuse and does not correspond to dermatomal or peripheral nerve patterns. Thus, these patients are often diagnosed as having psychogenic pain. Anatomic or pharmacologic nerve blocks may establish the diagnosis. Psychogalvanic reflex tests having psychogenic pain. Anatomic or pharmacologic nerve blocks may establish the diagnosis. Psychogalvanic reflex tests and thermography may be useful in documenting sympathetic nervous system hyperactivity.

Pharmacological treatments have included NSAIDs (tramadol), sodium channel blockers (intravenous [IV] lidocaine), N-methyl-D-aspartate [NMDA] receptor blockers (ketamine, dextromethorphan, memantine), calcium-regulating drugs (calcitonin, clodranate, alendronate), free-radical scavengers (transdermal dimethylsulfoxide [DMSO], oral N-acetylcysteine [NAC], α₂-adrenoceptor agonist (clonidine), and oral phenytoin (301,302).

The most effective invasive treatment of CRPS consists of an appropriate sympathetic nerve block using a local anesthetic agent (278,303). If the problem originates in the head, neck, upper limb, or upper thorax, a stellate ganglion or a cervical sympathetic block at the C6 level may be useful (304). If the pain is in the upper abdominal area, sympathetic blockade can be achieved using the celiac plexus block. Pain originating in the lower limb requires a lumbar paravertebral sympathetic block at the L2 level (277). Patients with a history of long-standing pain may receive only temporary pain relief from the local anesthetic blocks. IV regional bretylium has been reported to provide a statistically significant decrease in CRPS symptoms (305); however, neither IV regional guanethidine nor reserpine provided any better relief of CRPS symptoms than placebo in randomized, double-blind studies (306,307). If the patient does not receive lasting pain relief from sympathetic blocks or the IV regional technique, then permanent interruption of the sympathetic pathways may be accomplished by a neurolytic injection of the sympathetic trunk or surgical sympathectomy. Unfortunately, neurolytic techniques are not recommended for the long-term control of neuropathic pain that is due to nerve regrowth, central spinal reconfiguration, and neuroma formation. Pain normally returns in 6 to 12 months following neurolytic intervention (303). Alternatively, patients with neuropathic pain may also benefit from spinal cord stimulation. The system can be implanted if trial stimulation is successful (308–310).

The patient may also have nonsympathetically mediated neuropathic pain and musculoskeletal pain stemming from the initial injury and subsequent disuse. Often, the patient receives relief of pain at rest with the sympathetic pain, but when the limb is moved during physical therapy, the patient experiences significant pain. This pain is secondary to tight muscles or stiffness of the joint, where the noxious input is carried through somatic fibers. Peripheral nerve, plexus, or epidural block can treat this type of pain, and can be achieved with single injections or continuously with percutaneous catheters. In conjunction with sympathetic blocks or these other procedures, intensive physical therapy should be instituted (311) to maximize the benefits.

When treating a patient with long-standing pain that is due to CRPS, it is important to consider the psychological aspects. Patients who develop RSD may have enhanced sympathetic pain, secondary to psychological tension. Post-traumatic stress disorder is also a common but unidentified psychological problem in CRPS patients. Appropriate psychological and psychiatric consultation for relaxation training and biofeedback, in addition to other treatment modalities, is important. Psychological factors must be adequately treated; otherwise, the patient is unlikely to improve regardless of other interventions (312).

**Phantom Limb Pain**

Phantom limb pain involves an amputated portion of the body. The etiology of phantom pain appears to be related to peripheral, spinal, and central factors. Peripheral components of phantom limb pain appear to be the result of sensation due to loss of previously present peripheral nerve activity triggering changes in the central nervous system. Development of abnormal activity in a neuroma or dorsal root ganglion results in pain fiber transmission and an alteration in ion channel activity at site of injury resulting in pain fiber activation. Spinal-level factors involved in phantom limb pain include deafferentation of neurons and their spontaneous and evoked hyperexcitability resulting from the loss of large A-afferent nerve fiber input in the dorsal horn cells with an unopposed C-fiber input. Cortical reorganization has been observed in the primary motor and somatosensory cortex in humans after amputation. These reorganizational changes appear to be a significant factor in the occurrence and intensity of phantom limb pain. The remapping of regions of the brain receiving touch, pain, and temperature input from the amputated limb to other parts of the brain appears to be an etiology of phantom limb pain. Pain may be continuous, in character with intermittent exacerbations. It is often reported by patients as cramping, aching, or burning, with occasional superimposed electric-like components. Reported incidence of phantom limb pain is highly variable often due to the setting and circumstances in which the data were collected. Prospective research suggests approximately 82% of patients experience phantom pain following amputation. It is estimated that 65% experience phantom pain at 6 months, and 59% experience phantom pain 2 years after amputation. Severe disabling phantom limb pain incidence is reported at 10% after several years. Of note are reports that severe pain occurs more often with proximal amputations.
The current data do not suggest a predisposition for phantom pain among traumatic amputees, elderly amputees, those with pain in the amputated limb before amputation, or poor preamputation interpersonal relationships (313). Phantom limb pain does not appear to be correlated with amount of time after amputation or use of prosthesis. The incidence of phantom limb pain is independent of age, gender, level of amputation, employer, or age after 8-years old (313–318). Most studies have found no significant relationships between preoperative limb pain and persistent phantom pain (319,320). Multiple modalities, adjuvant medications, and anesthetic and surgical procedures have been used in the treatment of phantom limb pain with varying long-term success. Although more than 70 methods of treating phantom limb pain have been identified, successful treatment of persistent types of pain is not commonly reported (292,294,321–327). Transcutaneous nerve stimulation, tricyclic antidepressants, anticonvulsants, calcitonin, lidocaine, mexiletine, and mirror therapy have been used with varying success. Chemica sympathectomy or neurosurgical procedures also have had variable success. Treatments yielding a temporary decrease in pain include analgesics, anesthetic procedures, stump desensitization, physical modalities, and sedative/hypnotic medication. One survey reported that treatments reducing residual limb pain problems also resulted in decreased phantom pain (317). Studies of various therapeutic regimens have reported up to a 70% efficacy in the treatment of phantom limb pain, but reviews of most with long-term follow up suggest a limited response to any intervention. Animal studies report that functional reorganization of the primary somatosensory cortex correlates with the intensity of phantom limb pain after amputation of a limb (328). Underlying mechanisms for phantom limb pain include reorganization that appears to occur at the dorsal horn of the spinal cord, in the thalamus, and in the cortex.

Residual limb pain is pain at the site of the limb amputation. Residual limb pain is common, with approximately 57% of patients experiencing residual limb pain following amputation. It is estimated that 22% experience residual pain at 6 months, and 10% experience residual pain 2 years after amputation. Severe disabling residual limb pain incidence is reported at 5% after several years (313–316). Residual limb pain may be rated as more severe than the patient’s phantom limb pain. Residual limb pain was rated as the worst problem by 33% of the patients, phantom limb pain by 24%, back pain by 17%, and sites other than the involved limb by 26%. In the early evaluation of phantom limb pain, it is important that stump pain that is due to a neuroma is ruled out as the etiology of the pain reports. The pain caused by a neuroma is sharp, often jabbing, pain in the stump and is usually aggravated by pressure or by infection in the residual limb. Pain is often elicited by tapping over a neuroma in a transected nerve. The increased sensitivity of sprouts from cut peripheral nerves to noradrenaline and adrenaline may partially explain why adrenergic-influenced emotional states (i.e., stress or anxiety) occasionally provoke attacks of phantom limb pain.

Neuroma and Scar Pain

Injury to the nerve with a resulting neuroma or entrapment of the branches of the nerve in scar tissue can produce disabling pain. Neuromas are suspected when numbness appears in the distribution of a particular nerve and when pain is produced by palpation of the neuroma. It has been shown that the nerve fibers in the neuroma develop receptors, which respond to catecholamines with spontaneous firing and pain production. Painful neuromas are difficult to treat; many patients continue to have pain despite multiple attempts at surgical excision of the neuroma. Suspected scar pain can be evaluated by picking up the scar from the deeper tissues with two fingers and palpating. If this does not reproduce the patient’s pain, the pain is probably not originating in the scar tissue. Pain that is due to neuroma or scar tissue can be associated with CRPS. Diagnosis can be established by infiltration of the scar or the neuroma with a local anesthetic agent, resulting in complete pain relief (141).

Repeated injection of a local anesthetic agent has proven to be an extremely useful technique. This should be followed by appropriate physical therapy to the scar, usually ultrasound followed by stretching or deep massage of the scar. This type of treatment has provided permanent or prolonged pain relief for longer than 6 months in many patients. When the local anesthetics, with or without steroids, do not provide prolonged relief, other methods should be considered. Cryoanalgesia using a cryoprobe and freezing the neuroma for 1 minute at 20°C has been used with good success (329). The advantage of using a cryoprobe lies in the fact that it is a physical method of blocking the nerve without producing further neuroma or neuritis. Neurolytic agents such as phenol or alcohol have been used to relieve neuroma and scar pain. Incomplete block with these agents can result in neuritis, producing severe pain. These neurolytic techniques should be used only after repeated injections of local anesthetics produce consistent pain relief proportional to the duration of action of the local anesthetic agent. Although surgical revision of the scar is often considered, it is not very successful when the scar cannot be stretched out and there is significant nerve entrapment.

Cancer Pain

Cancer pain is acute or recurrent pain secondary to a chronic condition. Unbearable pain is one of the greatest fears and a major source of morbidity for patients with cancer (294, 330–332). Clinical experience suggests that patients with cancer pain are treated most effectively with a multidisciplinary approach, including multiple modalities, appropriate analgesic drugs, neurosurgical and anesthetic procedures, psychological intervention, and supportive care (12–14). The goals of pain therapy for cancer patients are a significant relief of pain to maintain the functional status they choose, a reasonable quality of life, and a death relatively free of pain.

Noninvasive treatment measures are often used singularly for mild to moderate pain and in combination with drug therapy for moderate or severe pain. Commonly used noninvasive measures include cutaneous stimulation, thermal modalities,
and behavioral intervention. Their advantages include low risk of complications, low cost, and few serious side effects.

Nonopioid, opioid, and adjuvant analgesic drugs are the primary therapy for patients with cancer pain. Anesthetic, psychiatric, behavioral, and occasionally neurosurgical approaches are commonly used with pharmacologic intervention. This combination of treatment is estimated to provide adequate pain relief in at least 90% of the patients with cancer pain (333).

The World Health Organization has recommended a systematic approach for the selection of pharmacologic agents for treating patients with cancer pain (334,335). The protocol is based on the premise that doctors and health care professionals should know how to use a few simple drugs well. A three-step analgesic ladder is the scheme that is used for analgesic selection and is summarized as follows:

1. Those patients who have mild to moderate pain should be initially treated with a nonopioid analgesic (see Table 49-4), such as aspirin or one of the other NSAIDs, and adjuvant analgesic (see Table 49-7) if indicated.

2. Patients who have moderate to severe pain or are not able to achieve adequate relief with step 1 should be started on a weak oral opioid (see Table 49-5), such as codeine, along with a nonopioid and an adjuvant analgesic if indicated. Pain secondary to soft tissue and bone metastases usually has a limited response to opioids. In these situations, NSAIDs are indicated at all levels of treatment.

3. For patients who have very severe pain or do not achieve adequate relief on step 2, the regimen should include treatment with a potent opioid, such as morphine or methadone (see Table 49-6) with or without an NSAID, and with or without an adjuvant analgesic if indicated (336).

Opioid analgesics are the drugs of choice for relief of intractable pain due to cancer (12). Slow-release morphine and methadone are extremely useful analgesics in the treatment of cancer pain. The object is to titrate the level of analgesic to the optimal dose that prevents the recurrence of pain. Oral medications should be used whenever possible because they facilitate ambulatory care, encourage greater independence, and do not represent heroic intervention to the patient. It is important to remember that the maximum recommended dosages of opioid medications were derived mainly from postoperative parenteral single-dose studies and are not applicable to administration by mouth in the long-term treatment of pain in advanced cancer. Dependency and respiratory depression should not be feared because they are seldom a problem. The most serious side effects of drug therapy, which should be prevented or treated, include constipation and nausea.

Supportive care is designed to maintain the cancer patient in an outpatient setting, with pain a common cause for readmission to the hospital. Nerve blocks and neurolytic procedures are most often used to provide pain relief in the thoracic and abdominal regions (140–141,278). These procedures are often performed in the outpatient setting and often provide pain relief for up to 6 months. Supportive care may involve a hospice, hospital-based care teams, visiting nurses, and social services. The primary goal of the physician is to maintain the quality of life for the cancer patient to the end, allowing the patient a death with dignity.

It is essential that during a cancer patient’s hospitalization, outpatient treatment, and everyday activities, they maintain their functional skills, strength, and mobility. Progressive immobilization is an insidious aspect of this disease and is often iatrogenic. Although radiation and chemotherapy may transiently render a patient unable to perform activities, the patient needs to be consistently evaluated and involved in active physical, occupational, corrective, and recreational therapy programs. The psychiatrist may offer multiple methods of combating immobility and its morbidity. For additional information on cancer rehabilitation, the reader is referred to Chapter 44.

There are multiple invasive measures to control cancer pain. The most common is the surgical removal of all or part of the tumor with the hope of relieving pain and obtaining a cure. Radiotherapy or chemotherapy also may relieve pain by shrinking the tumor. Common invasive anesthetic procedures include trigger point injections and nerve blocks. Myofascial pain syndromes are often the result of inactivity and disuse and are commonly found in the cancer patient. Neuropathic pain may result from direct tumor invasion or as a side effect of chemotherapy. Patients who have side effects with large doses of opioids may derive long-term benefit with a chronically implanted epidural catheter or an intrathecal catheter and pump to administer opioids and/or clonidine (337,338).

Multiple neurosurgical procedures have been used for the control of cancer pain (339).

Herpetic Neuralgia

Herpes zoster (shingles) is a reactivation of the varicella (chickenpox) virus, which has remained latent. The viral inflammation of the dorsal nerve root and ganglion causes vesicle formation and severe burning, aching, and lancinating pain in a radicular distribution. During the acute stage, a course of antiviral medication, along with NSAIDs and sympathetic blocks, provides excellent pain relief (340). The majority of patients recover from the acute episode in approximately 2 weeks without sequelae. However, some patients develop postherpetic neuralgia secondary to deafferentation from neuronal cell death with and without allodynia. Postherpetic neuralgia is uncommon in patients younger than 40 years of age but occurs in more than 50% of patients older than 60 years of age after an episode of herpes zoster (341). Skin biopsies of patients after herpes zoster show a decrease in density of cutaneous nerve endings in the patients with postherpetic neuralgia when compared with patients without postherpetic neuralgia (342–343). It has been suggested that an innervation level of approximately 650 neurites per mm² may be necessary to avoid postherpetic neuralgia (344).

Multiple reports indicate that postherpetic neuralgia may be prevented if the patient is treated with a sympathetic
block within 1 month after the onset of herpes zoster (140). Postherpetic neuralgia is, thus, a preventable syndrome, but it is a difficult problem once established. A significant number of patients experience pain relief with sympathetic blocks. If the syndrome is allowed to progress untreated for 3 months to a year, it becomes more difficult to treat (269).

The most effective treatment for the management of established postherpetic neuralgia is the use of a tricyclic antidepressant, such as amitriptyline, in small doses (269,293). Most patients obtain pain relief with 25 to 75 mg administered at bedtime. If the patient does not get complete pain relief with administration of tricyclic antidepressants alone, fluphenazine may be added, beginning with 1 mg at bedtime and progressing as needed to a maximum dosage of 1 mg three times a day. Anticonvulsants (such as gabapentin, phenytoin, and carbamazepine) have been administered with variable results (259, 277). Pain reduction has been reported with topical lidocaine (345–347) for herpetic neuralgia. Opioids and tramadol are considered second-line treatments in most circumstances. Topical capsaicin and valproate are suggested as third-line interventions. Peripheral nerve blocks and destructive neurosurgical procedures have not proved useful in treating established postherpetic neuralgia.

### Pain in Spinal Cord Injury

Pain related to spinal cord injury is often complex and multifactorial. Spinal cord injury pain may be broadly classified into nociceptive pain (musculoskeletal and visceral) and neuropathic pain located above level, at level, and below level of the spinal injury (Table 49-9). The reported prevalence of chronic pain in the population with spinal cord injury is variable with an average of 65%, approximately one-third rate the pain as severe (348).

Many of the nociceptive types of pain are variations of those found in patients without spinal cord injury and are treated in a similar fashion with adaptation to the underlying injury. However, patients with spinal cord injury are more susceptible to mechanical instability of the spine that is due to trauma of the musculoskeletal system, muscle spasm pain that is due to trauma to the neurologic system, as well as both overuse and nerve-compression syndromes primarily involving the upper extremities.

Mechanical instability of the spine is usually best treated with NSAIDs and opiates (as needed), as well as immobilization using orthosis or surgical fusion. These are effective treatments in most patients. Pain that is due to muscle spasm is best treated with antispasmodics, motor point blocks, or botulinum toxin to decrease spasticity. Secondary overuse syndromes are often relieved by rest and NSAIDs, with significant adaptation by the patient and modification of environmental factors. Nerve-compression syndromes are often relieved through the use of orthotics or surgical decompression. Visceral nociceptive pain is often due to afferent input from the sympathetic nervous system and may be due to underlying pathology. A headache that is due to autonomic dysreflexia is a medical emergency and should be treated appropriately.

### Table 49.9 Classification of Pain Related to Spinal Cord Injury

<table>
<thead>
<tr>
<th>Nociceptive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Musculoskeletal (often dull, aching, movement, decreased with immobilization, responsive to NSAIDs and opioids)</td>
</tr>
<tr>
<td>Secondary overuse syndromes</td>
</tr>
<tr>
<td>Bone, joint, muscle trauma, or inflammation</td>
</tr>
<tr>
<td>Mechanical instability</td>
</tr>
<tr>
<td>Bone joint, muscle trauma, or inflammation</td>
</tr>
<tr>
<td>Scar infection</td>
</tr>
<tr>
<td>Muscle spasm</td>
</tr>
<tr>
<td>Visceral structures (often dull, cramping, burning, constant but fluctuating pain in the abdomen with preserved innervation)</td>
</tr>
<tr>
<td>Regional calculus</td>
</tr>
<tr>
<td>Bowel or sphincter dysfunction</td>
</tr>
<tr>
<td>Dysreflexic headache</td>
</tr>
<tr>
<td>Neuropathic (often sharp, shooting, burning, electrical sensation (hyperalgesia, hyperesthesia))</td>
</tr>
<tr>
<td>Above level (located in region of sensory preservation)</td>
</tr>
<tr>
<td>Compressive mononeuropathies</td>
</tr>
<tr>
<td>Complex regional pain syndromes</td>
</tr>
<tr>
<td>Routinely treated as neuropathic pain of peripheral nerve origin</td>
</tr>
<tr>
<td>At level (located in segmental pattern of level of injury)</td>
</tr>
<tr>
<td>Nerve root compression (including cauda-equina)</td>
</tr>
<tr>
<td>Syringomyelia, spinal cord trauma/ischemia (transitional zone, etc.)</td>
</tr>
<tr>
<td>Presents with an ascending neurological deficit often with alterations in pain/temperature</td>
</tr>
<tr>
<td>Dual level cord and root trauma and double lesion syndrome</td>
</tr>
<tr>
<td>Below level (located diffusely below level of injury)</td>
</tr>
<tr>
<td>Spinal cord trauma/ischemia (central dysesthesia syndrome, etc.)</td>
</tr>
<tr>
<td>Phantom sensations</td>
</tr>
</tbody>
</table>

With neuropathic pain, evidence suggests that trauma and neural alterations of the pain pathways are primarily involved (349–353). Patients describe their pain as having one or more of the following components at the level of injury: lancinating, burning, or stabbing pain in a radicular distribution. This pain may also be related to spinal instability or nerve-root entrapment. Pain is typically relieved by opioids, neuropathic pain medications, and spinal decompression, if necessary. This includes cauda-equina injury, which is neuropathic pain of peripheral nerve origin that has a burning sensation routinely affecting the groin and bilateral lower extremities. Segmental deafferentation pain often occurs at the border of sensory loss. Transitional zone pain is often associated with allodynia and hyperalgesia and may respond to neuropathic pain medications, epidural or somatic root blocks, dorsal root entry zone (DREZ) procedures, or spinal cord stimulation. The combination of spinal cord and root trauma often results in severe burning pain over a region that is otherwise denervated. Pain with root evulsion may be responsive to neuropathic pain medications, DREZ procedures, or spinal cord stimulation. The segmental deafferentation pain may be secondary to spinal cord trauma or ischemia. Neuropathic pain below the
level of entry is often referred to as deafferentation or central dysesthesia syndrome, commonly described as burning, tingling, numb, aching, or throbbing. Pain is usually constant and unrelenting. Pain may be relieved by neuropathic pain-relieving medication, intrathecal opioids, and clonidine, and occasionally by dorsal column stimulator. This is located diffusely below the level of injury (354–358). Studies indicate that 66% of all patients with spinal cord injury have pain that is mild to moderate in severity. Approximately 37% of patients with high thoracic and cervical lesions and 23% of patients with low thoracic and lumbosacral spinal cord injury lesions experience severe pain (355–361). Pain after spinal cord injury is common, poorly understood, complex, and often difficult to treat. Also note that 50% of patients with cauda-equina syndrome report severe pain. Pharmacologic and nonpharmacologic treatments for pain in patients with spinal cord injury have a minimal number of subjects per treatment approach, and only a limited number of studies have been subjected to randomized, controlled trials.

The precise etiology underlying pain in spinal cord injuries is not known, but evidence suggests that trauma-induced alterations of the pain pathways are primarily involved (350–353). Hypersensitivity of the structures in the ascending pathway may play a role. Patients describe their pain as having one or more of the following components: burning in body parts below the injury, a deep, aching sensation over and around the site of injury, and radicular with lancinating characteristics. Burning pain in spinal cord-injured patients may be a variation of deafferentation pain occurring as a result of loss of inhibitory or augmentation of excitatory influences. The most effective treatments of this type of pain include gabapentin and SSRIs and neuroaugmentative techniques (362).

Spinal fracture site pain results from an alteration of body mechanics causing pain-sensitive structures to be stretched or compressed. This mechanical pain may be the result of vertebral end-plate fractures, annulus fibrosus tears, or internal disc herniation after a spinal fracture (363). Fracture site pain or mechanical pain is often exacerbated by activity. NSAIDs, trigeminal neuralgia (9%), peripheral neuropathic pain

(4%), spasticity pain (1%), pain of unknown origin (1%), and nociceptive pain (36%) occurring singularly or in combination (368). Patients often present initially with paroxysmal lancinating and intense burning pain that primarily affects the face, shoulder region, or pelvic girdle. In multiple sclerosis, the central pain usually involves the bilateral lower extremities (368). Treatment of pain resulting from multiple sclerosis has been limited. Minimal response has been shown to tricyclic antidepressants, carbamazepine, and phenytoin. Limited response has resulted with gabapentin (371). If the patients have intrathecal baclofen pumps, neuropathic pain adjuvant medications can be added to the infusion. For additional information on multiple sclerosis, the reader is referred to Chapter 25.

Poststroke Pain
Pain after a cerebral vascular accident may be secondary to multiple etiologies, including central pain, CRPS, pain that is due to spasticity, and pain from dysfunction of the affected extremities. Central pain is due to a lesion or dysfunction of the central nervous system that is often characterized as an agonizing, burning pain on the side contralateral to that of the lesion. Approximately 90% of all cases of central pain are caused by cerebral vascular accidents, and 78% of these are supratentorial. Central nervous system lesions producing central pain can be extremely small, especially if located in central pathways; approximately 8% of stroke patients report central pain during the first year. Almost 100% of central pain patients are affected by change in temperature perception. Pain is constant in 85% of the patients and intermittent but daily in 15%. Pain is primarily burning, aching, prickling, and lancinating. Patients make the point that central pain is not truly pain, but an unpleasant sensation that drastically affects their quality of life. Neuronal reorganization of the central nervous system architecture is most likely the cause of central pain. The plasticity of the central nervous system may result in changes in receptor structure and function that are likely to be delayed in onset. Central nervous system neurons are responsible for nociceptive perception and can exhibit long-lasting changes and responsiveness after transient or permanent damage in input. Either excitatory or inhibitory mechanisms can be affected, and as this responsiveness changes, so does modulation of the signals. Central pain that is due to thalamic infarction is often characterized as an agonizing, burning pain on the side contralateral to that of the lesion. Pain to minimal cutaneous stimulation as well as aggravation by emotional stress and fatigue are characteristic findings. Sensory alteration is variable in these patients, with minimal findings of motor weakness. Central pain that is due to lesions involving central thalamic spinal tracts may be manifest in pain distributed to the level of the tract involved. This results in loss of pain and temperature perception on the contralateral side at the level below the injury. Central pain of tract origin is similar to pain of thalamic origin but is usually less intense. This pain may be described as burning, pulling, or swelling. Information on central pain syndromes is scarce, with treatments limited in efficacy (372). Treatment options include sympathetic blockade (373), multiple medication
In treatment of central pain, no one pharmacological, surgical, or other therapeutic intervention has proven effective. After surgery, pain always recurs, usually within 3 to 6 months. Initial treatment includes antidepressants (amitriptyline, nortriptyline, trazodone), opioid analgesics, anticonvulsants (carbamazepine, phenytoin, gabapentin), transcutaneous electrical nerve stimulation, and psychological support. Although there are many targets for treatment, with at least five neurotransmitters implicated and mediated through first-, second-, and third-order neurons, no one pharmacological treatment has proven effective; hence, rarely is surgical intervention recommended. Central pain is also noted with Dejerine-Roussy syndrome. The second step of treatment includes adjuvant agents, including clonidine and β-blockers, as well as opioid agonist/antagonists, including talwin and mexiletine. Step three includes dorsal column stimulators, DREZ, ablation, thalamic stimulation, and deep brain stimulation.

Spasticity secondary to stroke may result in pain. This may be treated with medication or nerve blocks. Studies have suggested that patients with spasticity and pain in the hemiplegic upper limb may benefit from selective posterior rhizotomy in the DREZ. Poststroke pain that is due to dysfunction of the affected extremities is most often manifest in the upper extremity. This pain may be the result of shoulder subluxation, decreased range of motion because of adhesive capsulitis, or brachial plexus injury.

Other factors of affected limb dysfunction include bicipital tendinitis, arthritis, fracture, heterotopic ossification, and knee/ankle instability. Appropriate use of modalities, as well as orthotics and assistive devices, is indicated in the resolution if some of these problems are contributing to the patient’s pain. For additional information on stroke rehabilitation, the reader is referred to Chapter 23.

**Fibromyalgia**

FM is a common type of nonarticular, chronic pain involving the musculoskeletal system, with a prevalence ranging from 2% to 12% in the general population and primarily involving middle-aged women. In 1990, the American College of Rheumatology (ACR) established diagnostic criteria for FM that currently remains the standard. FM is described as chronic, widespread nonarticular musculoskeletal pain that may not be attributed to specific structural or inflammatory etiology, with diagnostic markers involving both sides of the body, above and below the waist, and axial skeletal system. In addition, there should be at least 11 tender points among 18 specific sites. It is notable that there have been concerns that the defined criteria are too subjective, and

---

**FIGURE 49-18.** Anatomical locations of tender points in fibromyalgia. At least 11 of 18 tender points must exhibit painful sensitivity to 4 kg of digital palpation pressure (4 kg = pressure to blanch the examiner’s thumb nail to its midpoint). The history of widespread pain for greater than 3 months and above sensitivity to 11 or more tender points has demonstrated moderately high sensitivity (88%) and specificity (81%) for patients with fibromyalgia.
that FM should be classified as a functional somatic syndrome and not a rheumatologic disorder (382).

In addition to pain, other complaints with FM are fatigue, cognitive dysfunction (memory loss, searching for words), psychological distress, sleep disturbance (insomnia and disordered sleep), stiffness, paresthesia, anxiety, and irritable bowel syndrome. FM does not cause joint pathology, but may occur in the presence of other rheumatologic disorders. There are no specific lab abnormalities associated with FM. The pathogenesis of FM is unknown, but evidence suggests that it may be secondary to abnormal processing of sensory stimuli in which sensory input is amplified at the brain and spinal cord, thus leading to an increased experience of pain (383). The sensory detection threshold is normal, with an increased sensitivity to most stimuli in patients with FM. The tender points are locations where the general population is more sensitive to palpation and demonstrate no specific pathology in the patient with FM. However, these points are more tender than normal in patients with FM due to lowered pain threshold (382,384). There is also an increased incidence of depressive and anxiety disorders associated with a common pathophysiological pathway in neurotransmission in patients with FM (385). Factors that trigger FM are frequently identifiable and include trauma, especially to the axial skeleton, surgery, infections, and active or chronic stress. Genetics also appears to play a role, as there is a higher incidence within families (384,386,387).

A multidisciplinary approach to treatment and an individualized treatment plan are important, as response to treatment varies widely. The current treatment protocols alleviate the peripheral sensitivity and diminish the central pain; they are palliation and not a cure. The goal of treatment is to maintain everyday activities and enhance quality of life. The best results are obtained using a combination of pharmacological and nonpharmacological therapies.

**Pharmacological**

Evidence-based recommendations for the pharmacological management of pain from FM include anticonvulsants (pregabalin), tricyclic antidepressants (amitriptyline, desipramine), dual-action, SSNRIs (venlafaxine, duloxetine), and atypical analgesics (tramadol) as first-line medications in alleviating the pain from FM (388–393). These medications may be used in combination, as each may have a different mechanism of action. Current Food and Drug Administration (FDA) approval for treatment of FM includes Pregabalin up to 150 mg orally three times a day and duloxetine up to 60 mg orally twice a day. No evidence suggests the efficacy of corticosteroids, melatonin, NSAIDs, opioids, or thyroid hormone in the treatment of FM (394).

**Nonpharmacological**

Nonpharmacological treatment includes patient education, exercise, and cognitive behavior therapy. Patients need to be aware that this is a chronic, but not life-threatening disorder and that the goal of treatment is to work with the physician and health care team to improve functional abilities and quality of life. The relationship between physician and patient is likely to be long term, optimized with open lines of communication to discover the best treatment for each individual. There is high-quality evidence, based on meta-analysis, that supervised aerobic exercise training at prescribed levels has positive effects on physical function and feelings of well being in patients with FM (395). However, studies also show that compliance with exercise as a treatment is quite low in this patient population (396). Factors influencing the patient’s ability to exercise or barriers to exercise should be addressed. Excessive or over-vigorous exercise programs tend to lead to pain, fatigue, and noncompliance; therefore, an individualized, slowly progressive aerobic exercise plan often decreases pain and fatigue while increasing a sense of well being with improved compliance. Cognitive behavior therapy may be helpful to manage the psychological aspects of FM and may help to manage stressors that trigger or increase symptoms. The reasons for this type of therapy should be fully understood by the patient to be a component of a multimodal approach to their pain (394–397).

Significant research remains to be done in terms of etiology and treatment of FM. In the meantime, the goal is to help patients maintain and improve their functional abilities, as well as enrich their quality of life.

**Low Back Pain**

Low back pain episodically affects nearly 75% of the population in most industrial nations. It is estimated that 15% of the workforce is affected by back pain each year (9). Approximately 90% of these incidences of back pain resolve in 12 weeks or less. It is estimated that the 10% of these workers who develop chronic low back pain (remaining off work for 3 months or longer) are responsible for 80% of the cost (lost wages, compensation, medical expenses, etc.) (398,399).

Strategies to prevent acute low back pain from progressing to chronic low back pain include use of conservative intervention, return to work as soon as possible, or continue working at a job the worker is physically able to accomplish during the recuperation period and rapid adjudication of disabilities/compensation claims (76,77). Comprehensive prevention and diagnostic programs have been identified (121,400). An extensive review of randomized, controlled trials in industrial low back pain over the past 20 years has repeatedly shown methodological limitations for interventions (6,121,187,200,401–403). For additional information on rehabilitation of the patient with spinal pain, the reader is referred to Chapter 33.

**CONCLUSIONS**

The clinical management of an individual with chronic pain is often most challenging and almost always a multifaceted problem. The diagnostic and clinical evaluation, detailed history of the pain report, assessment of the impact of pain on the patient, directed and comprehensive physical examination (to include the patient’s functional capabilities, diagnostic procedures as necessary, as well as consideration for any comorbid-
ties in clinical management) need to be individualized. This should be done to appropriately address each patient’s multifaceted problem in order to alleviate the pain and enhance the patient’s quality of life and functional capabilities. A one-size-fits-all treatment plan for patients with chronic pain of a specific type is rarely effective. Periodic reexamination, review of treatment efficacy, and reevaluation of each individual’s care plan are essential to minimize the pain and optimize the function in each individual. Concise, long-term documentation is important in the long-term management of patients with chronic pain so that treatment strategy can maintain direction in patients with complex disease process and can assist in the periodic review of the efficacy of the patient’s treatment. Most chronic pain patients have been seen by numerous capable physicians without amelioration of their disease process. The task is to evaluate a multifaceted disease process and come up with a comprehensive and structured treatment program for optimal results.

REFERENCES

1. Dicebat Bernardus Carnotensis nos esse quasi nanos, gigantum humeris insinctis, ut postimus plura et ei remota vis, non utique proprii visus acuminet, aut eminentia corporis, sed quia in altum subvenimur et extollimus magnitudinem gigantem transeuntibus a Latin to English Bernard of Chartres used to say that we are like dwarfs on the shoulders of giants, so that we can see more than they, and things at a greater distance, not by virtue of any sharpness of sight on our part, or any physical distinction, but because we are carried high and raised up by their giant size. John of Salisbury, (1120–1180) in Metalogicon (1159).


Spasticity and Muscle Overactivity as Components of the Upper Motor Neuron Syndrome

Spasticity is derived from the Greek word *spasticus*, which means “to pull.” It is a component of the upper motor neuron syndrome (UMNS), which is caused by a lesion proximal to the anterior horn cell; in the spinal cord, brainstem, or brain. It has both positive and negative components. Weakness, paralysis, and fatigue are the negative signs of the syndrome, whereas spasticity, athetosis, hypertonia, release of primitive reflexes, and dystonia are the positive. Hyperreflexia, spread of reflexes beyond muscles stimulated, hypertonia, co-contraction, clonus, and rigidity are often seen in association with spasticity (1). Another word that can be used to describe the positive signs of the syndrome is muscle overactivity. The treatment team involved in the management of spasticity may find it one of the most challenging issues confronting them in the care of patients with neurologic disability. The goal of the treatment is to address this muscle overactivity with the fewest side effects and least exacerbation of the negative components of the UMNS.

**Epidemiology of Spasticity**

Clinicians treating neurological disorders are often confronted with the issue of spasticity and muscle overactivity. There is limited information regarding the prevalence and incidence of spasticity in various conditions. As expected, it depends on the etiology of the UMNS. In spinal cord injury (SCI), Maynard et al. (2) reported a prevalence of 65% to 78% and in 2004, Rizzo et al. (3) reported that 85% of patients with multiple sclerosis (MS) had some spasticity. This information may well be dated, as the introduction of newer agents to slow the progression of the illness may affect the incidence of spasticity that is seen in MS. Studies performed at varying time frames after stroke differ in the incidence of spasticity reported post stroke. A study assessing people 3 months after stroke revealed a prevalence of 19% and 35% in those with hemiplegia (4). On the other hand, a rate of 38% was reported at 1 year post stroke (5).

**Normal Motor Function and Control**

**Requirements for an Effective Motor System**

To facilitate the discussion of spasticity pathophysiology, it is useful to discuss normal motor control. To function effectively, the motor system must be able to integrate sensory feedback, control reflex activity, and coordinate volitional movement. It is critical for the controller to have information concerning the position of muscles and joints in addition to muscle velocity. It is also critical for the system to rapidly respond to external forces to control and respond to reflex activity and to initiate and stop motor activity. Feedback must exist among the many pathways that pass through the cortical, subcortical, brainstem, spinal cord, peripheral nerve, and muscle. The most distal unit involved in motor control is the motor unit, a part of the peripheral nervous system that is critical for control based on excitation and inhibition of muscle fibers.

**The Motor Unit**

The motor unit was first described by Sherrington (6). It is comprised of an α motor neuron and all of the muscle fibers that are innervated by it. Not all motor units are the same, as they differ in recruitment patterns and firing rates. This is a result of the different demands and purposes of individual motor units. The units also differ by the types of muscle fibers that comprise them. There are two major fiber types: type I and type II. Type I fibers are small, red, and oxidative, and they fatigue slowly. Motor unit function reflects the fibers that they contain. Motor units full of type I fibers are responsible for the baseline tonic muscle activity. Type II fibers, on the other hand, are large, white, and anaerobic. These muscles are more powerful and can deliver greater speed and velocity than type I; however, they can fatigue easily. These units are brought in to increase the force or speed of a contraction. Additional motor units are a hybrid of the two fiber types (7,8). When working properly, motor units fire with coordination of agonist and antagonist systems, normal patterns of recruitment and deactivation (9). Katz and Pierrot-Desilligny (10) reported...
the problems that can be created by co-contraction of muscles and muscles firing out of phase or at angles different from their normal areas of activity. This loss of the normal recruitment and recruitment pattern may play a key role in spasticity.

Normal Regulation of the Motor Unit
The motor control system uses a feedback loop, integrating information about muscle activity, position, and velocity. This enables the motor system to control and coordinate the stimulation of agonist and antagonist muscles around a joint. Critical information required by the control system includes muscle length, velocity, muscle tension, and joint position. This is mediated by a combination of immediate monosynaptic reflexes and more complicated higher-level control involving spinal and supraspinal polysynaptic activity. This activity can either increase or inhibit activity at the motor unit level. Units fire when the net excitation minus inhibition reaches threshold (8).

Muscle Spindle and Golgi Tendon Organs
The muscle spindle plays a critical role in the provision of necessary information for proper motor control. It is attached in parallel to the main muscle mass, and it contains afferent type Ia and II fibers that communicate information concerning position and rate of change of a muscle to the spinal column. The γ motor neuron is an integral component of the muscle spindle. During normal motor function, the γ motor unit coactivates with the α motor neuron and maintains the spindle tension and efficiency (8). Were it not for the γ motor neuron, the muscle spindle would be unable to provide accurate information throughout the muscle’s range of motion (ROM), and the spindle would be under less tension when the muscle contracts. This is analogous to a volleyball net being supported by two poles. The muscle contracting would be similar to the poles being brought closer together.

The Golgi tendon organs also play an important role in coordinating muscle activity. During normal movement the stretch reflex, which is mediated by the firing of the Ia afferent fibers, must be suppressed to allow full movement about the joint. This reflex is monosynaptic and mediated by the Ia afferents and is triggered when a sudden stretch is applied to a muscle, such as tapping on the knee. Some suppression of muscle activity is mediated through the Golgi tendon organs. Found within the muscle tendons, through the Ib fibers and their related interneurons, the Golgi tendon organs limit muscle contraction by facilitating antagonists and inhibiting agonists. Thus, they serve to impose a ceiling effect on muscle contraction and prevent musculotendinous injury (11).

Spinal Interneurons
The spinal interneurons play a critical role in normal motor control and spasticity. The effects of the Ia and Ib fibers mentioned earlier are often mediated through and with the help of interneurons called Ia and Ib interneurons, respectively. Other interneurons, including the Renshaw cell and the propriospinal interneurons, are also an important part of the control process.

As mentioned earlier, the Golgi tendon organs generate a ceiling effect for maximum muscle tension. The Ib afferents from these organs connect to their respective Ib interneurons. These interneurons also receive suprathreshold propriospinal influences from above that facilitate antagonists and inhibit the firing of agonist muscles (8,10). The type Ia interneurons receive activation from the type Ia neurons from the muscle spindle. When activated, the Ia interneurons facilitate agonist activity and reciprocally inhibit antagonist muscles, preventing the futility of co-contraction. Ia interneurons are also under supraspinal influence, and this plays a critical role in strengthening reciprocal inhibition by the type Ia interneuron. The loss of supraspinal influence on the Ia interneurons plays a critical role in co-contraction and cerebral origin spasticity (12).

The process of recurrent inhibition involves the Renshaw cell, which receives input directly from the α motor neuron. This process shuts off agonist activity by its direct effect on the α motor neuron, in addition to facilitation of antagonist function mediated via the antagonist muscle’s Ia interneuron (13). Tight motor control requires the function of the Renshaw circuit, and a loss of its function may greatly compromise movements (10). Like many other neurons, spinal and supraspinal input influence Renshaw cell function. Renshaw cell inhibition is increased in SCI (14).

Supraspinal Influences
Supraspinal influences play a major role in both volitional movement and the pathophysiology of spasticity. Rothwell et al. (15) demonstrated corticomotorneuronal pathways. Originating from the primary motor cortex, the corticomotorneuronal pathways are specific for specialized coordinated functions such as truncal balance or initiation of fine coordinated hand movements. Loss of these fibers leads to a functional deficit more than to spasticity. Reducing tone from the hands may improve passive and gross hand function but may not greatly improve fine movement (16).

The corticospinal tract is the major motor tract and originates from many areas within the brain. These include the extrapyramidal cells from the prefrontal region, the supplementary motor region, the cingular gyrus, and the postcentral gyrus of the parietal lobe. The pontine medial reticulospinal and lateral vestibulospinal tracts are the main extensor pathways within the brain. The pontine system facilitates the α and γ motor neurons of the extensors of the limb muscles with some input into the system from the sensorimotor cortex. The lateral vestibular tract is located in the ventromedial portion of the cord and terminates at the spinal cord motor neurons. Stimulating the lateral vestibular tract affects the motor neurons of the flexor muscles differently from the extensors, with the α and γ motor neurons of the flexors inhibited and those of the extensors facilitated. The nucleus of the cerebellum also has an excitatory influence on extensor pathways (17).

Several pathways facilitate flexion. The medullary lateral reticular formation (MLRF) inhibits extensor pathways. The cortex facilitates its action, and cortical injury can lead to
net overactivity of the lower-extremity extensor system. The MLRF demonstrates its effect through its connections to the motor neurons, type Ia interneurons, and type I bf system. In cats, the corticospinal, corticoreticulothalamic, and corticorubrospinal, all show significant flexor facilitation. Through interneurons, the corticorubrospinal tract excites flexor motor neurons and inhibits extensors. In addition, the medullary reticulospinal tract is a predominant part of a largely flexor-oriented system (18).

PATHOPHYSIOLOGY OF SPASTICITY

Origin of Spasticity
So from where does spasticity come? Dietz and Berger (19) have suggested that intrinsic properties of the muscle itself could explain the changes seen with spasticity. However, Nielsen et al. (20) feel that the term spasticity should only be used when the disorder is a result of “alteration of central processing of sensory input and to exclude structural changes in the muscles.” Based on work from animal models, the concept of “gamma rigidity” was raised. This concept states that through overfiring of the γ motor neuron, the spindle would be too taut and the Ia interneurons would be hyperexcitable. Efforts at identifying this with microneurography failed to confirm this hypothesis (21). Delwaide (22) felt that spasticity resulted from a loss of descending, facilitatory, inhibitory influences that act on Ia interneuron inhibition. In other words, reciprocal inhibition mediated through the Ia interneuron inhibition requires facilitation from higher centers. With injury to the central nervous system (CNS), the interneuron is unable to shut off antagonist muscles firing, with resultant increased velocity-dependent resistance to movement. The concept of a hyperexcitable motor neuronal pool has also been recently raised. In essence, these neurons would be hypervigilant and initiation of firing would occur with less excitation. This may result from a loss of tonic inhibition secondary to a loss of supraspinal influences. Some have expressed the belief that the ionic properties of the membrane itself are changed as well. Other theories that may explain spasticity include central collateral sprouting (23), presynaptic disinhibition (22), and denervation hypersensitivity (24). Neurotransmitters may also play some role in spasticity; some suspects include serotonin and substance P. In animal literature, serotonin has been noted to prolong responses and facilitate extensor responses (25). In general, no singular theory has been able to fully explain the pathophysiology of spasticity.

Other Muscle Overactivities Seen in the UMNS
There are numerous types of muscle overactivity that result from the UMNS that require treatment. As Sheean (26) stated in his excellent review article, the different components are commonly lumped under the term spasticity. As a result, the word is often used generically for muscle over activity from the UMNS. It is important to differentiate them.

In his article, Sheean stressed the importance of understanding the different components of the UMNS and how it may result in different treatment approaches and results. Sheean divides the different components of the UMNS into three distinct categories based on their underlying etiology: (a) spinal reflexes (stretch, nociceptive and cutaneous), (b) efferent drives that are not totally dependent on afferent feedback from the periphery, and (c) disorders of voluntary muscle movement.

As a result of damage to the CNS, there is abnormal processing of spinal reflexes which results in many of the positive components of the UMNS. These reflexes are dependent on afferent feedback from the periphery including muscle stretch, pain or stimulation of the skin. The increase in stretch reflexes falls into this first category and includes tonic (baseline tone) and phasic (from sudden stretch e.g., deep tendon reflexes [DTRs] and spasticity) responses. Other related components include clonus, flexor reflex afferents that result from noxious stimulation and cutaneous reflexes such as the Babinski.

In the second category, Sheean included spastic dystonia and associated reactions. Spastic dystonia is a term that was created by Denny-Brown (27) to describe his observation that some of the spinal cord injured cats held their limbs in a flexed position. The position of the cats’ limbs was not a result of a stretch and therefore was not a result of a reflex activity. This was even more strongly demonstrated when the posture remained after the dorsal root was transected. The only explanation that he could find to account for the observation is a supraspinal efferent drive to the motor neuron. The commonly seen clinical scenario that corresponds to Denny-Brown’s cat model is the stroke patient whose arm rests in a flexed position. Associated reactions are another example of a positive UMNS that results from efferent supraspinal stimulation to the motor neuron. This is often seen when a person’s arm gets progressively higher when he or she tries to push up from a chair or walk a distance. It appears to be related to the amount of effort that the person is exerting and/or the severity of the tone in the extremity (26). One theory regarding the etiology of this phenomenon is an inability to stop the spread of muscle activity. Treating this problem may prevent a person from being thrown off balance, thus facilitating transfers and/or mobility. An important clinical point is that this disorder could be resistant to a treatment that reduces sensory afferent input such as a dorsal root rhizotomy (DRT).

The third category is disorders of voluntary movement. A disturbance in reciprocal inhibition fits this category, as the prevention of pathological reciprocal inhibition is an important component of normal motor control. In the normal state, the type Ia interneuron receives input from the spindle-based Ia neuron and with facilitation from above is supportive of agonist activity and inhibits antagonist firing (12). Two clinical examples of this normal activity is the relaxation of the elbow flexors during attempted elbow extension and the relaxation of the plantar flexors during attempted dorsal flexion.
anterior activation. By reducing these muscle overactivities, there is a reasonable likelihood of increasing volitional firing of the antagonist muscle that may improve volitional function. The importance of recognizing these conditions cannot be overemphasized.

**MANAGEMENT OF SPASTICITY AND THE UMNS**

An integrated team is required to successfully manage spasticity. Under the direction of a physician skilled in the management of spasticity, the team should be able to deliver the entire continuum of services. The remainder of this chapter addresses the indications and benefits of the different treatment modalities (Table 50-1). How do clinicians decide what treatment to offer? Often the question is answered by the skills of the treating clinician, in that a man with a hammer sees everything as a nail. A physician skilled in chemoneurolysis is far more likely to perform a nerve block than one less comfortable with the procedure. Physicians may be more inclined to prescribe an oral pharmacologic agent than to prescribe serial casting or splinting. Therefore, it is important for the treatment team to communicate effectively and access needed surgical consultations as appropriate. Of note, not all spasticity is dysfunctional. Some women with spasticity use their elbow flexor tone to hold their pocketbooks, whereas lower-extremity tone may assist in transfers, standing, and ambulation. Optimizing function should be the primary outcome parameter in the treatment of spasticity. To maximize quality of life improvements for patients undergoing spasticity management, the entire treatment team must monitor the efficacy and adverse effects of spasticity interventions and adjust treatment accordingly (Table 50-2).

---

**TABLE 50.1** Sample Passive and Active Goals

<table>
<thead>
<tr>
<th>Passive</th>
<th>Active</th>
</tr>
</thead>
<tbody>
<tr>
<td>Improved hygiene</td>
<td>Improvement of transfers</td>
</tr>
<tr>
<td>Ease of care</td>
<td>Improvement in ADLs</td>
</tr>
<tr>
<td>Positioning</td>
<td>Improved mobility</td>
</tr>
<tr>
<td>Facilitate casting or splinting</td>
<td>Decrease spasms</td>
</tr>
<tr>
<td>Reduce pain</td>
<td>Release inhibition of antagonists</td>
</tr>
<tr>
<td>Improvement of orthotic fit</td>
<td>Reduce contraction during voluntary movement</td>
</tr>
<tr>
<td>Reducing difficulty of donning</td>
<td>Improvement of balance by blocking associated reaction</td>
</tr>
<tr>
<td>orthotic</td>
<td>Healing of decubitus ulcer</td>
</tr>
<tr>
<td>Healing of decubitus ulcer</td>
<td>Improvement in foot position during stance phase of gait</td>
</tr>
<tr>
<td>Ease of dressing</td>
<td>Easier straight cathing</td>
</tr>
</tbody>
</table>

---

**TABLE 50.2** Commonly Used Oral Medications

<table>
<thead>
<tr>
<th>Medication</th>
<th>Daily Dosage (Range)</th>
<th>Mechanism of Action</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baclofen</td>
<td>10–300 mg</td>
<td>Presynaptic inhibition of GABA&lt;sub&gt;2&lt;/sub&gt; receptors. Is active both presynaptically and postsynaptically. Hyperpolarizes cell membrane</td>
<td>Risk of withdrawal seizures and hallucinations. Dose must be adjusted with renal disease</td>
</tr>
<tr>
<td>Diazepam</td>
<td>4–60 mg</td>
<td>Facilitates postsynaptic effects of GABA&lt;sub&gt;2&lt;/sub&gt; by opening chloride channels in membranes resulting in increased presynaptic inhibition secondary to hyperpolarization</td>
<td>Oldest class of medications used for spasticity that is still in common use. Can have very long half-life</td>
</tr>
<tr>
<td>Dantrolene</td>
<td>25–400 mg</td>
<td>Interferes with calcium release from sarcoplasmic reticulum</td>
<td>Only truly peripherally acting oral agent. LFTs must be watched carefully</td>
</tr>
<tr>
<td>Clonidine</td>
<td>Oral 0.05 mg bid—0.4 mg/d or 1–6 patch/wk</td>
<td>α&lt;sub&gt;2&lt;/sub&gt; Agonist. Decreases tonic facilitation via locus coeruleus and in spinal cord enhances presynaptic inhibition</td>
<td>Primary use in SCI population. Theoretical limitation to use in ABI secondary to interference with recovery</td>
</tr>
<tr>
<td>Tizanidine</td>
<td>1–36 mg</td>
<td>α&lt;sub&gt;2&lt;/sub&gt; Agonist. Blocks release of excitatory neurotransmitters and facilitates inhibitory neurotransmitters. Antinociceptive and reduces spinal reflexes</td>
<td>Now available in 2 and 4 mg tablets. Slow titration reduces sedation side effect that is major limiting factor</td>
</tr>
</tbody>
</table>
Functional Goals Should be the Target for Treatments

Spasticity and the UMNS are a result overall of an irreversible process within the CNS. Treatment decisions should be made based on the functional limitations imposed by spasticity and the UMNS, neither of which represents an independent disease state. Specific impairments or functional deficits such as pain, problems with position, hygiene, or mobility are the specific issues that therapy should address. Gans and Glenn (1) divided treatment goals into two categories. The first is the management of passive function, such as reduction of pain, positioning, hygiene, splint wearing, and prevention of contracture. The second is related to active functional activities; as they described it, “Diminished capacity of the patient to accomplish useful work with the motor system.” The goals of these treatment interventions are to improve volitional purposeful movement. Some examples include unmasking functional movement that is inhibited by antagonist spasticity, improving transfers, ambulation, and performance of activities of daily living (ADLs).

Treatment Options

Reduction of Noxious Stimulation

The first step in any program to manage spasticity is the reduction of noxious stimulation. Spasticity and muscle overactivity have been shown to be increased as a result of this input (28). Stimulation of the flexor reflex afferents may lead to an increase in pathologic activity (29). The term noxious stimulation encompasses a wide variety of conditions such as a pressure ulcer, ingrown toenail, contracture, kinked catheter, urolithiasis, urinary tract infection, deep venous thrombosis (DVT), heterotopic ossification, fecal impaction, sepsis, and fracture. This is just a partial list. Addressing these conditions should generally be the first approach in spasticity management.

Positioning

Proper positioning is an important component of spasticity management. Poor positioning can result in an increase in spasticity and in decreased ROM, contractures, increased noxious stimulation, pain, and exacerbation of a vicious cycle that can lead to worsening spasticity (30). This is especially true in the ICU and acute hospital (28). Proper goals for a positioning program include improvement in body alignment and greater symmetry. Benefits include easing of nursing care, facilitation of therapy, and maximization of a patient’s function. Postures that should be avoided include a leg scissoring posture (bilateral hip extension, adduction, internal rotation), windswept position (hip flexion, abduction, external rotation on one side and relative hip extension, adduction, and internal rotation on the other), and frog-leg position, which can exacerbate the spasticity. Positioning is also important in the wheelchair. Tone can be minimized by placing the patient with the hips and knees at 90 degrees and by maintaining good torso position (30). While its full implication for treatment has not yet been clarified, the joint angle position during assessment of spasticity contributes to the measurement of both tone and reflexive action at the ankle and elbow joints (31,32).

Stretch and Casting

The UMNS causes muscle shortening for several reasons. One reason is the immobilization of paralyzed muscles in shortened positions. This resultant decrease in longitudinal tension (muscle unloading) can predispose to contracture. Immobilization can also cause a reduction in protein synthesis in immobilized muscles, thus promoting atrophy (33). Spasticity and muscle overactivity also play a part in muscle shortening (34). This can in turn result in an increase in spindle activity and sensitivity (35). Gracies et al. (28) have discussed the need to promote the commencement of stretch early in the treatment of any neurologic condition. Stretch has the advantage of being a focal treatment that can combat the development of the previously mentioned muscle shortening and increase in spindle sensitivity (33).

Schmit et al. (36) have demonstrated the benefit of a relatively brief stretch in the management of spasticity. However, the benefit is short-lived, as the tone returns after a single contraction (37). Therefore, stretch needs to be applied for a longer period of time to have potential functional benefit. A study involving the use of a Lycra garment that provided a stretch of 3 hours demonstrated both an improvement in spasticity and good patient tolerance (38). Stretch has been shown to be useful in volitional movement in both agonist (39) and antagonist muscles (40). Finally, in a nonblinded convenience sample, Selles et al. (41) reported that a 4-week program utilizing a device to deliver “intelligent stretch” was effective in treating spasticity and contracture in stroke patients as measured by joint properties, gait speed, and subjective reporting. In summary, stretching activities have the advantage of being a local treatment, with limited risk that has demonstrated proven effect in the management of spasticity.

Chronic stretch via casting or splints changes reflexive activity and reduces the stretch reflex (42–44). Serial casting is defined as the use of successive casts to treat increased tone and contractures. Some of the original work regarding the effectiveness of serial casting in the management of spasticity and contracture comes from the group at Rancho Los Amigos (45), which reported their treatment of the lower extremities in 42 patients with a brain injury. It is not definitely known how often to change serial casts. Pohl’s group (46) performed a retrospective chart review investigating the differences between changing casts every 1 to 4 days versus 5 to 7 days. They found that at 1 month the changes in ROM between the two groups were comparable, but the complication rates were very different. The group with casts changed every 5 to 7 days had a complication rate of over 29% while the group with more frequent changes only had a rate approaching 9%. Almost 13% of the casting efforts had to be discontinued secondary to complications. A review of the literature discussing upper-extremity casting was unable to demonstrate a definitive answer supporting or rejecting casting as a treatment. Using casting to facilitate the action of chemical denervation has been addressed in several studies. Some preliminary studies have shown that using casting in combination with botulinum toxin treatment has increased the treatment effect in the stroke
(47) and cerebral palsy (CP) populations (48). The duration of the effect of serial and bivalve casting has been addressed in several studies. One study demonstrated that the effect was maintained 1 month after its use was discontinued (46). Its use in combination with botulinum toxin has been studied. In summary, casting can be considered a beneficial treatment option in the management of spasticity and contractures and is commonly used in programs with experienced clinicians; definitive recommendations regarding its use cannot be made at this time.

Physical Modalities

Physical modalities can play a role in the management of spasticity. Like stretching, they have the benefit of being benign interventions with localized treatment benefits. The use of these agents will most likely remain as part of a spasticity treatment program. Used correctly, the physical modalities can have an important role in spasticity management.

Cooling of muscles is beneficial in the management of spasticity (49,50). It both inhibits the monosynaptic stretch reflexes and lowers receptor sensitivity after it is removed (51,52). Cooling can be used in different ways. The quick icing technique, with ice applied with a light striking movement, results in facilitation of α and γ motor neurons and is used to facilitate antagonist function (53), whereas prolonged cooling can result in decreased conduction velocity and a reduction in the maximal motor complex motor action potential (49,54–56). The issue of cooling and muscle elasticity was addressed in a study that found a 3% to 10% decrease in elastic stiffness after a 30-minute ice cooling over the calf muscles (56). However, the effect lasted less than 1 hour. Another method of cooling delivery includes the use of an evaporating spray, such as ethyl chloride (57). With their short duration of action, the cooling modalities may have their greatest utility by therapist to reduce muscle overactivity to allow other therapeutic interventions.

Heat is another modality that can be applied in various forms. Ultrasound, paraffin, fluidotherapy, superficial heat, and whirlpool are some of the most common ways heat is applied. Heat’s effect has a short duration (58), and, like cold, its application should be followed immediately by stretching and exercise. The effects of heat on spasticity have been studied in only a limited way. Its major effect seems to be related to an increase in elasticity that may assist in stretching activities (59).

Deeper heating modalities have also been used in the management of spasticity. Wessling et al. (60) demonstrated that 1.5 W/cm² in combination with stretch resulted in a 20% greater distensibility than stretch alone.

Electrical Stimulation

Electrical stimulation is another modality that can help spasticity management. Transcutaneous electrical nerve stimulation (TENS) units have been shown to be useful in the management of pain. Through its nociceptive action and resultant reduction in pain, it was thought to reduce spasticity. Specifically, by reducing the flexor reflex afferents that are facilitated by nociceptive stimulation (29), Bajd et al. (61) demonstrated a reduction in SCI-related spasticity in three of six patients in a dermatomal pattern, while a group applying TENS in an acupuncture method demonstrated a substantial reduction in spasticity that was partially reversed by coadministration of naloxone (62). Other potential mechanisms of action for spasticity reduction include inhibition or fatiguing of spastic muscles and possible activation of antagonist muscles through the la interneurons (63).

Numerous electrical stimulation devices are now being used to facilitate motor recovery secondary to CNS pathology. While the literature concerning these devices deals primarily with recovery, some information regarding their effect on spasticity has been reported. A single-blinded study performed in 2005 evaluated the effect of functional electrical stimulation (FES) on acute stroke patients. This study reported that 15 sessions of 30 minutes of FES administered in a reciprocal fashion to imitate normal gait resulted in improved ambulation and spasticity as measured by the Composite Spasticity Scale (64). Chen et al. (65) tested the effect of surface stimulation on the gastrocnemius of chronic stroke patients showing electrophysiological changes and a trend to a reduced ankle Ashworth score. Aydin’s group (66) provided evidence that TENS may be of value as an adjunct for spasticity management in the stroke population. A recent study (67) on stroke patients demonstrated that ankle plantar flexor tone could be reduced with a combination of Bobath techniques and 9 minutes of stimulation to the dorsal flexor muscles. Some limited results have also been demonstrated with the use of electrical stimulation in people with CP (68,69). Finally, Krause et al. (70) demonstrated that an FES bike was beneficial in the management of spasticity secondary to MS.

Massage

Massage is a therapy that is often desired by patients and their families. However, a review of the literature does not reveal any strong scientific evidence that supports its efficacy and utility (53,71,72).

Pharmacologic Treatments

Four common methods are currently being used in the delivery of pharmacologic agents. The oldest method is delivery though the enteral system, either by mouth or via gastrostomy tube. Agents such as baclofen, benzodiazepines, or tizandine are delivered in this fashion. These agents undergo systemic absorption and demonstrate an effect throughout the entire body. A second method, which is closely related to enteral delivery, is the use of a transdermal system. An example of this is Catapress TTS. Medications administered in this fashion are also absorbed systemically and demonstrate their effects throughout the body. Transdermal systems differ from the enteral methods by having a more steady-state blood drug level throughout the day with less fluctuation. Intrathecal administration of active agents is a third method of drug delivery. By placing the medications closer to their site of action, systemic side effects are reduced and clinical efficacy is obtained.
with lower total doses. Baclofen, morphine, and clonidine are some of the more common medications that are delivered in this manner. Local injection of chemodenervation agents is the fourth method of drug delivery. Agents such as phenol and ethanol classically and now in the last decade the botulinum toxin products fall in this category. This last mode of administration is the best choice for treating a focal issue with a minimum of systemic effects, though some systemic absorption is still detectable.

Specific Components of the Pharmacologic Decision Process

Many factors contribute to the decision of which pharmacologic agent to use in a particular patient. A partial list includes etiology, time since onset of medical issue, prognosis, access to medical services, personal support system, concurrent medical problems, cognitive status, and financial resources. All these items are important. A brilliantly planned intervention that the patient cannot afford is of no value. Similarly, an aggressive outpatient therapy program is of no value if transportation cannot be obtained for the patient.

Etiology

Spasticity may present similarly despite originating from different etiologies. Nonetheless, responsiveness to various treatment interventions may vary depending on the etiology. As an example, enteral pharmacologic agents have been shown to be of great efficacy in the management of spasticity resulting from SCI or MS (73–77), whereas the benefit in spasticity caused by traumatic brain injury (TBI) or stroke is far less apparent. A further complicating factor in cerebral origin spasticity is the potential for impairment of recovery secondary to treatment (78), or a potentially intolerable cognitive side effect profile even when the agent may be effective (79).

Time Since Onset

As a general rule, more aggressive spasticity interventions are tried later in the course of an event. Medications that may impair recovery are less likely to be used early on. For patients who have low-level functioning post TBI, physicians are less likely to prescribe sedating antispasticity agents. Phenol neurolysis is rarely used early in recovery, as the scarring of muscle and nerve and long duration of action may be undesirable in a recovering patient. Orthopedic interventions are almost never offered early on, as there needs to be stabilization of the neuromuscular structures before permanent surgical intervention. Relative to intrathecal baclofen (ITB), there is now some controversy as to what is considered too early. A recent report from France demonstrated that ITB may be beneficial for recalcitrant spasticity when initiated in the first month post injury (80).

Functional Prognosis

When a patient’s prognosis for motor and functional recovery is very guarded, this may lead clinicians to attempt more aggressive, permanent interventions such as a rhizotomy. The need to facilitate care delivery with aggressive means of reducing spasticity may be preeminent over efforts to promote an unlikely recovery.

Support System

The availability of social support may be an important factor in the management of spasticity. A number of questions should be considered. Can medication administration be supervised in the cognitively impaired patient? A family member or other caregiver may be critical for the physician to safely prescribe medications. Are supervision and assistance available for transportation to therapy or assistance in safe utilization of splinting devices? Will the patient be able to follow up for ITB pump refills or will there be the risk of withdrawal if the patient fails to follow up for pump refills?

Cognitive Status

It is important to assess a patient’s cognitive ability when prescribing treatment. The clinician must address the patient’s ability to be compliant and remain safe while using a treatment modality. Will the patient be noncompliant with a medication and risk withdrawal seizures? Will there be a risk of skin breakdown with the use of a splint or serial cast?

Concurrent Medical Problems

The overall medical condition of the patient being treated must be considered. Patients with hypotension, syncope, balance disturbances, or ataxia may be unable to tolerate the side effect profile of certain agents. Would an oral agent cause hypotension and resultant syncope, exacerbate ataxia, coordination, or balance disturbance? Does the patient have chronic infections that would increase the risk of development of an infection with an indwelling catheter or ITB system?

Distribution of Spasticity

How diffuse is the area that needs treatment? Is there a focal or segmental area that needs treatment or is spasticity distributed diffusely throughout the entire body? If there are only discrete regions or only the discrete regions need to be treated even in the midst of diffuse spasticity, chemodenervation may be most appropriate. If the target condition is more systemic, then treatment that is more global will be necessary.

Financial Issues

ITB and botulinum toxin injections cost thousands of dollars. Paying out of pocket for some modalities is not realistic and the physician, patient, and family have to utilize third-party payers. Some insurance companies are requiring trials with less expensive agents such as oral antispasticity agents before approving toxin injections. Clinicians are often required to justify their decisions and recommendations.

Oral and Transdermal Medications

Oral and transdermal medications are commonly used in the treatment of spasticity. Table 50-3 summarizes the usage of these medications.
PART IV | SECONDARY CONDITIONS AND COMPLICATIONS

### Benzodiazepines in SCI

The greatest benefit for diazepam has been demonstrated in the SCI population. A double-blind crossover study with 22 patients with SCI-related spasticity demonstrated efficacy (86), whereas another study with 21 patients with MS or SCI showed that diazepam is superior to placebo in treating spasticity. Whether the benzodiazepines are better for complete or incomplete lesions is debatable. Whyte and Robinson (87) suggest that benzodiazepines are effective only for incomplete lesions, but this is still controversial (88,89). A survey performed at Veterans Administration SCI programs showed that 70% of prescribers routinely give benzodiazepines to their patients (90).

### Benzodiazepines in MS

Studies performed in patients with MS have compared the efficacy and side effect profiles of the benzodiazepines versus baclofen. The two agents had very comparable efficacies and tolerances. Sedation was found more often with the benzodiazepines, whereas the baclofen group had more side effects (81–83). A survey performed at Veterans Administration SCI programs showed that 70% of prescribers routinely give benzodiazepines to their patients (90).

### Benzodiazepines in Acquired Brain Injury (ABI)

Benzodiazepines are rarely used in the ABI population because of their potential for cognitive side effects as well as their potential to compromise motor recovery (78).

### Benzodiazepines in CP

Few studies are available concerning benzodiazepine use in the CP population. Engle (94) conducted a double-blind crossover study that demonstrated the efficacy of diazepam in the management of patients with CP. However, determining whether the improvements were behavioral in origin was questionable. Mathew and Mathew (95) demonstrated in a

---

**TABLE 50.3 Comparisons of Different Treatment Modalities for Spasticity**

<table>
<thead>
<tr>
<th>Spasticity Treatment</th>
<th>Indications</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Therapeutic modalities</td>
<td>Used by therapist for early management and facilitation of chemodenervation</td>
<td>Minimal side effects</td>
<td>Short duration of effect</td>
</tr>
<tr>
<td>Oral medications</td>
<td>Generalized tone, spasms, no focal region of spasticity</td>
<td>Systemic administration. Can treat large area of spasticity</td>
<td>Systemic side effects such as sedation, metabolic load</td>
</tr>
<tr>
<td>Botulinum toxins</td>
<td>Focal area of spasticity</td>
<td>Can treat spastic area without systemic side effects</td>
<td>Expensive, 3-mo duration when procedure needs to be repeated</td>
</tr>
<tr>
<td>Phenol</td>
<td>Focal area of spasticity</td>
<td>Can treat spastic area without systemic side effects. Much cheaper than Botulinum toxins and longer duration</td>
<td>Requires considerable skill of injector, risk of dysesthesias, painful procedure</td>
</tr>
<tr>
<td>Orthopedics procedure</td>
<td>Potential improvement in passive or active ADLs. Stable neurologically</td>
<td>Can be long-term repair</td>
<td>Surgical risk, loss of motor strength</td>
</tr>
<tr>
<td>ITB</td>
<td>Significant tone not adequately treated by other modalities</td>
<td>Baclofen gets to spinal cord with minimal systemic absorption</td>
<td>Surgical procedure, pump and catheter will eventually need replacement, high cost, risk of catheter dislodgement or kink</td>
</tr>
</tbody>
</table>

---

**Benzodiazepines**

The benzodiazepines were the first agents used in the management of spasticity. Of this class, diazepam (Valium) is most commonly used. Other agents include clorazepate (Tranxene) and clonazepam (Klonopin). Ketazolam (Loftran) is another benzodiazepine that has been trialed for spasticity and is available in Canada but not in the United States (81). The benzodiazepines’ mechanism of action is central in origin, acting on the brainstem reticular formation and spinal polysynaptic pathways (82). The benzodiazepines demonstrate their effect via GABA<sub>A</sub> (γ-aminobutyric acid), which opens membrane Cl<sup>−</sup> channels with resultant hyperpolarization. The net effect is a reduction of monosynaptic and polysynaptic reflexes and an increase in presynaptic inhibition (81).

Initial dosing for diazepam is 2 mg bid or 5 mg at bedtime, with a gradual titration upward to a maximum of 60 mg/day for adults. Benzodiazepines are well absorbed after enteral administration and peak at 1 hour. There is a relatively long half-life for benzodiazepines when accounting for its active metabolites, which ranges between 20 and 80 hours. The side effect profile can be quite problematic, including problems with addiction and withdrawal, ataxia, weakness, cognitive impairment, memory dysfunction, poor coordination, fatigue, and CNS depression that can be potentiated by alcohol. Research with diazepam has demonstrated improvements in painful spasms, hyperreflexia, and passive ROM. Evidence concerning functional improvement is limited.

Clorazepate has been shown to have a more favorable side effect profile than diazepam. In clinical trials, it was noted to have fewer problems with sedation and memory (83,84). Its half-life is relatively short but its active metabolite, desmethyl-diazepam, has a half-life of up to 70 hours. In obese patients, the half-life of clorazepate can extend beyond 200 hours (85). Doses of 5 mg bid have been used in clinical trials.
placebo-controlled study the efficacy of bedtime diazepam administration in improving ADLs and reducing the burden of child care in children with CP without significant adverse side effects. Nogen (96) studied diazepam and dantrolene in patients with CP and found benefit with both agents. Cruikshank and Eunson reported on the utility of intravenous diazepam in their management of planned discontinuation of ITB in three cases of individuals with CP (97).

**Baclofen**

Baclofen (Lioresal) is another agent that mediates its activity through the GABA system. It differs from the benzodiazepines by mediating its effect via GABAB rather than by GABA$_A$ and is active both presynaptically and postsynaptically. Its action presynaptically is to bind to the GABA interneuron, where it causes hyperpolarization of the membrane that prevents the influx of calcium and resultant release of neurotransmitter. When it binds postsynaptically, it hyperpolarizes the cell membrane by acting on the Ia afferents. As a result, baclofen is inhibitory on both the monosynaptic and polysynaptic reflex pathways. Baclofen is eliminated via the kidney, and its half-life is roughly 3.5 hours (81).

When initiating treatment, 5 mg bid to tid is recommended, and this can be increased to 10 mg/day/week. The Physician’s Desk Reference suggests a maximum dose of 80 mg/day, but while not routinely recommended, doses as high as 300 mg/day have been used safely (81). Baclofen-related side effects reported include sedation, fatigue, weakness, nausea, dizziness, paresthesias, hallucinations, and lowering of seizure threshold. The patient is at greatest risk when the agent is abruptly discontinued, as hallucinations and withdrawal seizures have been reported (87). Since baclofen primarily undergoes renal clearance, dosing may need to be adjusted with kidney-related issues (98). When switching from oral to either IV or ITB administration, one must be wary of potential withdrawal-related issues. This is based on the efficiency of localization that results from IV or intrathecal dosing, as there is a relatively low dose of baclofen in the brain as compared with the lumbar cord region.

**Baclofen in SCI and MS**

Much of the baclofen literature combines research on patients with SCI and MS; therefore it is appropriate to merge a discussion of baclofen in these populations, although some differences will be highlighted (73,98–101). Feldman et al. (100) reported that the use of baclofen in patients with MS demonstrated a significant reduction in spasticity as well as reducing painful flexor spasms. Flexor spasms in the SCI population also responded to baclofen administration (73,74,102–105). It is far more difficult to find studies that demonstrate functional improvement, given that investigations were unable to demonstrate improvements in ADLs and ambulation with administration of baclofen (102,104). Orsnes et al. (106) studied patients with MS treated with baclofen but again found no functional improvement. Nielsen et al. (107,108) studied the effect of oral baclofen on the soleus muscle. Treatment with baclofen reduced ankle stiffness and increased soleus response latency. However, it also was found to increase the weakness in soleus function, which may explain the lack of functional improvement.

**Baclofen in ABI**

There is limited literature that has noted positive effect with oral baclofen in the ABI (103,109,110). A double-blinded study in the elderly stroke population was discontinued because of treatment-related sedation (111).

**Baclofen in CP**

Milla and Jackson (112) conducted the one blinded crossover trial in the CP literature that demonstrated efficacy. Actual functional benefits were not seen, but decreased scissoring and improvements in ROM were noted. The authors reported few side effects and recommended dosing of a total of 5 to 10 mg total per day in divided doses for children 2 to 7 years of age. A recent publication looked at the effect of 4 weeks of oral baclofen on ten children with CP. The authors looked at the effect of medication on neuromuscular activation and ankle plantar flexor torque. Neuromuscular activation was measured using surface electromyography (EMG) generated during maximal voluntary contraction to the M-wave during supramaximal electrical stimulation of the tibial nerve. They noted a statistically significant improvement in neuromuscular activation but this was not accompanied with isometric plantar flexion torque. The authors hypothesize that the agent may be beneficial in the facilitation of strength training (113); however, further work is needed to address this question.

**Dantrolene Sodium**

Unlike many other agents, dantrolene (Dantrium) is an enteral medication that acts peripherally, at the level of the muscle itself. Its mechanism of action is to inhibit calcium release from the sarcoplasmic reticulum during muscle contraction. Rather than dampening neuronal activation, it reduces the strength of contraction. In addition to its action on the muscle extrafusal fibers, dantrolene reduces muscle spindle’s sensitivity by acting on the γ motor neuron (71). Dantrolene’s primary action is on fast-twitch fibers. Parameters affected by it include easier ROM and tone. Starting dose is 25 mg bid and can be increased weekly by 25 to 50 mg to a maximum of 400 mg/day (114). Dantrolene’s enteral half-life is approximately 15 hours and is given bid to qid (81). Dantrolene’s most well-known side effect is potential liver toxicity. However, overall this is a rare occurrence, with a rate of only 1.8% (115) when administered for more than 60 days. Even when discovered it is usually reversible. It is found most commonly in women over the age of 40, especially if they had been on high doses, greater than 300 mg for a long period of time (116). Fatal liver failure has been reported in 0.3% of those who received the medication. Therefore, it is critical for clinicians to follow liver function tests (LFT) when prescribing dantrolene. Blood tests should be performed weekly for the first month, monthly
for the remainder of the first year, and four times a year after that. In addition to liver toxicity, other problems associated with dantrolene include weakness, paresthesias, nausea, and diarrhea (81). Rare side effects also include anorexia, enuresis, visual disturbance, acne-like rash, inhibition of platelets, eosinophilic pleural effusion syndrome, and pericardial effusions (116,117). There has also been a single case report of a minimally conscious 43-year-old female with acontractility of the bladder (118).

**Dantrolene in SCI**

Since dantrolene is associated with weakness, there have been few trials reported with its use in SCI. An early study by Glass and Hannah (119) reported that it was more effective than diazepam in controlling spasticity but was associated with greater weakness. Studies have demonstrated improvements in ROM and tone but no functional improvement (120,121). Additionally, two cases are reported in the literature where SCI patients responded to dantrolene when treated for baclofen withdrawal (122,123). A recent 2006 Cochrane review failed to reveal any studies of dantrolene in SCI patients rigorous enough to meet the criteria (124).

**Dantrolene in MS**

There is limited literature describing the use of dantrolene in patients with MS. Again, this may well be due to a poor tolerance, given the undesirable side effect of additional muscle weakness. Two studies performed demonstrated improvement in tone and ROM. However, the benefits were outweighed by the clinical weakness found while on the medication (125,126). Its use in patients with MS cannot be recommended from the literature.

**Dantrolene in ABI**

Whyte and Robinson (87) have recommended dantrolene for use in the treatment of ABI-related spasticity. Chyatte et al. (127) reported on nine patients with stroke-related spasticity. Although they demonstrated no functional improvements in ADLs and mobility, they noted improved ROM, DTRs, and some upper-extremity function. Ketel and Kolb (128) reported that in their selected population of dantrolene responders, there was an exacerbation of spasticity accompanied by clinical deterioration when they were placed on placebo.

**Dantrolene in CP**

While reviewing the data concerning the use of dantrolene in patients with CP, Krach (129) reported on four studies that demonstrated good efficacy. Haslam et al. (130) reported decreases in DTRs and scissoring. Dosing in the pediatric population has been up to 12 mg/kg. Verrotti et al. (131) suggested that it could be efficacious when used in combination with diazepam and that it may be more effective than baclofen, with younger children demonstrating better tone reduction and older children showing improved motor movement.

**Clonidine**

Clonidine (Catapress) is an imidazoline derivative that is primarily an antihypertensive agent. It is a central acting \(\alpha_2\)-adrenergic agonist that has been shown to demonstrate some efficacy in spasticity management, primarily in SCI. It peaks in 3 to 5 hours when taken orally and has a usual half-life of 5 to 19 hours, with an extended half-life of up to 40 hours in persons with renal impairment. Clonidine’s clearance is primarily renal, with half of it first metabolized by the liver. Clonidine has two distinct mechanisms of action. First, it acts directly on the locus coeruleus and decreases tonic facilitation (81). Second, it also has a spinal mechanism, acting to enhance \(\alpha_2\)-mediated presynaptic inhibition (132–134). Clonidine doses as low as 0.1 mg orally are often effective in treating spasticity (133). A transdermal system that allows for more uniform blood levels and easier administration is also available. Side effects reported with clonidine include bradycardia, depression, lethargy, syncope, and hypotension (133,135). There is no literature to date describing clonidine use in patients with MS.

**Clonidine in SCI**

Two separate reports demonstrate clonidine’s potential efficacy in the SCI population (136,137). Nance et al. (134) showed that resistance to stretch was reduced in her small series of four patients treated with 0.2 mg/day. Donovan et al. (132) reported on the use of clonidine as an adjunct to baclofen in spasticity management. They reported an overall 56% response rate, with improvements noted in persons with paraplegia and tetraplegia with complete and incomplete lesions. The authors reported that three patients who had been responsive had to be discontinued secondary to postural hypotension.

**Clonidine in CP**

Lubsch et al. (138) published a retrospective chart review of 87 children with the diagnosis of spasticity secondary to TBI or CP who were on either baclofen or clonidine. Eighty-six persons were taking baclofen and only 31 were taken clonidine. Eighty percent of the population studied included in this review had the diagnosis of CP, with the remaining having the diagnosis of TBI. The two populations were not analyzed separately in the paper which further limits its utility. The paper did report on the safety and tolerability of clonidine administration to children with CP in doses ranging from 0.025 to 3.6 mg/day.

**Clonidine in ABI**

Clonidine can impair motor recovery after ABI (78), thus making its use in these patients somewhat controversial. Several publications address clonidine use in ABI including a case report concerning brainstem origin spasticity (139) and a six-patient case series (140). The Lubsch study (138) mentioned above some very limited information on the use of clonidine in ABI. However, only 20% of the population was TBI and thus, the information for ABI is even more limited than for CP. Further work needs to be done to clarify clonidine’s place as an antispasticity agent in this population before any evidence-based recommendations can be made.
Tizanidine

Like clonidine, tizanidine (Zanaflex) is an imidazoline derivative with α2-agonist effects and is the newest, widely used antispasticity agent. Its onset is rapid, 1 hour, and it has a very short 2.5-hour half-life and may require frequent dosing. It is cleared via liver metabolism and then is excreted by the kidneys. One mechanism of action is mediated through its effect on neurotransmitters. Tizanidine blocks the release of the excitatory amino acids, glutamate, and aspartate, in addition to facilitating the inhibitory neurotransmitter glycine (81). Animal research reveals other potential mechanisms of action that include antinoceptive activity (141–143), in addition to an ability to reduce spinal cord reflex activity (141,144).

The side effects of tizanidine can be quite troubling, with close to 15% of participants discontinuing the medication during clinical trials due to side effects (145). Drowsiness is reported in up to 50% of patients on clinical trials (145–147). The work of Meythaler et al. (79) may be the most telling, as patients who were obtaining benefit from tizanidine's anti-spasticity effect were lowering the dose to limit the side effects, despite a lower clinical efficacy. Another major side effect is dry mouth (145–147), with up to 11% of people complaining of it. Other complaints included fatigue, dizziness and hypotension (148), muscle weakness, nausea, and vomiting. There is also a potential for liver damage, and LFT should be evaluated before medication initiation and at 1, 3, and 6 months after initiating treatment. Initial dosing begins at 2 to 4 mg before bedtime and can be increased to a maximum of 36 mg/day (81).

**Tizanidine in SCI**

Tizanidine has been well studied in patients with SCI (75,149), although a Cochrane review published in 2006 (124) was only able to demonstrate one trial that met its strict criteria. The trial demonstrated a significant reduction in spasticity and an early reduction in spasm frequency. No change in muscle strength was observed. There were also significantly more side effects reported in the tizanidine group than the placebo group (124). Other important points from the Cochrane review are that there was no evidence of benefit from any of the other oral agents that they reviewed. They also reported no evidence of functional benefit from tizanidine use, but that the side effects were tolerable. Overall, it has been shown to be effective with an acceptable side effect profile with no weakness noted. Tizanidine is considered safe and effective in this population.

**Tizanidine in MS**

Several studies have demonstrated the efficacy of tizanidine in patients with MS (145,147,150,151). In work from the U.K. Tizanidine Trial Group (152), an improvement in clonus and spasms was seen but not in tone or function. The work by Lapierre et al. (147) demonstrated improvement in clonus and DTRs, but again no functional changes could be detected. Two studies compared tizanidine to baclofen and revealed that tizanidine was as effective as baclofen and better tolerated (145,151).

**Tizanidine in ABI**

Despite being a relatively new agent, tizanidine has undergone its most rigorous testing in ABI and may turn out to be the most effective agent in these patients. Meythaler et al. (79) reported on a series of 17 patients treated with an escalating dose of tizanidine. When they had reached their maximal dose, the patients demonstrated statistically significant decreases in their upper- and lower-extremity Ashworth scores and lower-extremity spasm scores. The changes were not significant in upper-extremity spasms or any of the reflex scores. The side effect profile, especially sedation, was problematic. The original plan was to reach a dose of 36 mg by 4 weeks; however, the maximal dose reached was a mean of 25.2 mg/day. This was lowered by an average of 2 mg later in the trial when the patients were able to self-select dosing. Once again there was no evidence of improved function, with the FIM and the Craig Handicap Assessment and Reporting Technique (CHART) showing no change with treatment. Bes et al. (146) performed a study that compared tizanidine to diazepam in the management of patients with chronic spastic hemiplegia. They reached their maximal level for each medication by week 6, and they underwent evaluation at week 8. Improvements in nonfunctional measures such as duration of stretch and muscle contraction as well as reduction of clonus were noted with both agents. In regards to functional measures, the participants demonstrated statistically significant improvement only while on tizanidine, as noted in walking distance on flat ground. They also reported that tizanidine was better tolerated than diazepam. Two recent reviews (153,154) suggest the use of tizanidine in the management of ABI-related spasticity. However, the clinical utility is still somewhat unclear as demonstrated by Meythaler et al. (79). This belief is further supported by the conclusions of a double-blinded placebo-controlled clinical trial comparing tizanidine to placebo and botulinum toxin. The results did not support the use of tizanidine for the management of upper-extremity spasticity secondary to ABI. The results demonstrated a clear superiority for the toxin treatment, while the tizanidine group was not shown to be superior to placebo (155). A final concern that needs to be highlighted is that tizanidine, like clonidine, is a member of the imidazoline family. Therefore, it also has the potential to impair neurological recovery after ABI (78).

**Tizanidine in CP**

There are no clinical trials published in the English literature that have tested the use of tizanidine in the management of patients with CP. A review article reported that it can be safely used in this population if the tablets are divided (129). The fact that the drug is available in scored 2 and 4 mg tablets is especially beneficial for dosing in this population. A study from Russia (156) reported that a dose of 1 mg three times a day for children under 10 and 2 mg three times a day for older children had significant benefit. In a double-blinded placebo-controlled study, 0.05 mg/kg/day of tizanidine was compared to placebo over a 6-month time period. Statistically significant
differences between placebo and active drug were noted in regards to reflexes and tone (157). Given the limited amount of data for this patient population, further research is clearly needed.

Other Medications
Since spasticity is a significant and recalcitrant problem, the list of medications that have been tried in its treatment is endless. It would be impossible to list them all. Instead the authors will mention a few medications that have potential benefit or are undergoing further investigation. Piracetam is a GABA and baclofen analog that is available in Europe and is used as a cognitive enhancer (81). In one well-designed study, Piracetam was tested against placebo in a double-blinded crossover trial. Patients demonstrated improved passive range, hand function, and ambulation with a minimal side effect profile (158). Gracies et al. (81) have recommended further evaluation.

The cannabinoids have been shown to have efficacy in neurologic populations and are currently under study (159,160). Specifically, improvements in spasticity have been seen in patients with MS or SCI, but with mixed results (116). Recently Collin et al. performed a double-blind placebo-controlled study over 6 weeks that compared placebo to whole plant cannabis-based medicine containing delta-9 tetrahydrocannabinol and cannabidiol. Significant improvements were obtained in the subjective measure and a trend toward significance in the Ashworth (161).

Gabapentin’s potential in spasticity management has also been examined in several studies in patients with MS and SCI (162–166). These studies suggested that at doses of 400 mg tid, gabapentin reduced spasticity with a benign side effect profile. Similarly, an animal model study demonstrated that intraperitoneal injection of gabapentin in rats with SCI reduced spasticity over a 24 hour period. The authors hypothesized that the effect was likely mediated through its effect on glutamate (167). Another anticonvulsant tiagabine, a GABA analog, was noted to decrease spasticity in an open-label trial (168). In their evaluation of outcome, the investigators noted a 50% reduction in tone on the modified Ashworth score, improvement in strength, coordination, and ROM. Two newer anticonvulsant medications, levetiracetam and pregabalin have also shown some efficacy in reducing spasticity (169–172). A retrospective review of 22 patients with spasticity resulting from TBI, MS, or CP noted that 12 patients reported reduction in spasticity with pregabalin (172). Reports have also noted reduced phasic spasticity, or “spasms” with the use of levetiracetam. In one retrospective review featuring 12 patients with MS, a significant improvement was seen in the Penn spasm score, with three patients also noting a reduction in their neuropathic pain (171). Additionally, a case report noted the elimination of painful paroxysmal spasms in a patient with “stiff person” syndrome treated with levetiracetam (170). A trial has also been proposed to treat cramps and/or spasticity in patients with amyotrophic lateral sclerosis (169).

Finally, an open-label trial with the antinarclepsy drug modafinil was noted to improve spasticity in seven out of nine patients with CP (173). A retrospective review of 30 patients with CP was conducted by the same investigators in 2004. Of the 30 patients, 23 reported diminished spasticity confirmed by increased joint mobility on physical exam; however, seven patients discontinued the medication because of an unacceptable side effect (174).

Nerve Blocks, Motor Points and Chemical Denervation
It was mentioned earlier that when the symptoms of spasticity are focal, segmental, or regional, they can be treated with chemical neurolysis or chemical denervation (CD). In simple terms, chemical neurolysis is the process of a clinician treating spasticity and the UMNS by creating a lesion in the lower motor neuron. This is true regardless of the agent used, be it phenol, ethanol, or one of the botulinum toxins. Treatment goals are very straightforward: focally treat the area of muscle overactivity with the most beneficial side effect profile. First, here are some working definitions: *Nerve block*: the application of a chemical to a nerve to impair function, either temporarily or permanently. *Chemical neurolysis or chemodenervation*: a type of nerve block where there is actual destruction of nerve tissue to give a longer-lasting block.

*Motor point block*: the condition that occurs when a portion of a nerve lower down on the nerve trunk (hopefully, below the sensory branches) is blocked to create a motor block with a more limited sensory involvement (175).

What are some of the indications for chemical denervation? Autti-Ramo et al. (176) suggested that chemical denervation should be performed as part of the planning process before performing any orthopedic hand surgery. Chemical denervation can be beneficial when one chooses realistic, obtainable goals that will benefit from local intervention and includes the planning and performance of appropriate follow-up services. In a skilled clinician’s hand, it can improve quality of life, reduce pain, increase ROM, break synergy patterns, and improve positioning and hygiene (175,177). In summary, chemical denervation is an effective intervention when used to treat a primarily focal region to address a problem caused by spasticity or another type of muscle overactivity, when the clinician has identified a clear goal of treatment (178).

Agents Available for Chemical Denervation
Phenol and ethanol were the original agents used for chemical denervation. For chemical denervation, phenol should be used in a concentration of between 5% and 7%. At that concentration, it has an immediate anesthetic effect with a subsequent neurolytic effect that matures within 2 days. Ethanol, on the other hand, should be used in concentrations ranging from 45% to 100% to create a neurolytic effect. When comparing the two agents, one will find much more literature describing the use of phenol than alcohol. However, alcohol is less toxic and is easier for clinicians to obtain. The typical duration of effect of chemical denervation with either agent is between 3 and 9 months, but it may last as long as 12 to 18 months.
Along in the process of FDA approval process. Botox has been available in the United States since 1989. Approved indications for Botox are blepharospasm, cervical dystonia, severe primary axillary hyperhidrosis, and cosmesis. Dysport and Xeomin are currently available in Europe. Meditoxin is available in Korea and its surrounding area, while CBTX-A is available in China and Brazil. Myobloc is approved for the treatment of cervical dystonia in adults in the United States.

Mechanisms of Action for the Chemical Denervation Agents

Ethanol and phenol's action are mediated via their ability to denature protein. It is a crude process that can demonstrate an effect on both motor and sensory fibers, as the action is on the nerve itself (28,175,180). By contrast, the “botulinum toxins” act on the neuromuscular junction, where they inhibit the release but not the production of acetylcholine (Ach). Ach exocytosis is an important component of transmission at the neuromuscular junction with the soluble NSF attachment receptors (SNARE) complex critical for this process. The different toxins act as proteolytic agents at different sites of the SNARE complex (Fig. 50-1). Toxin A is active on Snap-25.
(25-kD synaptosomal associated protein), whereas the B works on Synaptobrevin (also called VAMP), a protein attached to the Ach vesicle (181).

**Chemical Denervation with Phenol and Alcohol**

Numerous publications describe the efficacy of phenol and alcohol chemical denervation. Khalili et al. (182), Halpern and Meelhuysen (183), DeLateur (184), and Awad (185) are just some of the original authors who published showing the efficacy of phenol nerve blocks. Kong and Chua (186–188) have also demonstrated the efficacy of alcohol chemical denervation. This intervention has been used in various patient populations, including adult stroke, TBI, and CP (189–192).

Phenol and botulinum toxin may also be combined to treat individuals with more generalized spasticity. Gooch and Patton (193) demonstrated that this combination allowed more muscles to be treated with higher dosages in children with CP and other neurologic diseases, with minimal side effects. Additionally, Lee and Lee (194) recently reported utilizing ultrasound guidance to identify nerves during intraneural phenol injection with some increase in effective duration.

**Botulinum Toxin Therapy**

Van Ermengem isolated the *Clostridium botulinum* bacteria after musicians developed a paralytic syndrome subsequent to eating raw ham at a wedding. There are seven serotype toxins produced by the bacteria, identified as A through G. At the time of the writing of this chapter, only two formulations of botulinum toxins were approved in the United States and commercially available here (type A, Botox; type B Myobloc). At present, there are several formulations of the botulinum toxin’s at different stages of the FDA’s approval process, with Ipsen’s Dysport the most advanced at the time of drafting this chapter. No botulinum toxin formulation is approved by the FDA for the treatment of spasticity, so at this time all treatment for this condition is off label (180).

All seven serotypes of botulinum toxin demonstrate their action at the neuromuscular junction, where they inhibit the release, but not the production, of Ach. This is accomplished via proteolytic cleavage of either a membrane or soluble protein of the system involved in Ach exocytosis called the (SNAP—soluble N-ethyl-maleimide sensitive factor attachment protein Receptor complex) SNARE complex. The SNARE complex has three critical components: synaptobrevin, a protein attached to the Ach vesicle, SNAP-25 (25-kD synaptosomal associated protein), and syntaxin, which are attached to the cell membrane. Toxins B, D, F, and G cleave synaptobrevin, whereas toxins A, C, and E cleave SNAP-25. Toxin C cleaves syntaxin as well. In contrast to phenol, botulinum toxin is transported to the nerve terminals and EMG or electrical localization is needed to identify only the muscles injected (71,72).

As biologic agents the toxins are measured by their potencies. A unit of toxin is defined as the LD_{50} for an IP injection of toxin into an 18 GM Swiss-Webster Mouse. While all toxins’ potencies are based on their effect on the mouse, they do not necessarily have corresponding potencies in human. The effect on humans of 1 mouse unit of type A toxin is not the same as that of 1 mouse unit of type B toxin, as there are interspecies differences. There is no simple formula for converting a dose of type A toxin to a dose of type B toxin (71,72,180). There are also differences noted between the different type A toxins; however, it may be possible to determine ratios between the different A toxins. Wenzel et al. (195) reported that Botox was more potent than Dysport and that the ratio of dosing is 4 to 5/1. This is truly a very important issue when one is reviewing the literature or determining dosing for treatment.

**Clinical Uses of Botulinum Toxin**

There are now over 50 reported therapeutic uses for botulinum toxin, with only a few of these approved by the FDA. Therefore, the majority of them can be described as off-label indications supported by randomized controlled trials include esophageal achalasia, essential tremors, palmar hyperhidrosis, headache, and chronic anal fissure (196,197). Other conditions that botulinum toxin has been used for include the treatment of urologic dysfunction (196,198–200), poststroke shoulder pain (201,202), and temporal mandibular disorders (203,204).

With the exception of the palmar hyperhidrosis and possibly headache, toxin injection addresses symptoms secondary to dysfunctional muscle contraction. Focal treatment with botulinum toxin is often preferred to generalized treatment with an oral medication that can be a source of systemic side effects (198,199,203,204). This is especially pertinent in patients where fatigue, dizziness, and hypotension need to be avoided, such as those with MS (198). Treatment with botulinum toxin in disorders where the symptoms are primarily caused by an overactive sphincter muscle may be seen as an attractive alternative to surgery due to its reversible nature. Examples include detrusor sphincter dyssynergia, chronic anal fissure, and esophageal achalasia (196,198,199).

Botulinum toxin treatment of spasticity in upper or lower extremities has been studied widely (205–221). One of the first studies demonstrating the utility of the botulinum toxins in the management of spasticity was performed by Snow et al. (222) who demonstrated the utility of toxin in the management of adductor spasticity secondary to MS. Its efficacy in TBI was first suggested in two studies performed by Yablon and Wilson’s groups (223,224). Richardson et al. (225) reported on toxin A efficacy on patients after they failed conservative measures, demonstrated an improvement in parameters related to gait symmetry and velocity, and reported improvement in a mixed neurologic population 3 to 12 weeks after injection. Since then, multiple double-blind placebo-controlled studies have been performed investigating the use of botulinum toxin types A in the management of poststroke upper and lower limb spasticity (202,209,210,212,216,219,221,225,226). A recently published study by Simpson et al. (227) demonstrated the efficacy of and the superiority of botulinum toxin A over the oral medication tizanidine in the management of upper-extremity spasticity secondary to ABI. Finally, a recently completed evidence-based review makes the recommendation...
to consider the use of botulinum toxin A for the management of spasticity (228).

There have also been numerous studies that have addressed the efficacy of botulinum toxin A in the management of spasticity from various conditions including CP (205, 206, 211, 214, 217, 218, 229, 230), SCI (208), and MS (213, 222, 231). In general, the studies demonstrated a significant reduction in tone and improved ROM with some improvement in function (226). The only studies utilizing Myobloc to treat upper-extremity spasticity were unable to demonstrate significant improvement and reported numerous side effects (226, 232). Repeated injection with botulinum toxin type A has been shown to be safe and efficacious (209, 212, 216).

**Botulinum Toxin Varieties**

**Botulinum Toxin A (Botox).** Allergan markets the toxin type A under the trade name Botox. In 1989, the toxin was licensed by the FDA as an orphan drug for the treatment of strabismus, hemifacial spasm, and blepharospasm. For years, Botox was the only botulinum toxin available in the United States. In its original formulation lot (79–11), the protein load was 25 ng per 100 units. In December 1997, a new formulation lot (BCB2024) with a protein load of only 5 ng per 100 units was introduced to replace the old toxin. These two toxins were noted to have similar clinical efficacies, but the lower protein load was believed to have the potential for lower antigenicity (233, 234). The majority, but not all, of the studies performed with botulinum toxin A in the study of spasticity have been performed with the BCB2024 lot of Botox.

Botox comes packaged in a frozen lyophilized form of purified botulinum toxin A. It requires reconstitution with preservative-free saline prior to its use and should be used within 4 hours of reconstitution. The maximal therapeutic doses have been increasing as more information relative to safety has been obtained. Consensus guidelines from the NIH (235) suggested using maximum doses of 400 units in adults. Experienced clinicians have been pushing doses higher to improve clinical effect despite the above caveat. There still is a wide window for the risk-benefit ratio even if higher doses are used. For the primate, a lethal dose is 40 units/kg, or roughly 3,000 units for a 70-kg man (236).

**Botulinum Type A (Dysport).** Dysport, first released in 1991, is produced by Ipsen Ltd and is another botulinum toxin A. Like Botox, it is stored as a white lyophilized powder and requires reconstitution with saline prior to use. While it is not yet available in the United States, it is available in over 70 countries around the world. Dysport has been granted orphan drug status for the management of cervical dystonia by the FDA. In late 2007, Ipsen Ltd filed for a biologics license application. The FDA issued a Complete Response Letter in January 2009 that requested additional information, including the finalization of the Risk Evaluation and Mitigation Strategy (REMS) and of the draft labeling, as well as a Safety Update Report, but did not request any further studies (237). Ipsen reports optimism in bringing the product to the US market. Numerous studies have been performed that have demonstrated safety and efficacy of the agent in spasticity of various etiologies (229–231, 238–243). Currently, a large open-label trial of Dysport for the management of arm spasticity secondary to stroke is underway (244). Although Botox and Dysport are both botulinum toxin type A, they are not equivalent or interchangeable on a unit per unit dosage and the ratio between the two is 4 to 5/1 (195). As a result of a pilot study comparing three different doses of Dysport in the management of arm spasticity secondary to stroke (241), the recommended starting dosage for arm spasticity is 1,000 units.

**Botulinum Toxin B (Myobloc).** In December 2000, the FDA approved Elan Pharmaceutical’s botulinum toxin type B product, under the trade name Myobloc for the indication of cervical dystonia. Therefore, like other botulinum toxins, using it for spasticity will be off label. Unlike Botox and Dysport, Myobloc does not require reconstitution and is packaged as a liquid preparation at a pH of 5.6 that can be used right out of the vile (245). The original abstract reports by BeDell et al. (246), Hecht et al. (247), Jayasooriya et al. (248) and Moberg-Wolf and Walke (249) suggested some efficacy for the treatment of spasticity and shoulder pain. However, the more recent literature concerning the use of the B toxins in the management of spasticity has been disappointing. A double-blind placebo-controlled trial by Brashear et al. (232) demonstrated numerous side effects while being unable to show significant improvement in upper limb spasticity. A higher dose may have achieved better clinical effects but would have risked greater side effects as well. Barnes et al. reported on their clinical work with the B toxin for patients who were A resistant. They reported that secondary to either nonefficacy or unacceptable side effects, only 13 of 36 patients with cervical dystonia and 7 of 20 patients with spasticity were continuing to be treated with the type B toxin (250). In another comparison study of Toxin A (Botox) versus Toxin B (Myobloc) in the management of cervical dystonia, the toxins were shown to have equivalent benefits at 4 weeks post injection. However, the A toxin demonstrated superiority with a longer duration of effect (14/12.1 weeks) and superior side effect profile. In regards to the side effect profile, there was a lower rate reported with the A toxin in regards to dysphagia (19%/48%) and dry mouth (41%/80%) (251).

**Side Effects**

Chemical denervation with botulinum toxin does lead to occasional side effects. Systemic absorption doses occur and antibody formation is a real concern in the dystonia literature (252, 253). Herrmann et al. (254) reported a rate of 31.8% antibody formation in serum samples of children who had received multiple injections of either Botox or Dysport for the management of spasticity. Yablon et al. published a study in 2007 examining antibody formation in patients treated with at least one session of botulinum toxin A (Botox) lot (BCB2024) for the management of post stroke spasticity. They found that...
of the 191 patients sampled, only one (0.5%) demonstrated antibodies to botulinum toxin type A. They suggest that reduced risk of antibody formation may be due to the reduced protein load in the new batch of toxin (255).

Multiple studies have been published regarding the side effects of botulinum toxin injection. Reactions can be categorized as systemic or local and often depend on the area being treated. The most common systemic effects tend to be headache, flulike symptoms, fatigue, and nausea and are benign with no long-term sequelae (223). These have been reported after treatment of limb spasticity, oromandibular motor disorders, and urologic disorders (199, 204, 245, 246).

There are also case reports of iatrogenic botulism resulting in respiratory depression, dysphagia, generalized weakness, dysarthria, ptosis, and dry mouth in all treatment areas (199, 203, 204, 257, 258). Rare side effects also include a generalized skin exanthema following botulinum toxin type A injection and visual disturbances due to loss of accommodation secondary to parasympathetic dysfunction following Myobloc injection (255, 259). The increased frequency of dry mouth and the reports of visual disturbances have led authors to the conclusion that Myobloc has more parasympathetic side effects than the type A toxins (245, 260). The large multicenter studies of type A toxins also reported no increase in serious adverse events between placebo and treatment groups (212, 216, 221, 256).

Local side effects tend to reflect the area injected. There may be pain and erythema at the injection site (196) and local diffusion to surrounding muscles can cause a variety of symptoms depending on the anatomic location. Treatment of limb spasticity may cause weakness in the surrounding muscles, while injections for cervical dystonia or bruxism have been reported to cause dysphagia, dysarthria, or difficulty swallowing (203, 204, 245).

Recently, the FDA issued an early warning for both types of botulinum toxin available in the United States in response to reports of serious systemic side effects. The side effects specified were suggestive of botulism and included patients experiencing difficulty holding up their heads, prossis, dysphagia, respiratory insufficiency, weakness, numbness, as well as death. In the adult population, no deaths were reported, and none of the more severe side effects requiring hospitalization resulted in ventilatory support or a feeding tube. However, in the pediatric population, these measures were occasionally required, with some cases resulting in death. The doses used in the pediatric population ranged from 6.25 to 32 units/kg for Botox and 388 to 625 units/kg for Myobloc. In the adult population, Botox doses ranged from 100 to 700 units, while Myobloc doses varied from 10,000 to 20,000 units. This report is an early warning and further information will be obtained prior to a final determination by the FDA (261).

In summary, chemical denervation is a valuable treatment modality in the management of spasticity. Phenol and alcohol treatments require considerable skill to perform effectively, but clearly this method still has a place in the management of focal spasticity. Botulinum toxin A has a long record of success in the treatment of spasticity. The B toxin is still relatively new, and its suitability in the spasticity treatment armamentarium is uncertain as there is no large-scale study demonstrating its safety and efficacy in spasticity management. The authors are unable to give a definitive answer at this time which toxin is more effective or antigenic than the other; however, there is substantially more supportive literature for the A toxin. There is no evidence to suggest that alternating agents delay antibody formation. Final answers will have to be deferred to sometime in the future. There also is no fixed ratio of conversion when switching toxins, as they are different but related drugs, in the same fashion as calcium channel blockers. At this time, the clinician's skill and judgment will be an important factor in the dosing decision.

**Surgical Techniques**

**Intrathecal Medications**

**Baclofen**

One of the first reports on the efficacy of ITB was by Penn et al. (262), who reported its benefit in severe spasticity in patients with MS and SCI. The authors demonstrated that ITB was acutely beneficial, via a double-blinded placebo ITB injection over 3 days. They also were able to demonstrate a long-term treatment effect ranging from 10 to 32 months after pump insertion in an open-label follow-up study. In 1993, Albright et al. (263) demonstrated that ITB was also useful in the management of CP-related spasticity. It was approved by the US FDA for intrathecal therapy for management of severe spasticity of cerebral origin in 1996. ITB has been shown to be useful and well tolerated in the treatment of other neurologic populations, including those suffering from TBI (264), anoxia (265), and stroke-related spastic hemiplegia (266–270). Adequate management of spasticity from cerebral origin with ITB generally requires higher dosing compared to spasticity of spinal origin (271).

Potential candidates for ITB undergo screening for potential efficacy. Their spasticity is evaluated in the normal trial may not address functional issues adequately. A simple reduction in tone that is considered a success in the normal trial may not address functional issues adequately. A trial using a continuous infusion of ITB may be more appropriate in those cases. In 2007, Bleyenheuft et al. reported on their use of an external pump connected to an intrathecal
catheter on seven ambulatory patients with spasticity and CP. The authors found that not only did spasticity decrease but they were also able to assess the patient’s ability to ambulate while on ITB. They reported complications of a chemical meningitis and two CSF leaks but were positive about the overall procedure’s safety and efficacy (272).

The advantage of the ITB system is that it delivers a continuous supply of baclofen through a catheter that is directly connected to the subarachnoid space. As a result, there is reduced spasticity at a dose that is substantially lower than that required via oral administration (72). Dralle et al. (273) reported that required dosing was roughly 1% of that needed for oral administration. As a result, the CNS side effects that are commonly seen with oral antispasticity agents are greatly reduced with ITB (274). The placement of an ITB device should not be taken lightly. The system requires a commitment to a surgical procedure and the insertion of a foreign object that will have to be replaced in 3 to 7 years, based on current technology.

The ITB system, while often a good therapy choice, can be associated with adverse effects. Problems with the ITB device itself may include high cost (275), infection risk (276), CSF leaks (277), pump dysfunction, and kinking or disconnection of the catheter tubing (278). Problems associated with the ITB medication administration may include baclofen overdose, withdrawal and associated seizures (279–281), and chronic constipation. Some studies reveal concern for another long-term adverse effect of ITB therapy of rapid progression of scoliosis (282–284).

Originally, the most cephalad placement of the catheter was between the levels of T8 and T10 (285). More recently, surgeons have been placing it higher (264,286) and reported significant improvement in spasticity in both the upper and lower extremities. Broseta et al. (287) reported the safe administration of ITB, with the catheter placed as high as C4. The reason for limiting the cephalad catheter placement is concern for respiratory suppression, with baclofen administration near muscles of respiration. ITB has proven beneficial for spastic hypertonia in patients with poststroke unilateral symptoms, without reducing weakening the unaffected limb (268).

**Other Intrathecal Medications**

Baclofen is not the only medication that has been used intrathecally. Other pharmacologic agents that have been administered intrathecally, either alone or in combination with baclofen, include clonidine, midazolam, morphine, lidocaine, and fentanyl (288–292). Fentanyl has been shown to be effective in cases of developed resistance to ITB (290). Midazolam, a benzodiazepine, has been used intrathecally for the management of spasticity. Work in animals shows that the agent functions as a CNS depressant, analgesic, anticonvulsive, and myorelaxant (81). Some studies report neurotoxicity from intrathecal midazolam, with animal data showing potential toxicity to the blood-brain barrier, neurons, and the myelin sheath (293,294). As a result, its use in humans is controversial and has been demonstrated in case reports for pain management (81,295–298). Finally, midazolam use for spasticity management is also limited by its sedating properties.

**Intrathecal Morphine.** Morphine has also been used intrathecally in the management of spasticity, with reduction in tone noted with boluses of 1 to 2 mg (299) and doses of 2 to 4 mg administered daily in the treatment of SCI-induced spasticity (300). The authors also reported that only one of the 12 patients who had responded to treatment had become resistant after 3 years of treatment. Intrathecal morphine has been used successfully in combination with ITB to control spasticity and pain in patients with MS (301). Case report evidence also supports using intrathecal morphine to control spasticity as a temporary agent for a “drug holiday” from ITB in cases of ITB resistance (302). The significant side effect profile of intrathecal morphine limits its use, including development of tolerance, itching, GI disturbances, hypotension, and urinary retention (81). Lydon et al. (303) reported problems with gastric emptying, whereas Glass (304) reported that an intrathecal bolus dose of 0.4 mg induced respiratory depression. Itching is the most common side effect with intrathecal morphine, whereas respiratory depression, common with doses of 1 to 2.5 mg, is the most troubling. As a result, the role of intrathecal morphine in the management of spasticity is uncertain (81).

**Intrathecal Clonidine.** Intrathecal clonidine has also been used in the management of spasticity (288), especially in spinal cord patients (289). Remy-Netis et al. (305) reported that the use of intrathecal clonidine on patients with incomplete SCI showed significant electrophysiologic changes, in addition to three of eight patients demonstrating improvements in gait velocity. It has also been shown to demonstrate decreases in the stretch reflex in this population. The functional results of intrathecal clonidine are mixed. In the previously mentioned trial, three participants did indeed demonstrate an increase in their preferred walking speed, and three others demonstrated a decrement in their performance (loss of their ability to stand), most likely as a result of the functional benefits of their tone. This loss of the ability to stand was not dose related (289). Another potential benefit of intrathecal clonidine is its ability to reduce detrusor hyperreflexia in the SCI population (306,307).

The side effect profile for intrathecal clonidine is potentially problematic. In addition to the functional loss reported earlier, intrathecal clonidine can also cause cardiac problems (hypotension, bradycardia, and decreased cardiac blood flow) (289,308), dry mouth, and sedation. In particular, the hemodynamic side effects of intrathecal clonidine may be its most limiting factor (81).

**Other Neurosurgical Procedures**

In addition to intrathecal administration of pharmacologic agents, other neurosurgical procedures are available for the management of spasticity. DRR is a procedure that has
demonstrated efficacy in the management of spasticity. This procedure should be considered when the patient's condition has been recalcitrant to other more conservative measures or when financial or compliance issues direct treatment interventions. Much of the evidence demonstrating the efficacy of rhizotomy is found in the CP literature (309–312). DRR has been shown to be useful in other populations, such as those with MS (313,314). A radiofrequency rhizotomy was performed by Kasdon and Lathi in 1974, and the authors reported significant improvement in 24 of the 25 patients treated (315).

Initially, rhizotomy did not selectively ablate motor and sensory roots. The technique of interrupting a limited number of the most pathologic sensory rootlets and sparing the remainder was an innovation derived later in the history of the rhizotomy procedure. This method, called a selective DRR, reduces the propensity of anesthesia due to sensory nerve root ablations and leaves the motor system intact. Much work concerning this subject has been published by Sindou et al. (314,316,317). Studies have shown that the selective DRR continues to be beneficial years after surgery. Gul et al. (310) showed that 5 years after surgery, their patients still maintained improvements in tone and muscle strength. Mittal et al. (318) demonstrated that fine motor improvement seen after 1 year was still maintained at 3- and 5-year follow-up. Other more recent studies by Langerak et al. in 2007 and 2008 (319,320) demonstrated preservation of improved ROM in the lower extremities and ambulation 20 years after DRR, with none of the patients reporting the need for help with ADL. DRR has also been shown to reduce the need for orthopedic surgery, particularly soft-tissue procedures, which is best achieved when performing the DRR in younger patients with CP (321).

While the functional benefits of DRR appear to be maintained many years after the procedure, several studies have evaluated the long-term effects, particularly on the progression of spinal deformities. Steinbok (322) showed in a study of 104 children with spastic CP who underwent selective DRR that 54.8% had scoliosis greater than or equal to 10 degrees and six had a more significant scoliosis greater than 35 degrees. These numbers, however, were within the range of incidence of scoliosis in patients with spastic CP who did not undergo the surgical procedure, and only three of the patients required surgical intervention to correct the deformities. Nonetheless, the authors concluded that awareness of the progression of spinal deformities after selective DRR may be an important consideration to consider. Farmer and Sabbagh (323) studied 101 patients with a mean follow-up of approximately 6 years, showing that scoliosis was present in 45%, but only in two cases did the curve reach above 20 degrees, and no cases required bracing or fusion. Overall, Farmer and Sabbagh concluded that the natural history of spinal deformity in children with spastic CP was not worsened by selective DRR, and in fact the procedure before the age of 5 has a protective effect on lordosis.

Numerous complications have been reported as a result of selective dorsal root DRR, but they are often transient. Complications include bronchospasm, aspiration pneumonia, urinary retention, ileus, sensory loss, hypotonia, and bowel dysfunction (324–326). Nonetheless, the authors of these studies concluded that the benefits of the procedure far outweigh the potential complications.

There has been some controversy regarding the number and selection of rootlets to be ablated. Mittal et al. (327) stated that using electrophysiologic stimulation to identify nerve rootlets was very effective in choosing pathologic rootlets for ablation. They also reported the rootlet scoring was 93% reproducible and felt their technique resulted in great reduction in tone with only very limited loss of strength. On the other hand, Sacco et al. (328) reported that they had similar results to selective DRR when they performed nonspecific DRR on ten patients. Finally, Bertelli et al. (329) reported a new method in the treatment of upper-extremity spasticity. They performed a brachial plexus DRR and reported reduction in tone without any sensory difficulties. In summary, selective DRR remains a possible treatment option for spasticity, particularly in patients with CP; but its benefits, long-term outcomes, and proper case assignment remain controversial (330).

For patients unable to benefit from any of the above interventions, other aggressive modalities such as myelotomy or cordotomy are available. These procedures are generally reserved for intractable cancer pain, and most studies describing myelotomy or cordotomy involve pain management (331–333). Bischoff has advocated the use of myelotomy or cordotomy in severely spastic patients; however, these more aggressive interventions are controversial because they may cause dysfunction of the bowel and bladder (72).

**Orthopedic Procedures**

Although the preceding discussion highlights the numerous pharmacologic and therapeutic modalities available in the management of spasticity, sometimes orthopedic interventions are still necessary. Rhizotomy, oral agents, and intrathecal agents are unable to act directly on muscle and tendon, the way a chemoneurolytic or orthopedic procedure can. This may be especially true in the management of childhood-related spasticity (334), where orthopedic surgical intervention can be effective in improving both function and the limitation of complications.

Orthopedic procedures are often irreversible, and physicians sometimes experience resistance from patients and their family. As a consequence, realistic goal setting is critical and requires that the family, health care providers, and patients, all be in agreement. It is also important for the treatment team to understand the pathophysiology, kinesiology, and overall patient function. Improper preprocedure planning can result in limited benefit or even a worsening of the patient's function and quality of life. When planned, executed, and given appropriate follow-up, orthopedic interventions can improve positioning or hygiene and can play a major role in active functioning tasks such as ambulation (72).

In the limited space available here, it is impossible to discuss all surgical options. The goal is to assist the reader in developing general principles for management decisions.
As mentioned earlier, careful planning is critical before any operation. If available, dynamic EMG can be exceptionally beneficial in the decision-making process (72). This point was demonstrated effectively by a 2003 study performed by Keenan et al. The authors investigated the role of dynamic motor control studies on their planning of orthopedic intervention on 21 patients with spastic elbow deformity. Two orthopedic surgeons were asked to plan a surgical intervention before and after motor studies were performed. As a result of the evaluation, the surgical plans were changed 57% of the time after the dynamic testing was performed (335).

When this type of testing is not available, previous results from chemoneurolysis and temporary nerve blocks, in addition to clinical evaluation, may assist in the decision process. The evaluation should look for muscles co-contracting or firing out of phase. These muscles may benefit from lengthening or a transfer of distal attachment site.

An example is a brachioradialis muscle that fires during attempted extension. Transferring the muscle's distal insertion to the extensor surface will allow it to be of assistance rather than a hindrance. Another example is foot inversion during gait. This movement can be caused by a host of muscles, such as the gastrocnemius complex, tibialis anterior, tibialis posterior, and flexor digitorum longus, and flexor and extensor hallucis longus. Dynamic EMG (336) or selective botulinum toxin injection may help identify the offending muscle or muscles.

Based on these tests and clinical evaluation, surgery can be planned and performed with the best chance for functional benefit. Increased activity involving elbow flexion may require the treatment of one, two, or all three elbow flexors that can be treated with lengthening procedures (337). When addressing the shoulder for spasticity, pain, or function, it is important to correctly identify and target specific muscles to optimize the intervention (335). The tibialis anterior is best known for its critical function related to ankle dorsiflexion. It can also, when overfiring, cause inversion of the foot during swing phase. The split anterior tibial tendon transfer procedure, which divides the tendon's distal attachment on the medial surface and anchors it into the medial and lateral aspect of the foot, reduces the inversion moment with limited loss of dorsiflexion strength (338). Both passive and sometimes functional improvement can be seen when hyperactivity and knee flexion contractures are treated with a release of the hamstrings with positional improvement (339). With shortening of the Achilles tendon, the performance of a Z-plasty procedure that lengthens the tendon may be of benefit. However, there may be some loss of strength at the ankle joint as a result. Keenan et al. (340) showed that transfer of the flexor hallucis and the digitorum longus to the os calcis improved ankle strength and reduced the need for long-term use of a brace. An iliopsoas tenotomy may be necessary when there is significant spasticity or a hip flexor contracture of greater than 20 degrees after the application of more conservative measures (339). In patients with severe spasticity from MS or CP, the problem of hip subluxation may be encountered. This can be treated with adductor myotomies, or if not successful, femoral osteotomy (72).

More aggressive intervention may be appropriate in more severely affected patients. A 2007 report on eight severely impaired, nonambulatory patients with hip and knee flexure contractures over 90 degrees, evaluated the results of knee disarticulation and hip flexor release. It demonstrated significantly improved positioning, hygiene, passive ROM at the hip, quality of life, and decreased pain. In addition, all patients improved to the point that they were able to use a wheelchair after surgery (341).

Patients with severe spasticity often receive multiple types of treatments. Orthopedic surgery is often contemplated in combination with, or instead of, ITB or DRR. Gerszten et al. (342) reviewed the CP literature. They found that 18 out of 28 patients who were to undergo orthopedic procedures after ITB placement did not require it, as their spasticity had substantially improved. Grabb and Doyle (334) generally agree with this recommendation for their ambulatory patients with spasticity and structural issues, as they perform ITB or DRR before orthopedic intervention. Their exception was the lower-level ambulator, where they feel the tone is a functional benefit to ambulation and they are more likely to perform soft-tissue release first. For the nonambulatory patient, they recommend ITB before orthopedic intervention. Additionally, a 2006 review concluded that selective DRR reduces the requirement for orthopedic surgery compared to historical controls and that when performed at a young age (2 to 4 years old), selective DRR reduces the need for future orthopedic intervention (321).

CONCLUSION

Muscle overactivity resulting from the UMNS continues to be a problem that can negatively affect a person's recovery and function. New treatment modalities have been developed that have greatly increased the options available to the treatment team. It is important for the physician to work together with other clinicians, the patients, and their caregivers to develop a comprehensive treatment approach with realistic and meaningful goals. With the rapid development of new treatments, greater improvement for even a longer period of recovery is possible. Clinicians should think creatively when choosing their treatment interventions and working in partnership with their patients strive for greater functional recovery.

REFERENCES


271. FDA. Early Communication about an Ongoing Safety Review Botox and Botox Cosmetic (Botulinum toxin type A) and Myobloc (Botulinum toxin type B). 2003;29(4):2348–2354.


308. Canavero S, Bonicalzi V, Clemente M. No neurotoxicity from long-term (>5 years) intrathecal infusio
Voiding dysfunctions are commonly encountered in patients who are referred for rehabilitation. These voiding problems may result from medications, cognitive changes, physical impairments, or neurologic etiologies. Timely identification of voiding dysfunctions, treatment, and follow-up are important. This is particularly true in the rehabilitation setting, where voiding dysfunctions may cause patient embarrassment, interruption of therapy, and increased morbidity and ultimately may make the difference between reintegration into the community and being confined to a home or nursing home.

ANATOMY AND PHYSIOLOGY OF THE UPPER AND LOWER URINARY TRACTS

Upper Urinary Tracts
The kidney is composed of two parts: the renal parenchyma, which secretes, concentrates, and excretes urine; and the collecting system, which drains urine from multiple renal calyces into a renal pelvis. The renal pelvis then narrows to become the ureter. The point at which the renal pelvis becomes the ureter is known as the ureteropelvic junction (1).

The ureter travels down to the bladder. It is approximately 30 cm in length in the adult. The ureter has three areas of physiologic narrowing that take on clinical significance with respect to possible obstruction from stones. These areas are the ureteropelvic junction, the crossing over of the iliac artery, and the ureterovesical junction (2,3). The ureterovesical junction is the place where the ureteral orifice opens into the bladder. Its function is to allow urine to flow into the bladder but prevent reflux into the ureter. This can be accomplished because the ureters traverse obliquely between the muscular and submucosal layers of the bladder wall for a distance of 1 to 2 cm before opening into the bladder (Fig. 51-1). Any increase in intravesical pressure simultaneously compresses the submucosal ureter and effectively creates a one-way valve (4). Presence of ureteral muscle in the submucosal segment also has been shown to be important in preventing reflux (5). However, if there is high intravesical pressure the submucosal segment can be compressed and prevent urine from draining down into the bladder, causing a physiological obstruction.

Normal Urine Transport from the Kidneys to the Bladder
Urine transport is the result of both passive and active forces. Passive forces are created by the filtration pressure of the kidneys. The normal proximal tubular pressure is 14 mm Hg, and the renal pelvis pressure is 6.5 mm Hg, which slightly exceeds resting ureteral and bladder pressures. Active forces are the result of peristalsis of the calyces, renal pelvis, and ureter. Peristalsis begins with the electrical activity of pacemaker cells at the proximal portion of the urinary collecting tract (6).

For the ureter to propel the bolus of urine efficiently, the contraction wave must completely coat the ureteral walls (7). Ureteral dilation for any reason results in inefficient propulsion of the urine bolus, and this can delay drainage proximal to that point. This can result in further dilation and, over time, lead to hydrenephrosis.

Lower Urinary Tracts
Anatomically, the bladder is divided into the detrusor and the trigone. The detrusor is composed of smooth muscle bundles that freely crisscross and interlace with each other. Near the bladder neck, the muscle fibers assume three distinct layers. The circular arrangement of the smooth muscles at the bladder neck allows them to act as a functional sphincter. The trigone is located at the inferior base of the bladder. It extends from the ureteral orifices to the bladder neck. The deep trigone is continuous with the detrusor smooth muscle; the superficial trigone is an extension of the ureteral musculature (see Fig. 51-1) (4).

There is no clear demarcation of the musculature of the bladder neck and the beginning of the urethra in a man or woman. In a woman, the urethra contains an inner longitudinal and outer semicircular layer of smooth muscle. The circular muscle layer exerts a sphincteric effect along the entire length of the urethra, which is approximately 4 cm long.

In a man, the urethra runs the length of the penis. It begins at the meatus and is surrounded by the spongy tissue of the corpora cavernosus. In addition to the corpora cavernosus, which runs on the underside of the penis, the penis is made up of two corpora cavernosa that contain the spongy erectile tissue. The urethra is divided into the posterior or prostatic urethra, extending from the bladder neck to the urogenital diaphragm, and the anterior urethra, which extends to the meatus. The junction between the anterior and posterior urethra is known as the membranous urethra.
Urinary Urethral Sphincters

Traditionally, the urethra has been thought to have two distinct sphincters, the internal and the external or rhabdosphincter. The internal sphincter is not a true anatomic sphincter. Instead, in both men and women, the term refers to the junction of the bladder neck and proximal urethra, formed from the circular arrangement of connective tissue and smooth muscle fibers that extend from the bladder. This area is considered a functional sphincter because there is a progressive increase in tone with bladder filling so that the urethral pressure is greater than the intravesical pressure. These smooth muscle fibers also extend submucosally down the urethra and lie above the external rhabdosphincter (8).

In a man, the external or urethral rhabdosphincter often is diagrammatically illustrated as a thin circular band of striated muscle forming a diaphragm just distal to the prostatic urethra (i.e., membranous urethra). In an anatomic study, however, Myers et al. (8) reconfirmed earlier studies showing that the urethral external striated sphincter does not form a circular band but has fibers that run up to the base of the bladder. The bulk of fibers are found at the membranous urethra (9). This sphincter

---

is under voluntary control. The striated muscular fibers in both man and woman are thought to have a significant proportion of slow-twitch fibers with the capacity for steady tonic compression of the urethra. In a woman, striated skeletal muscle fibers circle the upper two thirds of the urethra (9).

**STRUCTURE AND FUNCTION OF THE MALE AND FEMALE CONTINENCE MECHANISM**

In a man, the structures responsible for continence at the level of the membranous urethra include the mucosa, longitudinal smooth muscle of the urethra, striated sphincter, and levator ani musculature. Traditionally, the striated sphincter has been considered responsible for maintaining continence. However, experimental paralysis of the striated sphincter and levator ani following surgery for prostate outlet obstruction did not result in incontinence. This demonstrated the important role of the smooth muscle fibroelastic component of the membranous urethra. The increased tone at the bladder outlet (i.e., internal sphincter) also helps maintain continence (10).

In a woman, there are three important factors in maintaining continence:

- Adequate pelvic floor support from the endopelvic fascia and anterior vagina.
- Good sphincter function.
- Maintenance of the intra-abdominal position of the proximal urethra.

During an increase in intra-abdominal pressure, continence is maintained by the downward-moving pelvic viscera compressing the urethra against the layer of endopelvic fascia and distribution of the increase of intra-abdominal pressure to the proximal intra-abdominal urethra. The urethral epithelium, which is sensitive to estrogen, is believed to help maintain continence by forming a mucosal seal (9).

**NEUROANATOMY OF THE LOWER URINARY TRACT**

Urine storage and emptying is a function of interactions among the peripheral parasympathetic, sympathetic, and somatic innervation of the lower urinary tract. Additionally, there is modulation from the central nervous system (CNS).

**Bladder Neuroanatomy**

**Efferent System**

The parasympathetic efferent supply originates from a distinct detrusor nucleus located in the intermediolateral gray matter of the sacral cord at S2 to S4. Sacral efferents emerge as preganglionic fibers in the ventral roots and travel through the pelvic nerves to ganglia immediately adjacent to or within the detrusor muscle to provide excitatory input to the bladder. After impulses arrive at the parasympathetic ganglia, they travel through short postganglonics to the smooth muscle cholinergic receptors. These receptors, called cholinergic because the primary postganglionic neurotransmitter is acetylcholine, are distributed through the bladder. Stimulation causes a bladder contraction (11,12).

The sympathetic efferent nerve supply to the bladder and urethra begins in the intermediolateral gray column from T11 through L2 and provides inhibitory input to the bladder. Sympathetic impulses travel a relatively short distance to the lumbar sympathetic paravertebral ganglia. From here, the sympathetic impulses travel along long postganglionic nerves in the hypogastric nerves to synapse at \(\alpha\) - and \(\beta\)-adrenergic receptors within the bladder and urethra. Variations in this anatomic arrangement do occur; sympathetic ganglia sometimes also are located near the bladder, and sympathetic efferent fibers may travel along the pelvic as well as the hypogastric nerves (Fig. 51-2) (11,12).

Sympathetic stimulation facilitates bladder storage because of the strategic location of the adrenergic receptors.

afferents that stimulate voiding are those that pass to the sacral cord via the pelvic nerves. These afferents include two types of afferents, small myelinated A-delta and unmyelinated (C) fibers.

The small myelinated A-delta fibers respond in a graded fashion to bladder distention and are needed for normal voiding. The unmyelinated (C) fibers have been termed silent C-fibers because they do not respond to bladder distention and therefore are not essential for normal voiding. However, these silent C-fibers do exhibit spontaneous firing when they are activated by chemical or cold temperature irritation at the bladder wall. Additionally, the unmyelinated (C) fibers (rather than A-delta afferents) have been found to “wake up” and respond to distention and play an important role in stimulating uninhibited bladder contractions in following suprasacral SCI. This role of the unmyelinated (C) fibers has been demonstrated in studies using intravesical administration of capsaicin and resiniferatoxin. Both of these agents are potent C-fiber afferent neurotoxins. In non-SCI animals (with A-delta afferents), there was no blockage of bladder contractions with bladder distention. However, in SCI animals (with “awake” C-fiber afferents) have been found to “wake up” and respond to distention and play an important role in stimulating uninhibited bladder contractions in following suprasacral SCI.

Bladder Neurotransmitters
It is known that there are more transmitters than acetylcholine and norepinephrine, including nitric oxide, vasoactive intestinal polypeptide, endogenous opioid peptides, and neuropeptide Y.
These transmitters may work independently or help modulate the classic neurotransmitters. Nitric oxide and vasoactive intestinal polypeptide have smooth muscle relaxant effects. This helps explain the concept of atropine resistance. It has been found that a single neurotransmitter-blocking agent such as atropine fails to suppress 100% of the bladder or urethral activity (14,18). This explains why a combination of agents may be more effective than a higher dose of a single agent.

**Urethral Sphincter Innervation**

The external urethral sphincter classically has been described as having somatic innervation, allowing the sphincter to be closed at will. Somatic efferents originate from a pudendal nucleus of sacral segments from S1 to S4. Somatic efferents then travel through the pudendal nerve to the neuromuscular junction of the striated muscle fibers in the external urethral sphincter.

The internal urethral sphincter has been described as being under control of the autonomic system. This area has a large number of sympathetic α-receptors, which cause closure when stimulated. Animal studies have revealed that nitric oxide is an important parasympathetic neurotransmitter mediating relaxation of the urethral smooth muscle (8,18).

The distinction between the internal and external sphincter is becoming less clear. Elbadawi and Schenk (19) reported histochemical evidence of a triple innervation pattern of the external sphincter (i.e., rhabdosphincter) in five mammalian species, with dual sympathetic and parasympathetic autonomic components superimposed on the somatic component. Sundin and Dahlstrom (20) demonstrated sprouting and increasing adrenergic terminals after parasympathetic denervation in cats. Crowe et al. (21) reported a substantial invasion of adrenergic nerve fibers in smooth and striated muscle in the urethra in SCI patients with lower motor neuron lesions.

**Influences of the Central Nervous System on the Lower Urinary Tract**

Facilitation and inhibition of the autonomic nervous system are under control of the CNS. There are several theories of how this occurs. Denny-Brown and Robertson (22) suggested that micturition was primarily mediated by a sacral micturition reflex. According to their theory, descending nervous system pathways modulate this micturition reflex (21). Barrington, Bradley, and de Groat thought that facilitative impulses to the bladder originated from a region of the anterior pons termed “Barring-ton’s center” (23).

De Groat and associates additionally stressed the importance of the sympathetic nervous system in facilitating urine storage (24). Carlson provided evidence that this pontine mesencephalic area also plays a role in coordinating detrusor and sphincter activity. Stimulation of Barrington’s center significantly decreased electromyographic (EMG) activity in the periurethral-striated sphincter while causing a bladder contraction (25).

Transection experiments in cats suggest that the net effect of the cerebral cortex on micturition is inhibitory. This also is true for the basal ganglia and corresponds to clinical findings of detrusor hyperreflexia in those with basal ganglia dysfunction (e.g., Parkinson’s disease). The cerebellum is thought to maintain tone in the pelvis floor musculature and influence coordination between periurethral striated muscle relaxation and bladder emptying (15,25).

**Normal Voiding Physiology**

Micturition should be considered as having two phases: the filling (storage) phase and the emptying (voiding) phase. The filling phase occurs when a person is not trying to void. The emptying phase occurs when a person is attempting to void or told to void.

During filling (filling or storage phase), there should be very little rise in bladder pressure. As filling continues, low intravesical pressure is maintained by a progressive increase in sympathetic stimulation of the β-receptors located in the body of the bladder that cause relaxation, and stimulation of the α-receptors located at the base of the bladder and urethra that cause contraction. Sympathetic stimulation also inhibits excitatory parasympathetic ganglionic transmission, which helps suppress bladder contractions. During the filling phase, there is a progressive increase in urethral sphincter EMG activity (26). Increased urethral sphincter activity also reflexly inhibits bladder contractions. When a bladder is full and has normal compliance, intravesical pressures are between 0 and 6 cm H₂O and should not rise above 15 cm H₂O. Filling continued past the limit of the viscoelastic properties of the bladder results in a steady progressive rise in intravesical pressure (27). This part of the filling curve usually is not seen in a person with normal bladder function, because this much distension would cause significant discomfort and not be tolerated.

When a patient is told to void (voiding or emptying phase), there should be cessation of urethral sphincter EMG activity as well as a drop in urethral sphincter pressure and funneling of the bladder neck. There is no longer reflex inhibition to the sacral micturition center from the sphincter mechanism. This is followed by a detrusor contraction. The urethral sphincter should remain open throughout voiding, and there should be no rises in intra-abdominal pressure during voiding. In younger individuals, there should be no postvoid residual, although postvoid residuals may increase with aging.

**Geriatric Voiding Physiology**

The aging process often affects voiding physiology. The kidneys undergo an age-related decrease in glomerular blood flow and renal blood flow (28). The elderly also experience a loss in concentrating ability and often excrete most of their fluid intake at night, even in the absence of medical conditions such as prostate outlet obstruction, diabetes, remobilization of lower-extremity edema, and use of evening diuretics (29).

Detrusor overactivity has been reported as the most common type of voiding dysfunction in incontinent elderly men and women (29). More recently, this has been called an overactive bladder. This may be associated with a central nervous lesion (such as a stroke, head injury, cervical disk disease, etc.).
Pediatric Voiding Physiology

Voiding physiology also changes with age in children. In the newborn, the sacral micturition reflex is primarily responsible for voiding. Because the brain stem is intact, there is coordination of the bladder contraction with sphincter relaxation; however, there is little inhibition of the micturition reflex from the cerebral cortex. As the child grows, the voided volume increases and voiding frequency decreases. By 3 years of age, most children have some voluntary control of voiding. This control usually is complete by the age of 4 years. There are some neurologically intact children, however, for whom complete control of voiding may take 5 or 6 years (34).

CLASSIFICATION OF VOIDING DYSFUNCTION

There have been a wide variety of classifications to describe voiding dysfunctions. These classifications have been based on neurologic lesion (e.g., Bors-Comarr, Bradley), urodynamic findings (e.g., Lapides), functional classification (e.g., Wein), and combination of bladder and urethral function based on urodynamics (e.g., International Continence Society) (35–39). The International Continence Society classification: The Standardisation of Terminology of Lower Urinary Tract Function: Report from the Standardisation Sub-committee of the International Continence Society has become widely accepted not only because it provides a standardization of terminology describing bladder and urethral function but also because a description of each of these terms is provided (39). Table 51-1 shows some highlights of

<table>
<thead>
<tr>
<th>Table 51.1 Urodynamic Terminology—Highlights From International Continence Society Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bladder</strong></td>
</tr>
<tr>
<td>Bladder—function</td>
</tr>
<tr>
<td>Normal (during filling)</td>
</tr>
<tr>
<td>Detrusor overactivity (during filling)</td>
</tr>
<tr>
<td>Neuropathic detrusor overactivity</td>
</tr>
<tr>
<td>Idiopathic detrusor overactivity</td>
</tr>
<tr>
<td>Normal (during emptying)</td>
</tr>
<tr>
<td>Underactive (during emptying)—inadequate magnitude or duration of a contraction to empty the bladder</td>
</tr>
<tr>
<td>Acontractile (no contractions during urodynamics)</td>
</tr>
<tr>
<td>Areflexic (no contractions due to neurogenic cause)</td>
</tr>
<tr>
<td>Bladder—sensation</td>
</tr>
<tr>
<td>Normal</td>
</tr>
<tr>
<td>Increased</td>
</tr>
<tr>
<td>Reduced</td>
</tr>
<tr>
<td>Absent</td>
</tr>
<tr>
<td>Nonspecific</td>
</tr>
<tr>
<td>Bladder capacity</td>
</tr>
<tr>
<td>Small (350 mL)</td>
</tr>
<tr>
<td>Average (350–650 mL)</td>
</tr>
<tr>
<td>Large (650 mL)</td>
</tr>
<tr>
<td>Compliance</td>
</tr>
<tr>
<td>Low (hypocompliance)</td>
</tr>
<tr>
<td>Normal (350–650 mL with little pressure rise)</td>
</tr>
<tr>
<td>High (hypercompliance)</td>
</tr>
<tr>
<td>Urethra</td>
</tr>
<tr>
<td>Normal</td>
</tr>
<tr>
<td>Incompetent (during filling)</td>
</tr>
<tr>
<td>Obstructive (during voiding)</td>
</tr>
<tr>
<td>Mechanical (stricture, bladder outlet obstruction)</td>
</tr>
<tr>
<td>Overactive (detrusor/[external] sphincter dyssynergia)</td>
</tr>
</tbody>
</table>
Incontinence
Caused by the bladder
Detrusor overactivity
Decreased capacity
Low bladder wall compliance
Normal (cognitive/mobility issue)
Caused by the outlet
Incompetent sphincter
Retention
Caused by the bladder
Detrusor areflexia
Large capacity/high compliance
Normal (cognitive/mobility issue)
Caused by the outlet
High voiding pressure with low flow rate
Internal sphincter dysynergia
External sphincter dysynergia
Overactive sphincter mechanism (i.e., sphincter or pseudosphincter dysynergia)
Retention and incontinence
Caused by the bladder
Uninhibited contractions with underactive detrusor
No contractions
Normal (cognitive/mobility issue)

**TABLE 51.2 Urodynamic and Functional Classification**

The urodynamic terminology from the International Continence Society Classification, Table 51-2 shows Wein's classification. This classification is helpful at directing treatment because it is based on a clinical problem (incontinence or retention), which can then be applied to specific urodynamic findings. A urodynamic basic SCI data set has recently been developed, which should be helpful, collecting information in a standardized way regarding urodynamic bladder and sphincter function in those with SCI (40).

**VOIDING DYSFUNCTIONS FOUND IN COMMON NEUROLOGIC DISORDERS**

**Suprapontine Lesions**

Any suprapontine lesion may affect voiding. Lesions may result from cerebrovascular disease, hydrocephalus, intracranial neoplasms, traumatic head injury, Parkinson's disease, and multiple sclerosis. It should be noted that multiple sclerosis is unique among the suprapontine lesions because it also affects the white matter of the spinal cord and often has a relapsing and remitting nature. The expected urodynamic finding following a suprapontine lesion is detrusor hyperreflexia without detrusor-sphincter dysynergia. Because of various factors such as medications, prostate obstruction, and possible normal bladder function but poor cognition, the voiding dysfunctions may be very different from expectations. Voiding dysfunction following cerebrovascular accident (CVA), Parkinson's disease, and multiple sclerosis has been studied more extensively than those associated with other suprapontine lesions and is reviewed in the following discussion.

**Cerebrovascular Accidents**

After a CVA, some patients initially have acute urinary retention. The reason for this detrusor areflexia is unknown. Urinary incontinence, however, is the most common urologic problem following an acute CVA. Various series have reported that 40% to 60% of patients are incontinent 1 week post-CVA (41–43). In the inpatient rehabilitation setting, a 33% incidence of incontinence during the first 3 months post-CVA has been reported (44). It also has been well documented that this problem significantly improves or resolves in the majority of patients. At 1 month, the percentage of incontinent patients dropped to between 29% and 42%. By 6 months to 1 year post-CVA, 14% to 15% of patients still were incontinent, which is similar to the 15% to 30% incidence in the general geriatric population (41–43). Risks for incontinence poststroke include age greater than 75 years, dysphagia, motor weakness, and visual field defects. At 2 years poststroke, incontinent patients versus continent patients had higher case fatality rates (67% vs. 20%), higher institutional rates (39% vs. 16%), and greater disability (39% vs. 5%) (43). A recent study has further categorized poststroke urinary incontinence as urge incontinence or impaired awareness urge incontinence. Those with impaired awareness urge incontinence had poorer attention and outcomes. Impaired awareness incontinence was also found to be a strong predictor factor for mortality and nursing home residence at 1 year. It would be expected that those with impaired awareness incontinence would not do well (45).

Detrusor overactivity with uninhibited bladder contractions is the most common urodynamic finding following a stroke. It has been reported to occur 70% to 90% of the time (42,46). One hypothesis for this finding is the release of the spinal micturition reflexes from the inhibitory higher centers. Symptoms, however, often do not correlate with urodynamic findings. Linsenmeyer and Zorowitz evaluated 33 consecutive incontinent patients who were 1 to 3 months post-CVA. They found that whereas 82% of men had uninhibited contractions, 43% also had urodynamic evidence of outlet obstruction. Six percent of the incontinent group had no bladder contractions and twelve percent had normal urodynamic findings (46). Voluntary sphincter contractions (i.e., pseudodyssynergia) to keep from voiding should not be misinterpreted as true detrusor-sphincter dysynergia. In a review of 550 patients, Blaivas (47) reported that patients with CVAs do not develop true detrusor-sphincter dysynergia. EMG studies by Siroky and Krane (48) gave similar results.

**Parkinson’s Disease**

Symptoms of bladder dysfunction have been reported in 37% to 72% of patients with Parkinson’s disease. These symptoms may be frequency or urgency (57%), obstruction (23%), or a combination of the two (20%). Detrusor overactivity has
been the most common urodynamic finding (72% to 100%) (49–51). Detrusor overactivity is thought to occur because of loss of the inhibitory input from the basal ganglia on the micturition reflexes; however, detrusor instability also has been associated with benign prostatic obstruction. Detrusor areflexia may result from bladder decompensation through a combination of bladder outlet obstruction and chronic use of anticholinergic and \(\alpha\)-adrenergic medications (51).

EMG studies of the external sphincter reveal that patients may have pseudodysynergia or bradykinesia but not true detrusor-sphincter dysynergia (49–51). The majority of patients (63% to 75%) have normal sphincter function (49–51).

**Multiple Sclerosis**

Only 6% of patients with multiple sclerosis first present with urologic symptoms (52). Bemelmans and associates, however, reported that 50% of asymptomatic patients with early multiple sclerosis had urodynamic abnormalities that needed further follow-up, and 50% of these required therapeutic intervention (53). As the disease progresses, urologic symptoms become common, eventually affecting at least 50% of men and 80% of women (54). The type of voiding dysfunction often is difficult to predict because of the diffuse involvement and changing nature of the disease.

Goldstein et al. reported that in a series of 86 symptomatic patients, 49% had incontinence, 32% had urgency and frequency, and 19% had obstructive hesitancy and retention. They also documented that patients with similar neurologic findings may have different voiding dysfunctions and that urologic signs and symptoms do not accurately reflect the voiding dysfunction (55). Wheeler et al. found that 55% of patients who were studied had changes in their urodynamic picture. The urodynamic pattern varied from detrusor areflexia to detrusor overactivity and vice versa (56).

Because suprapontine and suprasacral plaques occur most frequently, detrusor hyperreflexia is the most common urodynamic finding; however, as many as 50% of patients have poorly sustained uninhibited bladder contractions with inefficient bladder emptying. Detrusor areflexia is found in approximately 20% of patients with urologic symptoms. This is believed to be a result of sacral plaque involvement (57).

True detrusor-sphincter dysynergia may occur in multiple sclerosis when there is involvement of the suprasacral spinal cord. Approximately 15% to 20% of patients develop detrusor-sphincter dysynergia. Blaivas and Barbalias (58) reported that this was an ominous sign because of the potential for upper-tract damage and development of reflux as a result of the increased intravesical pressures needed to force urine past the dysynergic sphincter. Upper-tract pathologic processes, including pyelonephritis, renal calculi, reflux, and hydronephrosis, have been reported to occur in 10% to 20% of patients with multiple sclerosis (58,59).

**Suprasacral Spinal Cord Lesions**

Traumatic SCI is the most common suprasacral lesion affecting voiding. Other suprasacral lesions include transverse myelitis, multiple sclerosis, and primary or metastatic spinal cord tumor.

Patients with suprasacral spinal cord lesions would be expected to have detrusor hyperreflexia with detrusor-sphincter dysynergia. However, in cases of partial lesions, occult lesions of the sacral cord, or persistent spinal shock, this is not always the case (60).

Traumatic suprasacral SCI results in an initial period of spinal shock, in which there is hyperreflexia of the somatic system below the level of injury and detrusor areflexia. During this phase, the bladder has no contractions, even with various maneuvers such as water filling, bethanechol supersensitivity testing, or suprapubic tapping. The neurophysiology of spinal shock and its recovery are not known. Recovery of bladder function usually follows recovery of skeletal muscle reflexes. Uninhibited bladder contractions gradually return after 6 to 8 weeks (61).

Clinically, a person with a traumatic suprasacral SCI may begin having episodes of urinary incontinence and various visceral sensations, such as tingling, flushing, increased lower-extremity spasms, or autonomic dysreflexia with the onset of uninhibited contractions. As uninhibited bladder contractions become stronger, the postvoid residuals decrease. Rudy et al. (62) reported that voiding function appears optimal at 12 weeks postinjury. However, detrusor hyperreflexia has been reported to have a delayed onset of up to 22 months postinjury. Eventually, all of these patients did develop uninhibited contractions (63). Bors and Comarr (35) considered the bladder “balanced” when postvoid residuals were less than 20% of the total bladder capacity in those with detrusor hyperreflexia. Graham (64) reports that 50% to 70% of patients will develop balanced bladders without therapy. Unfortunately, high intravesical voiding pressures usually are required for the development of a balanced bladder. These high pressures may cause renal deterioration.

Traditionally, it has been thought that there is decreased activity of the external urethral sphincter during acute spinal shock. However, Downie and Awad (65) noted in dogs that with surgical transection between T2 and T8, there was no change in the activity of the periurethral striated musculature despite detrusor areflexia. In humans, Nanninga and Meyer (66) found that in 44 patients in spinal shock with suprasacral lesions, all had a positive bulbocavernosus reflex, and 30 of 32 had sphincter activity despite detrusor areflexia within 72 hours of injury. Koyanagi et al. noted that external sphincter electrical activity was not affected during acute spinal shock but was likely to increase after recovery from spinal shock. This increase was more marked in those with high suprasacral lesions than in those with low suprasacral lesions (67).

Detrusor–external sphincter dysynergia often occurs following suprasacral lesions. Blaivas et al. (68) noted that it occurred in 96% of patients with suprasacral lesions. They found several different patterns of striated sphincter dysynergia. Rudy et al. proposed that detrusor-sphincter dysynergia is an exaggerated continence reflex. The continence reflex is the normal phenomenon of increasing urethral sphincter
activity with bladder filling. They believed that the patterns described by Blaivas et al. (62) represented variations of the single continence reflex.

In addition to the detrusor–external sphincter dyssynergia, internal sphincter dyssynergia also has been reported, often occurring at the same time as detrusor–external sphincter dyssynergia.

**Sacral Lesions**

There are a variety of lesions that may affect the sacral cord or roots. These include spinal trauma, herniated lumbar disk, primary or metastatic tumors, myelodysplasia, arteriovenous malformation, lumbar stenosis, and inflammatory process (e.g., arachnoiditis). In Pavlakis et al. series, trauma was responsible for conus and cauda equina lesions more than 50% of the time. The next most common cause was L4/L5 or L5/S1 intervertebral disc protrusion. The incidence of lumbar disc prolapse causing cauda equina syndrome is between 1% and 15% (69). Damage to the sacral cord or roots generally results in a highly compliant acontractile bladder; however, particularly in patients with partial injuries, the areflexia may be accompanied by decreased bladder compliance, resulting in progressive increases in intravesical pressure with filling (70). The exact mechanism by which sacral parasympathetic decentralization of the bladder causes decreased compliance is unknown (70,71).

It has been noted that the external sphincter is not affected to the same extent as the detrusor. This is because the pelvic nerve innervation to the bladder usually arises one segment higher than the pudendal nerve innervation to the sphincter (72). The nuclei also are located in different portions of the sacral cord, with the detrusor nuclei located in the intermediolateral cell column and the pudendal nuclei located in the ventral gray matter. This combination of detrusor areflexia and an intact sphincter helps contribute to bladder overdistention and decompensation.

**Peripheral Lesions**

There are multiple etiologies for peripheral lesions that could affect voiding. The most common lesion is a peripheral neuropathy secondary to diabetes mellitus. Other peripheral neuropathies that have been associated with voiding dysfunction include chronic alcoholism, herpes zoster, Guillain-Barré syndrome, and pelvic surgery (73,74). A sensory neuropathy is the most frequent finding in diabetics. Urodynamic findings, including decreased bladder sensation, chronic bladder overdistention, increased postvoid residuals, and possible bladder decompensation, may result from bladder overdistention secondary to decreased sensation of fullness. Andersen and Bradley (75) reported that in their series, mean bladder capacity was 635 mL, with a range of 200 to 1,150 mL. An autonomic neuropathy also may be responsible for decreased bladder contractility. Guillain-Barré syndrome and herpes zoster are predominantly motor neuropathies. Transient voiding symptoms, predominantly urinary retention, have been reported to occur in 0% to 40% of patients and are thought to represent involvement of the autonomic sacral parasympathetic nerves. Detrusor hyperreflexia occasionally has been found in those with Guillain-Barré syndrome (76). Voiding dysfunctions resulting from pelvic surgery or pelvic trauma usually involve both motor and sensory innervation of the bladder (75).

**COMPREHENSIVE EVALUATION OF VOIDING DYSFUNCTION**

**Neurowurologic History**

The urological history should first focus on the patient’s voiding symptoms. The history should establish whether the onset of the current symptoms is new, has become worse, or has remained unchanged since the neurologic insult. This will allow for more meaningful discussions when patients and family ask about “returning to normal.” A preexisting problem such as urinary frequency may cause incontinence from decreased mobility.

Because symptoms often correlate poorly with the actual voiding problem, it is best not to initiate pharmacological or surgical treatment based solely on symptoms. Katz and Blaivas (77), in a prospective study of 425 consecutive patients, found that the clinical assessment based on symptoms did not correlate with the objective urodynamic findings in 45% of patients. They believed that urodynamic findings were predictive of the urodynamic diagnosis in only 25% of those believed to have storage problems, and in 54% of those believed to have storage and emptying problems. It has also been well documented that urodynamics should be included in the evaluation of those with SCI since bladder and sphincter function cannot be predicted by history and physical examination. Ouslander et al. (78) found in the geriatric female population that presenting symptoms were predictive of the urodynamic diagnosis in only 55% of those with pure urge incontinence. It has also been well documented that urodynamics is essential in the evaluation of those with SCI since bladder and sphincter function cannot be predicted by history and physical examination (79).

Significant past history includes additional medical problems that may contribute to present problems, such as diabetes, previous CVAs, hypertension, and use of diuretics. The past history also needs to focus on surgery that may affect voiding, such as previous transurethral resection of the prostate, surgery for stress incontinence, or pelvic surgery. Questions about past and present bowel and erectile function should be asked. Potentially reversible causes of voiding disorders need to be investigated. A helpful mnemonic coined by Resnick and Yalla (80) to describe the reversible causes of incontinence in the elderly is DIAAPPERS. These same factors also may be responsible for problems with retention. The mnemonic can be broken down as follows:

Delirium
Infection
Atrophic vaginitis, urethritis
Pharmaceuticals
Psychological
The physiatric history that has particular significance for voiding dysfunction is hand function, dressing skills, sitting balance, ability to perform transfers, and ability to ambulate. These factors not only play a role in why a person may be incontinent but also are important considerations in developing management strategies.

Neurourologic Examination

The neurourologic physical will not give objective evidence about the bladder and sphincter function but may suggest potential contributory causes of a voiding dysfunction. The examination should focus on the abdomen, external genitalia, and perineal skin. The rectal examination is important to evaluate for cancer. Prostate obstruction cannot be determined by prostate size alone since it is not the overall size of the prostate but the amount of prostate growing inward that causes obstruction. Therefore, a urodynamic study, which measures the actual pressure within the bladder during a contraction and the resulting flow of urine, is important to objectively diagnose outflow obstruction.

In postmenopausal women, the urethra and vaginal introitus should be examined for atrophic changes suggestive of estrogen deficiency. In women, the examination also should focus on the degree of pelvic support. A determination of masses producing extrinsic compression on the bladder should be made during the vaginal examination.

The mental status portion of the neurourologic examination should, as a minimum, evaluate the patient's level of consciousness, orientation, speech, long- and short-term memory, and comprehension. Voiding disorders may be secondary to or made worse by disorientation, inability to communicate the desire to void, or lack of understanding when the patient is told to void.

The sensory examination should focus on determining the level of injury in those with SCI. Especially important is establishing whether the level of injury is above T6, which would make the patient prone to autonomic dysreflexia. Sacral sensation evaluates the afferent limb (i.e., pudendal nerve) of the sacral micturition center. Loss of pinprick and light touch sensation in the hands and feet is suggestive of a peripheral neuropathy.

The motor examination helps establish the level of injury and degree of completeness in those with SCI. Hand function should be assessed to determine the ability to undress or possibly perform intermittent catheterization (IC). Upper- and lower-extremity spasticity with sitting, standing, and ambulating also needs to be evaluated. Anal sphincter tone is also important. Decreased or absent tone suggests a sacral or peripheral nerve lesion, whereas increased tone suggests a suprasacral lesion. Voluntary contraction of the anal sphincter tests sacral innervation, suprasacral integrity, and the ability to understand commands.

Cutaneous reflexes that are helpful to the neurourologic examination are the cremasteric (L1 to L2), bulbocavernous (S2 to S4), and anal reflex (S2 to S4). Absence of these cutaneous reflexes suggests pyramidal tract disease or a peripheral lesion. The bulbocavernosus reflex is a useful test to evaluate the sacral reflex arc. However, it may be unreliable. A false negative often results from a person being nervous and already having his or her anal sphincter clamped down at the time of the examination. Muscle stretch reflexes also should be evaluated. A sudden increase in spasticity may indicate a urinary tract infection (UTI). In addition, pathologic reflexes (e.g., Babinski reflex) may help localize the neurologic lesion.

Urologic Assessment of the Upper and Lower Urinary Tract

Indications for Testing

A variety of tests can be performed to evaluate the upper and lower urinary tract. The exact types of tests and follow-up depend on the disease process, the patient's clinical course, and any preexisting urologic problems needing further follow-up.

If the disease process is one that is not known generally to affect the upper tracts, such as a stroke, hip replacement with retention, or peripheral neuropathy, then the evaluation can be directed at the lower urinary tract. Evaluation of the upper tracts in these individuals should be undertaken if there is any suggestion of upper-tract involvement, such as pyelonephritis or hematuria.

Patients with disease processes that may affect the upper tracts, such as multiple sclerosis, should undergo baseline testing of the upper tracts and then periodic screening. Emphasis is directed primarily at the lower tract with the use of urine analysis, culture and sensitivity, postvoid residual, and urodynamics. Testing usually is done annually, but may be needed more or less frequently depending on the patient's clinical course.

Spinal cord–injured patients, particularly those with potential high intravesical voiding pressures, need constant surveillance of the upper tracts as well as lower tracts. Although there is no agreement on exactly which tests and the frequency at which testing should be done, there is agreement that upper- and lower-tract testing is necessary. The American Paraplegia Society has developed recommendations for the urologic evaluations of those with SCI (81).

Institutions often will have SCI patients undergo a yearly evaluation for the first 5 to 10 years, and if their upper tracts are stable, then evaluations are every other year. However, there is evidence that bladder function continues to change even after 20 years postinjury, suggesting that yearly evaluations should be considered (82). People with an indwelling suprapubic or Foley catheter often will get yearly cystoscopy to rule out stones and bladder tumors.

Specific Upper- and Lower-Tract Tests

Tests designed to evaluate the upper tracts include an intravenous pyelogram (IVP), renal ultrasound, 24-hour urine
creatinine clearance, and quantitative mercaptoacetyltriglycine (MAG 3) renal scan and computerized tomography (CT). When ordering a test one must consider whether the information is needed about the function of the upper tracts or about the anatomy of the upper tracts. For example, a renal ultrasound is excellent to detect anatomical changes; however, it will not suggest any problems with renal function until hydronephrosis develops. Conversely, a renal scan detects stasis of the upper tracts before hydronephrosis develops, but is unlikely to detect a small kidney stone.

A 24-hour urine creatinine clearance and quantitative renal scan evaluate upper-tract function, whereas renal ultrasound and CT are used to evaluate upper-tract anatomical features. An IVP evaluates both function and anatomy. Despite evaluating both function and anatomy, there are a number of disadvantages to an IVP. These disadvantages include potential allergic reactions, radiation exposure, and patient inconvenience, specifically getting an IVP laxative preparation the night before the test. Therefore, IVP is used infrequently.

The primary purpose of quantitative MAG 3 radioisotope renal scan is to monitor renal function and drainage. It has been found to be a safe and effective modality in those with SCI (83,84). While serum creatinine and creatinine clearance are easy to obtain and inexpensive, there are a number of possible problems when using them to monitor kidney function in those with SCI (85). Since individuals with SCI frequently have less muscle mass than able-bodied individuals, their serum creatinine should be less than that of able-bodied individuals. Therefore, a normal serum creatinine lab value may actually represent a high value for those with SCI. In addition, serum creatinine will not rise until there is at least a 50% decline in renal function. When ordering a 24-hour urine collection be aware that overcollection of urine (>24 hours) will overestimate kidney function and undercollection will underestimate kidney function (85).

The renal ultrasound is helpful for detecting hydronephrosis and kidney stones. When evaluating renal anatomy, ultrasound has largely replaced IVP (86). The major advantages of ultrasound are that it is noninvasive and does not involve any contrast agents. The major disadvantages of ultrasound are that it is user-dependent and does not show renal function (87).

If further anatomic definition is needed to evaluate for stones or tumors, CT should be considered. It has largely replaced IVPs in a number of institutions. In a prospective study of nonenhanced helical CT scans versus IVP, CT correctly identified 36 of 37 ureteral stones with one false positive. CT had a sensitivity 97%, specificity of 96%, and accuracy of 97% at detecting ureteric stones. This was double that of IVP (88).

Tests to evaluate the lower tracts include cystogram, cystoscopy, and urodynamics. Because each of these involves instrumentation, it is best to obtain a urine culture and sensitivity test, and give antibiotics if positive before the testing. An untreated infection or bacterial colonization has the potential to cause bacteriemia and increased bladder overactivity.

Some indications for cystoscopy in those with voiding disorders include hematuria, recurrent symptomatic UTIs, recurrent asymptomatic bacteriuria with a stone-forming organism (i.e., Proteus mirabilis), an episode of genitourinary sepsis, urinary retention or incontinence, pieces of eggshell calculi obtained when irrigating a catheter, and long-term indwelling catheter. Cystoscopy also is indicated when one is removing an indwelling Foley catheter that has been in place for more than 2 to 4 weeks or changing to a different type of management, such as IC or a balanced bladder. Cystoscopy can reveal a pubic hair or eggshell calculus that may be missed on radiography and serve as a nidus for bladder stones.

Urodynamics

Urodynamics provides objective information on voiding function (Table 51-2).

Urodynamics in general terms is defined as the study of normal and abnormal factors in the storage, transport, and emptying of urine from the bladder and urethra by any appropriate method (40). When deciding on an appropriate urodynamic test, one needs to consider whether information is needed about the filling phase, emptying phase, or both phases of micturition.

The following are some of the more common indications for an urodynamics evaluation:

- Recurrent UTIs in a patient with neurogenic bladder
- Urinary incontinence
- Urinary frequency
- Large postvoid residuals (i.e., retention)
- Deterioration of the upper tracts
- Monitoring of voiding pressures
- Evaluation and monitoring of pharmacotherapy

The physician's presence is important to help direct the urodynamics study. Typical decisions include how much water to put in the bladder, whether to repeat the study, and whether to have the patient sit or stand to void. Observing the patient during urodynamics also will help in getting an idea of factors that might influence the test, such as patient anxiety or inability to understand when told to void.

Blood pressure monitoring is particularly important in SCI patients prone to autonomic dysreflexia. Urodynamics is particularly helpful for detecting autonomic dysreflexia in men with SCI at T6. Autonomic dysreflexia may occur with bladder distention or more commonly when bladder distention provokes an uninhibited contraction. This causes the sphincter to contract, which causes a significant rise in blood pressure and other symptoms of autonomic dysreflexia. However, 43% of the men with SCI at T6 and above may have “silent dysreflexia” (elevated BP without any symptoms) during voiding. This would not be detected without urodynamics and simultaneous blood pressure monitoring (89).

In order to have an accurate urodynamic evaluation, it is important that a person does not have a UTI. Bladder wall inflammation is likely to cause the bladder to lose some of its compliance, resulting in a smaller bladder capacity than normal.
The inflammation is also likely to trigger uninhibited contractions and cause the bladder to be more overactive than usual. A recent prospective study found that 9.7% of SCI individuals who had asymptomatic bacteriuria developed a symptomatic UTI post testing. Nearly 40% of SCI individuals with sterile urine developed asymptomatic bacteriuria post testing (90).

It would be expected that one or two doses of an antibiotic would clear the bacteria from the urinary tract and reduce the risk of an infection. However, it has been our practice to obtain a urine culture and sensitivity 1 to 2 weeks prior to the test. Those with pyuria or a symptomatic UTI are treated for 5 days prior to testing with the goal not only to eradicate the bacteria in the bladder but also to give adequate time to reduce inflammation of the bladder wall. Those with sterile urine or asymptomatic bacteriuria are given one or two doses of an antibiotic prior to testing. A person who presents for testing with cloudy urine or other symptoms of a UTI is rescheduled.

**Evaluation of Bladder Filling (Storage Phase)**

The simplest type of bladder test to evaluate bladder filling is known as a bedside cystometrogram. This test involves attaching a cylinder such as 50 mL filling syringe without the plunger to a Foley catheter. Water is then poured into the cylinder and allowed to drain by gravity into the bladder. The blood pressure and volume of fluid going into the bladder are recorded. The Foley is sometimes attached by means of a Y-connector to a manometer, which is used to measure the actual rise in water pressure. This test can be used to evaluate sensation (whether or not a person is aware of the bladder being filled), stability (whether or not there is a rise in the column of water signifying a bladder contraction), and capacity (the volume at which the bladder contraction occurs). It can also be used as a screening test to determine if an SCI patient has come out of spinal shock. There are several limitations to the bedside cystometrogram, however. It is difficult to determine if small rises in the water column result from intra-abdominal pressure (i.e., straining) or a bladder contraction. An iatrogenic bladder contraction can be elicited if the tip of the Foley catheter rubs against the trigone pressure sensors, which can then trigger bladder contractions. Most important, the voiding phase cannot be evaluated (91).

The carbon dioxide urodynamics has been largely replaced with water-fill urodynamics. Although the gas is cleaner and neater to use than water, the major disadvantage is that the voiding phase of micturition cannot be evaluated. Therefore, this test is of little use when trying to evaluate bladder and sphincter function during emptying.

**Evaluation of Bladder Emptying**

One of the easiest screening tests to evaluate bladder emptying is a postvoid residual; however, it should not be used to characterize the specific type of voiding dysfunction. The postvoid residual can be determined with catheterization or bladder ultrasound. A younger person should have no postvoid residual; however, an elderly person with no voiding symptoms may have a postvoid residual of 100 to 150 mL. A normal postvoid residual does not rule out a voiding problem. For example, a postvoid residual may be normal despite significant outflow obstruction (e.g., benign prostate hypertrophy, sphincter-detrusor dyssynergia) as a result of a compensatory increase in the strength of detrusor contractions or of absent bladder contractions in the presence of increasing intra-abdominal pressure (e.g., Valsalva maneuver, Crede maneuver). Caution also has to be taken in interpreting a large postvoid residual. It may be abnormal because it was not taken immediately after voiding, because of poor patient understanding, or because of an abnormal voiding situation (e.g., the patient was given a bedpan at 2:00 A.M.).

A multichannel water-fill urodynamics study is the gold standard to evaluate bladder function because it measures both the filling and the emptying phase of micturition. Multichannel refers to the fact that each of the various urodynamics parameters is measured as a separate channel such as detrusor pressure, abdominal pressure, and flow rate. Urodynamic studies also may incorporate urethral pressure recordings, urethral sphincter or anal sphincter EMG, videofluoroscopy, and the use of various pharmacologic agents, such as bethanechol (Fig. 51-4).

**FIGURE 51-4.** Waterfill urodynamics setup. Simultaneous monitoring of various urodynamics parameters is shown. Intravesical pressure minus intra-abdominal pressure will produce the detrusor pressure ($P_{\text{det}}$). A: Intravesical pressure, $P_{\text{ves}}$. B: Urethral sphincter pressure, $P_{\text{ur}}$. C: Urethral sphincter electromyography. D: Intra-abdominal pressure, $P_{\text{abd}}$. E: Urine flow rate.
Multichannel Water-fill Urodynamic Study

A water-fill urodynamic study evaluates two distinct phases of bladder function. The first is the filling (storage) phase, during which water is being infused into the bladder. Urodynamic parameters that can be evaluated during this phase include bladder sensation, bladder capacity, bladder wall compliance, and bladder stability (whether or not there are uninhibited contractions). The second portion of the study is the voiding (emptying) phase. The voiding phase is considered to begin when a person is told to void. In those who have neurogenic bladders and reflexly void, the voiding phase is considered to begin when the person has an uninhibited contraction and voiding begins. Urodynamic parameters that can be evaluated during the voiding phase include opening or leak-point pressure (bladder pressure at which voiding begins), maximum voiding pressure, urethral sphincter activity (EMG or actual pressure), flow rate, voided volume, and postvoid residual. In those who have the potential for autonomic dysreflexia, changes in blood pressure before, during, and after voiding can also be evaluated.

With an empty bladder, there should be no sensation of fluid within the bladder. During the filling phase, the first sensation that a person has of having a full bladder (first sensation of fullness) usually occurs with 100 to 200 mL within the bladder. The sensation of fullness occurs around 300 to 400 mL, and the onset of urgency usually occurs between 400 and 500 mL. There is, however, variability in bladder capacity, which ranges between 400 and 750 mL in adults. There should be little to no rise in the intravesical pressure, which indicates normal bladder wall compliance. Additionally, there should be no involuntary bladder contractions during this part of the study.

During the voiding phase, the detrusor pressures usually are less than 30 cm H$_2$O in women and between 30 and 50 cm H$_2$O in men. A normal maximum flow rate is 15 to 20 mL/s and should not be less than 10 mL/s in any age group. The patient should have at least 150 mL in the bladder because the flow rate depends on the voided volume (44). The flow usually has a bell-shaped curve, progressively increasing to its maximum rate and then decreasing. The urethral sphincter should remain open throughout voiding, and there should be no rises in intra-abdominal pressure during voiding. As previously discussed, there should be no postvoid residual, although postvoid residuals increase with age.

A single elevated postvoid residual during urodynamics should be interpreted with caution because the patient may be nervous and voluntarily stop the urine stream. Several catheterized or ultrasound postvoid residual tests should be done to confirm an increased urodynamic postvoid residual (Fig. 51-5). Urodynamics is able to characterize specific types of voiding patterns (Fig. 51-6).

Special Considerations in Children

At one time, urodynamic evaluation was delayed until a child was school-aged and definitive corrective surgery was to be performed. However, reflux and renal deterioration often occur during the first 3 years of life. McGuire and colleagues reported that there was a high incidence of renal deterioration in patients with urethral leak-point pressures greater than 40 cm H$_2$O (92). Therefore, it is recommended that all myelodysplastic newborn children be evaluated as soon as possible (93).

It is difficult to obtain high-quality water-fill urodynamic studies on children younger than 4 or 5 years old. In younger children, it sometimes is necessary to use sedation or general anesthesia. It is important that children feel comfortable with the physician, nurses, and test. As a general principle, the amount of additional information gained from insertion of EMG needles usually is not enough to warrant the risk of obtaining poor urodynamics results from a crying, fearful child. This is especially true if it is anticipated that the child will come back for follow-up studies.

MANAGEMENT OF VOIDING DYSFUNCTIONS

A useful way to organize management of voiding dysfunctions is to base treatment options on a modification of the Wein classification: incontinence caused by (a) the bladder...
or (b) the outlet (e.g., bladder neck, sphincter, or prostate) or retention caused by (a) the bladder or (b) the outlet. Management can be categorized as behavioral, pharmacological, surgical, or supportive. Table 51-3 shows various treatment options. Although each of these modalities is listed separately, it is important to note combinations within the same category (e.g., behavioral—timed voiding, fluid restriction, biofeedback) and separate categories (e.g., behavioral—timed voiding, pharmacologic—anticholinergics).

It is important to characterize the type of voiding dysfunction that a person has with an urodynamics evaluation, particularly when considering pharmacological and surgical options (see Table 51-3). In addition to the type of voiding dysfunction, the physician needs to consider the type of disease process (i.e., progressive, stable, or remitting), cognition, mobility, family support, and medical conditions when recommending a bladder-management program. Empirical pharmacotherapy should be discouraged because there is a risk of potential side effects of drugs that may have no benefit or make the problem worse.

The following are goals of management in patients with voiding dysfunctions:

- Prevent upper-tract complications (e.g., deterioration of renal function, hydronephrosis, renal calculi, pyelonephritis).

## Table 51.3 Treatment Options for Voiding Disorders

<table>
<thead>
<tr>
<th>Incontinence Caused by the Bladder</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Behavioral:</strong> Scheduled (timed) voiding, limited fluid intake, biofeedback</td>
</tr>
<tr>
<td><strong>Pharmacologic:</strong> Oral anticholinergics, antispasmodics, tricyclic antidepressants, ODAVP (Vasopressin), intravesical oxybutynin, intravesical C-fiber afferent neuromodulators (capsaicin, resiniferatoxin), calcium antagonists, prostaglandin inhibitors</td>
</tr>
<tr>
<td><strong>Surgical:</strong> Augmentation cystoplasty, urinary diversion, interruption of innervation, neurostimulation, botulinum toxin injections into the bladder wall</td>
</tr>
<tr>
<td><strong>Supportive:</strong> Diapers, external condom catheter, intermittent catheterization, indwelling catheter</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Incontinence Caused by the Sphincter</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Behavioral:</strong> Scheduled voiding, pelvic floor exercises, biofeedback</td>
</tr>
<tr>
<td><strong>Pharmacologic:</strong> Alpha-adrenergic agonists, estrogen, injectable periurethral bulking agent</td>
</tr>
<tr>
<td><strong>Surgical:</strong> Artificial sphincter, urethral suspension, neurostimulation</td>
</tr>
<tr>
<td><strong>Supportive:</strong> Same as with bladder</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Retention Caused by the Bladder</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Behavioral:</strong> Scheduled voiding (↓ cognition/mobility), suprapubic tapping (bladder hypocontractility), Valsalva, Credé</td>
</tr>
<tr>
<td><strong>Pharmacologic:</strong> Cholinergic agonists, intravesical prostaglandin, narcotic antagonists</td>
</tr>
<tr>
<td><strong>Surgical:</strong> Spincterotomy, neurostimulation (if bladder contraction present)</td>
</tr>
<tr>
<td><strong>Supportive:</strong> Intermittent catheterization, indwelling catheter</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Retention Caused by the Sphincter/Outlet</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Behavioral:</strong> Biofeedback, suprapubic tapping, anal stretch/scissoring</td>
</tr>
<tr>
<td><strong>Pharmacologic:</strong> Alpha-adrenergic blockers, baclofen, diazepam, dantrolene</td>
</tr>
<tr>
<td><strong>Surgical:</strong> Spincterotomy, botulinum toxin injections, pudendal neurectomy, bladder outlet surgery, urethral stents, balloon dilation</td>
</tr>
<tr>
<td><strong>Supportive:</strong> Same as with bladder</td>
</tr>
</tbody>
</table>

Investigational use.
Prevent lower-tract complications (e.g., cystitis, bladder stones, vesicouretal reflux).

Develop a bladder management program that will allow patient to reintegrate most easily back into the community.

**Therapy for Incontinence Caused by the Bladder**

**Behavioral Treatment Options**

Many patients with incontinence caused by the bladder benefit from a scheduled (timed) voiding regimen. Patients with incontinence resulting from poor cognition, aphasia, or poor mobility but normal bladder function often are also helped by being placed on a commode or offered a urinal at set intervals. This is known as timed voiding. Patients who have an overactive detrusor may decrease incontinent episodes by voiding by the clock rather than waiting for a sense of fullness. They are taught to void before reaching their full bladder capacity because uninhibited contractions often become more forceful and frequent as the bladder is reaching its full capacity. Once the contraction begins it is very difficult to get undressed and go to the bathroom in time.

Another type of behavioral intervention is bladder training. This is done by progressively increasing the time between voiding by 10 to 15 minutes every 2 to 5 days until a reasonable interval between voiding is obtained (94). Bladder training often is effective for a person who has recovered or is recovering from a neurologic lesion (e.g., head injury, stroke) with improved bladder function but is voiding frequently out of habit or from fear of incontinence based on past experience.

**Pharmacological Treatment Options**

Pharmacologic treatment often is needed in addition to timed voiding in patients with incontinence caused by detrusor overactivity. There are currently a wide variety of medications available, which have anticholinergic effects. If a person does not tolerate one type of anticholinergic, he or she may tolerate another type of anticholinergic medication. Anticholinergic agent’s primary action is to block acetylcholine receptors competitively at the postganglionic autonomic receptor sites. Some agents, such as oxybutynin, also have a localized smooth-muscle antispasmodic effect distal to the cholinergic receptor site and a local anesthetic effect on the bladder wall (95).

Some of the more common potential side effects of anticholinergic medications include dry mouth, pupillary dilatation and blurred vision, tachycardia, drowsiness, and constipation from decreased gastrointestinal motility. Newer anticholinergic agents have been developed to be more selective to the bladder receptors, or have a slow sustained release or topical patch developed in an attempt to lessen anticholinergic side effects, particularly dry mouth and constipation. It has not been shown whether these agents maintain their effectiveness for an entire 24-hour period. Twenty-four-hour effectiveness is important in those with SCI at or above T6 who have the potential to develop autonomic dysreflexia if their medications “wear off” and they begin to develop uninhibited contractions. In those with neurogenic bladders the object is frequently to completely shut down the bladder and cause retention so IC can be performed. Therefore, those with neurogenic detrusor overactivity often require more than the “standard” doses used for able-bodied individuals.

Tricyclic antidepressants sometimes are used alone or in combination with anticholinergic agents. These medications are thought to have a peripheral anticholinergic effect and a central effect. They have been found to suppress uninhibited bladder contractions, increase bladder capacity, and increase urethral resistance (96). There have been several reports of severe autonomic dysreflexia in SCI patients secondary to overdistention of the bladder with urine. Therefore, caution should be taken in giving these medications that depend on uninhibited contractions to void (reflex voiders).

Intravesical instillation of medications is sometimes used because oral anticholinergic medications have a number of side effects. The major advantage of intravesical medications is that there are minimal to no systemic side effects due to less systemic absorption from the bladder wall. This is particularly helpful for those with a neurogenic bowel, because anticholinergic medications frequently cause constipation that could lead to fecal impaction. Anticholinergic medications may also cause a dry mouth, which can be particularly difficult to tolerate for those trying to limit their fluids because of being on IC.

Intravesical lidocaine has been shown to be effective at suppressing uninhibited bladder contractions in those with overactive bladders (97). Although it has a rapid onset of action, it does not have a long duration of action. Therefore, this medication is best reserved for use in acute problems. For more long-term suppression of uninhibited bladder contractions, oxybutynin can be used for intravesical instillation (98).

It not only has an anticholinergic effect but also a topical anesthetic effect. This medication is effective at suppressing uninhibited bladder contractions, but it still has the disadvantage of only being effective for 4 to 6 hours. One study reported using 5 to 10 mg dissolved in 15 to 30 mL of normal saline instilled into the bladder three to four times a day; this dose noted in the seven men resulted in an improvement in body image and enhanced sexuality because of the significant improvement in incontinence (99). Another investigator examined 32 patients, comparing standard dosages of intravesical oxybutynin (0.3 mg/kg body weight per day) with increasing dosages in steps of 0.2 mg/kg body weight up to 0.9 mg/kg body weight per day. Twenty-one of thirty-two (66%) patients became continent with the standard dose. Seven of the eleven failures at the lower dose became continent with a median dose of 0.7 mg/kg body weight for an overall success rate of 28 of 32 (87%). Four of the eleven (12.5%) had no improvement, and two of the eleven patients had side effects with a dosage of 0.9 mg/kg body weight per day (100). Despite its effectiveness, a number of individuals abandon using intravesical instillations because it is so labor intensive.

Intravesical instillations may, however, assume a more important role of helping to control uninhibited contractions with the development of longer-acting agents. Of particular interest is afferent C-fiber neurotoxins. The prototype
medication is capsaicin, which is effective at suppressing uninhibited contractions for several months at a time (101). Unfortunately, capsaicin frequently causes discomfort or suprapubic pain, urgency, hematuria, and autonomic dysreflexia, which can last to up to 2 weeks postinstillation.

A newer afferent C-fiber neurotoxin, resiniferatoxin, is being investigated. It is 1,000 times stronger than capsaicin and is long-acting. It has an extremely rapid onset of action at desensitizing the C-fiber afferent neurons, which causes minimal discomfort when it is instilled. In a study, 14 patients with detrusor hyperreflexia were instilled with 100 mL (or the bladder capacity if lower than that volume) of 50 to 100 nm resiniferatoxin instillation in 10% alcohol in saline. Treatment improved or abolished incontinence in 9 of 12 (75%) patients. Mean cystometric capacity increased from 182 to 330 mL. Maximal detrusor pressure was not modified by treatment. The effects were long-lasting, up to 12 months in seven patients (102). It is our understanding that Food and Drug Administration (FDA) multicenter trials evaluating resiniferatoxin in the United States have been stopped due to funding issues.

Desmopressin acetate has been found to decrease the number of episodes of nocturia in patients with multiple sclerosis. However, further studies are needed to determine its usefulness in the elderly because of the high prevalence of contraindications such as renal insufficiency, heart failure, and risks of inducing hyponatremia and fluid retention (29).

Another modality that has been gaining increasing popularity to help suppress uninhibited contractions in those with an overactive bladder is botulinum-A toxin injected into the bladder wall. While there have been encouraging data from early open-labeled and randomized, controlled trials with regards to efficacy and tolerability in both nonneurogenic and neurogenic bladder overactivity, the most experience has been with neurogenic bladder overactivity (103–105). Botulinum toxin inhibits acetylcholine release at the neuromuscular junction, which in turn blocks neuromuscular contraction and relaxes muscles that are either spastic or overactive. Doses ranging between 100 to 300 units have been confirmed by cystometry to suppress an overactive detrusor (103–105). Because it may take 1 to 4 weeks to completely deplete the acetylcholine at the neuromuscular junction, maximal effects at quieting the bladder do not occur immediately. Since there is reinnervation and sprouting at the neuromuscular junctions, the effects usually wear off after 3 to 6 months, so the injections usually need to be repeated. It appears that injections into the bladder wall may last closer to 6 months. Recent systematic reviews confirmed that botulinum toxin injections into the detrusor provide a clinically significant improvement in adults with neurogenic detrusor overactivity and incontinence refractory to other pharmacological therapy. It was well tolerated. The reviews point out that more studies are needed to evaluate issues such as the optimal dose, number and location of injections, and duration of effect (103–105). The panel for the SCI Consortium Guidelines on Management For Adults with SCI listed the use of botulinum injections into the bladder (detrusor) muscle as a bladder management option for those with SCI on IC with detrusor overactivity (106). Botulinum toxin to treat detrusor-sphincter dyssynergia is discussed in the pharmacologic agent section “Therapy for Retention Caused by the Outlet or Sphincter.”

In summary, there is a wide variety of pharmacologic agents that may be used by themselves or in conjunction with treatment modalities. A number of pharmacological agents discussed in this section as well as in the following sections are still undergoing investigation or “off-label” use. One reason is that individuals with “neurogenic bladders” are frequently excluded from initial FDA trials or do not make up a large enough population to have drug companies initiate trials in those with SCI. When using any pharmacologic agents for treatment, potential side effects and contraindications must be weighed against potential benefits.

Surgical Treatment Options
Bladder Augmentation

Bladder augmentation is a surgical technique that is frequently used to create a large bladder capacity with low intravesical pressures. Because this is a surgical procedure, other alternatives, such as pharmacologic treatment, should first be tried.

Surgical treatment is sometimes needed to improve bladder capacity in adult patients who are incontinent and want to perform IC. The SCI Consortium Bladder Management Panel recommended that bladder augmentation be considered for individuals with SCI who have (a) intractable involuntary bladder contractions causing incontinence, (b) the ability and motivation to perform IC, (c) the desire to convert from reflex voiding to an IC program, (d) a high risk for upper-tract deterioration secondary to hydronephrosis and/or ureterovesical reflux as a result of high-pressure detrusor-sphincter dyssynergia (106). They also recommended to consider avoiding bladder augmentation in individuals with (a) inflammatory bowel disease, (b) pelvic irradiation, (c) severe abdominal adhesions from previous surgery, and (d) compromised renal function (106).

An extensive preoperative evaluation is important. The history should include questions about any gastrointestinal problems. Urodynamics should be done to evaluate bladder and sphincter function. Various treatments may be needed to treat the sphincter if there is a low leak point pressure. Screening laboratory work should include liver and renal function. To help reduce the risk of significant acidosis and metabolic abnormalities, bladder augmentation is best reserved for those with serum creatinine less than 2.0 mg/dL. A cystogram should be done to evaluate for vesicoureteral reflux. Ureteral reimplantation may be considered if there is significant reflux. Upper-tract evaluation is important both to rule out any problems and to serve as a baseline for follow-up post augmentation (107).

There are a number of different techniques of bladder augmentation in which different segments of bowel can be used. The most common type of bladder augmentation is the clam cystoplasty. This procedure involves isolating a piece of
intestine, being careful to keep it attached to its mesentery, detubularizing it, and sewing it onto the bladder, which is first partially bivalved. Various bowel segments can be used and depend on the surgeon’s preference.

There are predictable metabolic abnormalities depending on the segment being used. The stomach mucosa has secretory epithelium with little absorptive function. Gastric mucosa secretes hydrochloric acid in conjunction with systemic bicarbonate release. Therefore, hypochloremic metabolic alkalosis can result if the stomach is being used, particularly if there is poor renal function. However, the stomach has the least absorptive properties and is best if one is concerned about metabolic acidosis from reabsorption of urinary solutes through the bowel wall. However, this is technically more difficult than an intestinal segment closer to the bladder.

The jejunal mucosa is different from the ileum and large intestine in that it secretes sodium and chloride and may result in hyponatremia, hypochloremia, and hyperkalemia. This segment is most likely to result in metabolic abnormalities and rarely used in diversions. The ileum and colon have similar transport mechanisms. Ammonia and chloride are reabsorbed. This can lead to hyperchloremic metabolic acidosis.

The most frequent changes that occur after this are an increase in mucus noted in the urine, possible metabolic changes, abnormal drug absorption (especially those that are absorbed by the gastrointestinal tract and excreted unchanged by the kidneys, such as Dilantin and certain antibiotics), osteomalacia from chronic acidosis, and stones, particularly in those with urea-splitting organisms and hyperchloremic metabolic acidosis. Long-term consequences of bowel attached to bladder are unknown. There have also been case reports of cancer in those with bladder augmentations, ileal conduits, and colon conduits. These have been adenocarcinomas, undifferentiated carcinomas, sarcomas, and transitional cell carcinomas (108–114). There have been a few reports of patients who had an augmentation cystoplasty more than 10 years previously developing adenocarcinoma in the bladder (110).

Mast and associates reported a 70% success rate at stopping incontinence (mean length of follow-up 1.5 years), and this was increased to 85% with the addition of an artificial sphincter in those with low sphincter resistance (111).

Complications included recurrent UTI (59%) and stone formation (22%). As a result of complications, however, further surgery was required for 44% of the patients. Another study evaluated clinical outcome and quality of life after enterocystoplasty in 18 patients with neurogenic bladders and 3 with contracted bladders because of radiation cystitis (112). Enterocystoplasty was performed using a 40-cm segment of terminal ileum. Mean bladder capacities improved from 1 to 6 months. One patient had a bladder rupture and another required surgery. The panel recommended that there should be caution considering urinary diversions with in SCI individuals who are (a) too debilitated to undergo a major surgical procedure or (b) have one of the following conditions: inflammatory bowel disease, pelvic irradiation, severe abdominal adhesions from previous surgery, or (c) compromised renal function (106).

**Bladder Autoaugmentation**

Another method of surgically increasing bladder size without mucus formation and use of a bowel segment is a bladder autoaugmentation. This is also called a detrusor myomectomy. Through an abdominal approach, the bladder muscle is stripped away from the inner mucosal lining. Without this muscle lining, the bladder mucosa gradually stretches to become a large diverticulum, thus increasing bladder capacity. Autoaugmentation has the advantage of not causing metabolic and absorption problems as described previously for bladder augmentation using a bowel segment. However, this procedure is often technically difficult, particularly in those with a neurogenic bladder who frequently have a small, a heavily trabeculated bladder, and it offers only an approximate 25% increase in bladder capacity. The capacity does not immediately increase, but gradually increases over time. One investigator reported on bladder autoaugmentation of 50 men with neurogenic bladders. Bladder capacity increased over a period of 1 to 6 months. One patient had a bladder rupture and two were reported to have “failed” because of psychological reasons (113).

**Urinary Diversions**

Another method of management is urinary diversion. Although there are many types of diversions, they can be grouped into two types: standard (noncontinent) and continent diversion. Preoperative workup should be the same as that for bladder augmentation.

Noncontinent urinary diversion may be used as an alternative to augmentation cystoplasty or continent diversion when hand function does not permit self-catheterization. The most common standard noncontinent diversion is an ileal conduit. Ten to fifteen centimeters of ileum, along with its mesentery, is isolated from the ileum. The isolated segment of ileum is closed off at one end, and the other end is brought out through the abdominal wall and everted as a nipple stoma. The ureters are implanted onto the side of the ureters.

The SCI Consortium bladder panel recommended that urinary diversion be considered for individuals with SCI who have (a) lower urinary complications secondary to indwelling catheters, (b) urethrocystaneous fistulas, (c) perineal pressure ulcers, (d) urethral destruction in females, (e) hydrenephrosis secondary to a thickened bladder wall, (f) hydrenephrosis secondary to vesicoureteral reflux or failed reimplant, and (g) bladder malignancy requiring a cystectomy. The most common reason for noncontinent urinary diversions in those with SCI is to divert the urinary stream away from the perineum because of impaired healing of a decubitus because of urinary incontinence, or a urethral stricture or fistula. The panel recommended that there should be caution considering urinary diversions with in SCI individuals who are (a) too debilitated to undergo a major surgical procedure or (b) have one of the following conditions: inflammatory bowel disease, pelvic irradiation, severe abdominal adhesions from previous surgery, or (c) compromised renal function (106). Continent diversions may be used in those who need to have their urinary stream diverted and who have adequate hand function to perform IC. Continent diversions are divided into two types. The first are orthotopic diversions, in which the
bowel reservoir is anastomosed to the urethra, and the second is a continent catheterizable pouch.

Orthotopic diversions are much like bladder augmentations and used to increase bladder capacity. Because they are attached to the urethra, people can catheterize themselves through their urethra in the same manner that they would if they were catheterizing their bladder. Continent catheterizable pouches have the advantage that the stoma can be placed in a location that makes catheterization easier. For example, the stoma can be created within the umbilicus so that a person does not have to undress to catheterize himself or herself. The most difficult part of the continent diversion is the creation of a continent mechanism.

The SCI Consortium panel recommended that a continent urinary diversion be considered in individuals with SCI; (a) in whom it is not feasible to augment the native bladder, (b) who cannot access their native urethra because of congenital abnormalities, (c) who have spasticity (making it difficult to catheterize their urethra), obesity, contracture, or tetraplegia, or who require closure of an incompetent bladder neck, (d) females with tetraplegia in whom a chronic indwelling catheter has caused urethral erosion, (e) males with SCI with unsalvageable bladders secondary to urethral fistula and sacral pressure ulcers, and (f) individuals with bladder cancer requiring cystectomy (106).

The most commonly used bowel segment for this is the ileocecal valve. The right colon without a segment of small bowel to increase volume is used for the pouch, and the terminal ileum is used to create the catheterizable limb (107,108).

Postoperatively patients need to catheterize their pouch frequently to prevent rupture. Because there is an increase in mucus, patients need to be taught how to irrigate their pouch. Irrigations can be decreased over time, but irrigation is recommended at least once a month as there may be malabsorption of bile salts because of the use of the ileal cecal valve. This increase in bile salts in the colon may cause diarrhea. This is best treated with oral administration of cholestyramine (114).

Long-term follow-up of bladder augmentations and urinary diversions includes regular monitoring of the upper tracts and careful monitoring of blood chemistries and renal function. Cystoscopy is used to monitor for stones or tumors (107,108).

Surgical methods also have been designed to interrupt innervation to the bladder. This can be done centrally (e.g., subarachnoid block, cordectomy), peripherally (e.g., anterior or anteroposterior rhizotomy), or peripherically (e.g., extensive mobilization of the bladder) (115–117). Although there usually is a successful short-term outcome, decreased compliance or detrusor hyperreflexia may return. This may result from an increased sensitivity of receptors following decentralization (24). Impotence usually occurs after these procedures.

**Neurostimulation**

Neurostimulation is a relatively new area of research. Electrical stimulation has a number of uses in treating those with voiding dysfunction. It has been used both to facilitate storage by decreasing uninhibited bladder contractions and to improve voiding by helping to trigger uninhibited contractions. Neurostimulation to inhibit bladder contractions has had its widest use in able-bodied individuals with overactive bladders. Ohlsson and Frankenberg-Sommar (118) reported an average 49% increase in bladder capacity by stimulation of the pudendal nerve with anal and vaginal electrode plugs. Tanago (119) reported success at selective sacral root stimulation to increase sphincter tone, which in turn suppresses detrusor activity.

In those with neurogenic bladders, investigators continue to try to improve voiding through the use of neurostimulation. Techniques include placing electrodes on the bladder itself, the pelvic nerves, conus medullaris, sacral nerves, and the sacral anterior roots. Of these, sacral anterior root stimulation has been most successful.

Brindley et al. (120) developed an anterior root stimulator that produces micturition by stimulating the sacral nerve roots. The largest experience is with surgically implanted Finetech-Brindley sacral afferent stimulator; there have been an estimated 800 implants over 15 years.

It is important to stress to patients that sacral stimulation is not indicated unless an individual has an intact sacral reflex arc and a bladder capable of having bladder contractions (106). Stimulation of the sacral afferent nerves causes reflex activation of the efferent nerves to the sphincter. However, this reflex accommodates so that fatigue of the sphincter occurs and the pressure generated in the urethra is overcome by the bladder contraction. A posterior rhizotomy is often performed at the same time as the sacral implant to abolish uninhibited bladder contractions, abolish contractions of the sphincter, and improve bladder wall compliance. The disadvantage of the posterior rhizotomy is the loss of reflex erections and reflex ejaculations, loss of perineal sensation, and loss of reflex bladder contractions (106,121). Van-Kerrebroeck reviewed the worldwide experience with the Finetech-Brindley sacral stimulator. In 184 cases, of which 170 were using the stimulator, 95% had postvoid residuals less than 60 mL. There was no deterioration of the upper tracts. Two thirds of men reported stimulated erections, but only one third used these for coitus (122). Sacral stimulation is further discussed in Chapter 65.

**Supportive Treatment Options**

**Protective Undergarments**

Protective undergarments are often helpful either as primary management or as backup to management for timed voiding. Custom-fit diapers are much more cosmetic than in the past. Major drawbacks include expense, patient embarrassment, difficulty getting them on and off, and potential skin breakdowns if they are not changed within 2 to 4 hours after getting wet. Insurance companies often do not reimburse for undergarments.

**External Condom Catheters**

External condom catheters often are a good option for men with detrusor overactivity or normal bladder function with
incontinence secondary to mobility or cognitive factors. An external condom catheter is not recommended unless the bladder is able to contract and empty with a minimal postvoid residual (with or without treatment of the sphincter—see section Management of Urinary Retention Due to the Sphincter). An advantage of an external condom catheter over protective undergarments is that condom catheters need to be changed only once a day and are often covered by insurance plans. Major drawbacks include wearing of a leg bag, potential for penile skin breakdown, condom catheter falling off, and slight increase in bladder infections compared to those who do not wear an external condom catheter. A reliable external collecting device for women has not been developed.

**Intermittent Catheterization**

IC, usually also requiring pharmacological or surgical intervention, is another effective way to manage patients with detrusor overactivity and incontinence. Gutman and Frankel popularized sterile IC for SCI patients in the 1960s. They reported that in 476 SCI patients monitored over 11 years on sterile IC, only 7.4% developed hydrourephrosis, 4.4% had vesicoureteral reflux, 1.7% developed kidney stones, and 0.6% developed bladder stones (123). In the mid-1970s, Lapides et al. reported on the effectiveness of IC using clean technique in those with SCI. They attributed the success of IC to keeping the bladder from getting overstretched. One reason for this was the ease of performing clean IC compared with sterile technique (requiring gloves and somewhat bulky catheter kit at the time when these studies were done) so that patients were more likely to catheterize themselves and prevent bladder overdistention than with sterile technique (124). Maynard and Glass (125) reported that 80% of patients on clean IC monitored for 60 months continued this technique, suggesting low morbidity and high patient acceptance. Since these studies, it should be noted that sterile IC has become simplified using “a sterile catheter in a bag” so there is probably no significant difference between sterile and clean “IC.”

The important principles of IC are to restrict fluids to 2 L/day and to catheterize frequently enough to keep the bladder from becoming overdistended (<500 mL). Bladder overdistension is a major cause of bladder infections.

The SCI Consortium bladder panel recommends to consider avoiding in individuals with SCI who have one or more of the following: (a) inability to catheterize themselves or a caregiver who is unable to perform catheterization, (b) abnormal urethral anatomy such as stricture, false passages, and bladder neck obstruction, (c) bladder capacity less than 200 mL, (d) poor cognition, little motivation, and inability or unwillingness to adhere to the catheterization time schedule or the fluid intake regimen, or (e) adverse reaction toward having to pass the catheter into the genital area multiple times a day, and (f) tendency to develop autonomic dysreflexia with bladder filling despite treatment (106).

Use of prophylactic antibiotics is discouraged for most individuals on IC. Antibacterial prophylaxis has been shown to significantly reduce the probability of laboratory infection but not the probability of clinical infection, although a trend was noted toward fewer clinical infections (125). In addition to the cost and potential side effects of antibiotics, a major concern of prophylactic antibiotics is the development of resistant organisms.

**Indwelling Urethral Catheterization**

With shorter lengths of hospital stays, many people with SCI are discharged from the hospital while their bladder is still in “spinal shock.” Those with other rehabilitation diagnosis such as a stroke or hip replacement may have poor mobility or difficulty with dressing, which may cause urinary incontinence because of the inability to get to the bathroom in time. It has been recommended that indwelling catheterization be considered for individuals with SCI who have one or more of the following: (a) poor hand skills, (b) high fluid intake, (c) cognitive impairment or active substance abuse, (d) elevated detrusor pressures, (e) lack of success with other less invasive bladder management methods, (f) need for temporary management of vesicoureteral reflux, and/or (g) limited assistance from a caregiver, making another type of bladder management not feasible (106).

A major advantage of having an indwelling catheter is the ability to have more independence. In SCI individuals discharged home on IC, the most common reason for switching to an indwelling catheter is the inability at being independent at performing IC. One study showed that 80% of those with tetraplegia prefer an indwelling catheter for social and practical methods (126). Because there is no satisfactory external collecting device for women, an indwelling catheter is particularly useful in women with poor hand function who have urinary retention and/or urinary incontinence.

Those without SCI who are not independent at transfers and dressing and their family members also often benefit having in an indwelling catheter when initially discharged from the acute care hospital. The indwelling catheter allows the family to deal with one less issue as the person is recovering and getting stronger. A common misconception is that the indwelling catheter causes a person to develop a “lazy bladder” or forget how to void. The opposite is actually true. The indwelling catheter increases a person’s bladder tone, especially if he or she has had a situation of chronic overdistention.

The above considerations may be used for those without SCI. In the elderly and those with nonneurogenic bladders, the Centers for Medicare and Medicaid Services (CMS) Guidance Tag F-315 emphasize that there must be a medically justified reason for an indwelling catheter in place for more than 14 days. These reasons include (a) documented PVR are in a range over 200 mL, (b) inability to manage the incontinence/retention with IC, (c) persistent overflow incontinence, (d) symptomatic infections, and/or renal dysfunction, (e) contamination of stage 3 or 4 pressure ulcers with urine that has impeded healing, and (f) terminal illness or severe impairment that makes positioning or clothing changes uncomfortable or that is associated with intractable pain (127).
 Principles of management of an indwelling urethral (Foley) catheter include (a) an oral fluid intake of at least 2 L/d unless there is a medical reason for fluid restriction, (b) keeping the catheter taped up to the abdomen of men when they are lying down to decrease the risk of a hypospadias due to downward pulling off the catheter during an erection, (c) daily cleaning the urethral meatus of incrustations with soap and water, (d) preventing reflux of urine into the bladder by not raising the drainage bag above the level of the bladder, (e) allowing effective urine drainage into the leg bag by not allowing the leg bag get more than half full, and (f) changing the Foley catheter every 2 to 4 weeks. In those with frequent encrustations, weekly catheter changes may be helpful. Catheter size should be limited to 14 to 16 Fr, especially in women. There is concern that larger diameter catheters may cause urethral dilatation and possible erosions. Health care providers will sometimes put in larger diameter catheters in women in an attempt to prevent leaking around the catheter rather than treating the cause (i.e., UTIs, detrusor overactivity). While the larger catheter may temporarily help the problem, the urethra frequently becomes dilated and the leaking around the catheter reoccurs.

In those with detrusor overactivity, especially with suprasacral SCI, the bladder wall can develop a lot of tone from the catheter rubbing against the bladder wall provoking uninhibited bladder contractions. The high tone in the bladder wall can in turn cause a functional obstruction at the ureteral orifices, resulting in decreased drainage from the upper tracts. Another result of the uninhibited contractions is autonomic dysreflexia. Therefore, periodic monitoring with urodynamics and the use of a pharmacological modality to quiet the bladder discussed in the section titled “Therapy for Incontinence Caused by the Bladder” should be strongly considered, particularly if there is urodynamic evidence of uninhibited bladder contractions, poor bladder wall compliance, and/or autonomic dysreflexia. Prophylactic antibiotics are not recommended for a patient with an indwelling catheter because of the risk of developing resistant organisms (106).

Before removing an indwelling Foley catheter, we believe it is best to obtain urine culture and sensitivity tests and to treat the patient with a culture-specific appropriate antibiotic. This is to decrease the risk of bacteremia if the patient is unable to void and develops a distended bladder. If a Foley catheter has been in place for 4 to 6 weeks before switching a patient over to IC, cystoscopy is recommended to remove eggshell calculi and debris that may have collected. Retrospective studies have reported that risks are higher with an indwelling Foley catheter compared with IC. These risks include the development of bladder stones; hematuria; bacteremia, especially if the catheter becomes obstructed; meatal erosions; penile scrotal fistulas; and epididymitis (128). However, it is important to note that in retrospective chart reviews, it is difficult to know if the indwelling catheter was put in place because of the complication or caused the complication. Prospective studies are lacking.

It has also been reported in retrospective studies that the risk of bladder cancer in those with SCI is increased in those using IC compared with those who were not. The risk has been reported to occur after 10 years of an indwelling catheter (106). Although not necessarily the predominant histologic type of cancer, squamous cell carcinoma is more common in those with SCI using indwelling catheterization. However, it should also be noted that squamous cell cancer is rare, so even though there is a greater incidence in those with SCI, the number of individuals who develop squamous cell bladder cancer is actually low (109). In a recent study the age standardized incidence of bladder cancer in SCI patients was 30/100,000. This was the same incidence for the general population in the area where the study was conducted (129). It is recommended that those with indwelling catheters have more frequent cystoscopic examinations than those who do not have in an indwelling catheter (106). This is done primarily to evaluate and remove bladder stones; however any questionable areas of the bladder can be biopsied at the same time. We generally perform cystoscopy once a year in those with an indwelling catheter. This is done more frequently if a person has recurrent bladder stones.

**Indwelling Suprapubic Catheterization**

If an indwelling catheter is going to be used as a long-term bladder management option, some men prefer to switch to a suprapubic catheter. Suprapubic catheters have a number of advantages to indwelling catheters because they pass directly into the bladder and not pass through the urethra. Some of the major advantages are that the catheter can be changes without undressing, the tubing is less likely to be sat on or kinked between a person’s legs, for those with sensation, it is much more comfortable, it is less likely to interfere with sexual activity (especially men), and is much easier to clean around (especially woman). It is also considered completely reversible because once the suprapubic catheter is removed, the tract closes, usually within 1 to 2 days. A suprapubic catheter has been found to be safer than an indwelling Foley catheter because it decreases the risk of epididymitis, urethral stricture disease, and urethral irritation (128). It has been reported that with the exception for bladder stones, the results of suprapubic catheterization were comparable to paraplegic SCI patients managed with continuous IC. They concluded that suprapubic catheterization is a valuable option of urinary management for tetraplegic patients (130).

It has been recommended that suprapubic catheters be considered for individuals with (a) urethral abnormalities, such as stricture, false passages, bladder neck obstruction, or urethral fistula, (b) urethral discomfort, (c) recurrent urethral catheter obstruction, (d) difficulty with urethral catheter insertion, (e) perineal skin breakdown as a result of urine, (f) leakage secondary to urethral incompetence, (g) psychological considerations such as body image or personal preference, (h) a desire to improve sexual genital function, and/or (i) prostatitis, urethritis, or epididymo-orchitis (106).

A recent prospective study revealed that in those with an SCI, there is a gradual decline in renal function, no matter what type of bladder management is used. Those with indwelling catheters did better than those on IC (131). Another study
are often necessary. These include sphincter dyssynergia, various treatments for the sphincter infections compared with IC.

The same principles of management for an indwelling urethral catheter apply to a suprapubic catheter with the exception that a 22 to 26 Fr catheter is used. The larger catheter has the advantage of being less likely to become kinked or get clogged.

**Reflex Voiding**

Another method of management that can be used in men with an intact sacral micturition reflex is reflex (spontaneous) voiding. This type of voiding involves having the person wear a condom catheter attached to a leg bag. The bladder has a spontaneous uninhibited contraction when it reaches a certain bladder volume. However, the volume that “triggers” the uninhibited bladder contraction is different for each person. One advantage to an external condom catheter and a leg bag is that is does not require good hand function. A caregiver can place on the condom catheter in the morning and not have to change it until the next day. Some assistance is needed, however, because even though the leg bag holds a 1,000 mL it should be emptied when it is half full. Similar to an indwelling catheter, another advantage is that there is no limit on fluid intake, in contrast to those who use IC as their method of bladder management. Major disadvantages are potential penile skin breakdown and having to wear an external condom catheter and leg bag and possibly a slightly increased risk of bladder infections compared with IC.

Because men who reflexly void usually have detrusor-sphincter dyssynergia, various treatments for the sphincter are often necessary. These include &alpha;-blockers, urethral stents, botulinum toxin injections into the sphincter, and sphincterotomy (106). These options are discussed in more detail in the section titled “Urinary Retention due to the Sphincter.” Those with reflex voiding need to have their upper tracts and lower tracts monitored. Although there is agreement that elevated voiding pressures cause upper-tract problems, there is no consensus as to what voiding pressure will cause this damage (133–135).

One study revealed that the most important voiding parameter causing upper-tract stasis in those who reflex void was the duration of the bladder contraction (136). Particularly before invasive treatment options, further studies are needed before treatment can be based solely on a numeric value of a specific urodynamic parameter. Currently, it is best to consider all of the variables and whether or a not a person is having problems prior to making changes. Problems may include autonomic dysreflexia, recurrent bladder infections, vesicoureteral reflux, kidney or bladder stones, kidney infection, or deterioration in bladder function (progressively higher postvoid residuals) or renal function (137).

**Therapy for Incontinence Caused by the Sphincter**

Preexisting occasional urinary incontinence from a weak sphincter mechanism, commonly found in elderly females, may become worse following a loss of mobility. Those with sacral and cauda equina injuries may also have urinary incontinence due to the sphincter. However, those with lower motor injuries will frequently have enough intrinsic sphincter tone to prevent the occurrence of urinary incontinence unless the bladder becomes overdistended or there is increased abdominal pressure (during transfers).

Urinary incontinence due to the sphincter also occurs with suprasacral SCI in those who have had a urethral stent or sphincterotomy or have urethral damage from being over-stretched by a long-term catheter or large-diameter urethral catheter. Care needs to be taken trying to increase the sphincter tone in a person with a suprasacral SCI because of exacerbation of detrusor sphincter.

Treatment for urinary incontinence due to the sphincter may be needed in those with SCI injuries if they have had prior treatment of their sphincter and want to switch to a different type of bladder management. An example would be a person who has had a sphincterotomy and wants to switch to IC or a suprapubic catheter.

**Therapy for Incontinence Caused by the Outlet or Sphincter**

**Behavioral Treatment Options**

Timed voiding sometimes is helpful in patients with mild to moderate incontinence who have normal bladder function but an underactive urethral sphincter mechanism. The object is to have the patient void before the bladder reaches full capacity. At full capacity, intravesical pressure is more likely to overcome the urethral pressure, resulting in leakage.

Pelvic floor (i.e., Kegel) exercises also may be tried in neurologically intact patients with mild to moderate stress incontinence caused by the sphincter (138). There is a great variation in the number of sets and repetitions described by various authors, with the total number of exercise contractions varying from 8 to 160/d (139). Exercises sometimes are combined with commercial biofeedback units (140). Patients have to be highly motivated since effects may not be seen for 4 to 8 weeks.

**Pharmacologic Treatment Options**

Alpha-adrenergic agonists may be useful at improving minimal to moderate stress incontinence caused by the sphincter. Wyndaele has reported success at decreasing urinary leakage around the Foley catheter in incomplete SCI women with patulous urethras (98). Ephedrine and phenylpropanolamine were two commonly used agents. Ephedrine causes a release of norepinephrine as well as directly stimulating &alpha;- and
\(\beta\)-receptors. Phenylpropanolamine is pharmacologically similar to ephedrine but provides less CNS stimulation (141). Phenylpropanolamine is rarely used because of reports that it may result in an increase in the risk of hemorrhagic stroke, especially in women (142).

A 4- to 6-week course of estrogen supplementation may be helpful in postmenopausal women with atrophy of the urethral epithelium or irritative symptoms from atrophic urethritis (143). Its beneficial effect may result from improving the local mucosal seal effect or increasing sensitivity or improving the number of \(\alpha\)-adrenergic receptors (144). Potential side effects and contraindications need to be weighed against potential benefits of using any agents to treat incontinence caused by the sphincter (141).

Periurethral collagen injection therapy is FDA-approved for those with intrinsic urethral sphincter deficiency. Clinical trials have focused primarily on non-SCI individuals. In a review, silicone particles and carbon spheres gave improvement at 12 months equivalent to collagen. A comparison of paraurethral and transurethral methods of delivery of the bulking agent found similar outcome but with a higher rate of early complications in the paraurethral group. The review concludes that in women with extensive comorbidity precluding anaesthesia, injection therapy may represent a useful option for relief of symptoms for a 12-month period, although two or three injections are likely to be required to achieve a satisfactory result (145). Bulking agents may be useful in SCI individuals who have decreased sphincter tone.

Significant progress has been made using a person’s own stem cells, which are injected into the urethra for those with weak urinary sphincters and stress urinary incontinence (SUI). It has been reported that at 1 year after implanting the cells, 94 of the 119 women (79%) were completely continent, 16 (13%) had a substantial improvement, and nine (8%) a slight improvement. Four patients were lost to follow-up. The incontinence and incontinence quality of life (I-QOL) scores, and the thickness, contractility, and EMG activity of the rhabdosphincter were significantly improved after treatment. The authors concluded that these results show the efficacy and safety of transferring autologous myoblasts and fibroblasts in the treatment of female SUI, after a follow-up of 1 year (146).

Before bulking agents or other treatment options are used, it is essential that detrusor overactivity or poor bladder compliance be ruled out with urodynamics and treated if present. Otherwise, increasing the urethral sphincter tone to prevent urinary incontinence may increase intravesical pressures, which may in turn result in back pressure and poor drainage from the upper tracts.

**Surgical Treatment Options**

In patients with a selective injury affecting just the sphincter mechanism, such as postprostatectomy or pelvic fracture, surgical implantation of an artificial urethral sphincter should be considered. It has been recommended that surgery should be delayed at least 6 months to 1 year to make sure there is no spontaneous return of sphincter function. Artificial sphincters are used infrequently in the adult SCI population, because they potentially can cause upper-tract damage in those with detrusor hyperreflexia and high intravesical pressure. In addition, there is an increased risk of prosthesis infection or erosion of the cuff in SCI patients because of frequent episodes of bacteriuria. Light and Scott (147) reported that 24% of their SCI patients developed infection requiring removal of the device.

For women with stress incontinence caused by the sphincter, or intrinsic sphincter damage, such as from a long-term indwelling catheter, a variety of surgeries have been developed to anatomically improve the urethral support and position. These procedures can be performed transabdominally, transvaginally, and even without surgical incisions. One- to three-year follow-up success rates have been reported to be 57% to 91% (148). A potential problem is that the operation works too well and causes retention. Patients therefore should be aware of the possibility of needing to perform postoperative IC. Other surgical options for those with intrinsic sphincter damage include surgical closure of the bladder neck followed by urinary diversion with an abdominal stoma that can be catheterized or the insertion of a suprapubic tube.

**Supportive Treatment Options**

Supportive options are similar to those for incontinence caused by the bladder. Specifically, these include diapers, external condom catheters, and indwelling catheters.

**Therapy for Retention Caused by the Bladder**

**Behavioral Treatment Options**

Timed voiding combined with increasing intravesical pressure either manually (i.e., Crede maneuver) or through increased intra-abdominal pressure (i.e., Valsalva voiding) may allow bladder emptying in individuals with an underactive or areflexic detrusor. In patients with weak uninhibited bladder contractions, suprapubic bladder tapping may be used to trigger a contraction (149).

Crede refers to pushing down with a closed fist in the suprapubic area with enough force to express urine from the bladder. Valsalva refers to bearing down with intra-abdominal pressure with enough force to push urine out of the bladder. The SCI Consortium bladder management panel recommended that one should consider not using Credé and Valsalva as a primary method of bladder management in those with SCI. If these maneuvers are going to be used, it is best reserved for those who are unable to perform IC and have decreased urethral sphincter activity, such as elderly women or SCI men with lower motor lesions and sphincterotomy (106).

Credé and Valsalva maneuvers may cause exacerbation of hemorrhoids, rectal prolapse, or hernia. Increasing intra-abdominal pressure in those with sphincter-detrusor dyssynergia often worsens the dyssynergia (150). Vesicoureteral reflux is a contraindication to this type of voiding.
Pharmacological Treatment Options
Bethanechol chloride, which provides relatively selective stimulation of the bladder and bowel and is resistant to rapid hydrolysis by acetylcholinesterase, is used to augment bladder contractions. A review of the literature shows that bethanechol is most useful in patients with bladder hypocontractility and coordinated sphincter function (151). Light and Scott (152) reported that it failed to induce bladder contractions in SCI patients with detrusor areflexia. Sporer et al. (153) found that bethanechol increased external sphincter pressures by 10 to 20 cm H₂O in SCI men. Therefore, it should not be used in those with sphincter-detrusor dyssynergia. It also is contraindicated in patients with bladder outlet obstruction. Potential side effects and contraindications must be weighed against potential benefits when pharmacologic agents are used to improve emptying (154).

Two investigational agents to improve bladder emptying are prostaglandins and narcotic antagonists. Intravesical prostaglandin F2α was noted to increase detrusor pressures in SCI patients with suprasacral lesions (155). Narcotic antagonists are thought to block enkephalins, which are believed to inhibit the sacral micturition reflex (156).

Surgical Treatment Options
There have been reports of surgically reducing the size of the bladder to decrease the postvoid residual; however, there is no effective way surgically to augment bladder contractions by operating on the bladder itself (157). Cloning and tissue engineering may be helpful in the future. Attempts at improving bladder emptying by performing a sphincterotomy in SCI men with urinary retention from detrusor areflexia have been reported. However, this is generally not recommended because there is a high failure rate at decreasing postvoid residuals if the bladder does not have contractions (158).

Supportive Treatment Options
A successful method for management of failure to empty caused by the bladder in those with hand function is IC. In those who are unable to perform IC, an alternative is an indwelling urethral or suprapubic catheter. Principles of management have been previously discussed.

Therapy for Retention Caused by the Outlet or Sphincter
Behavioral Treatment Options
Timed voiding and biofeedback methods have not been reported as successful methods of treatment in patients with neurogenic sphincter-detrusor dyssynergia. Biofeedback has been reported to be successful in patients with voluntary pseudosphincter-detrusor dyssynergia. These patients, who often are children, voluntarily tighten their sphincters during voiding, resulting in large postvoid residuals and UTIs (161).

In SCI patients with neurogenic detrusor overactivity and detrusor-sphincter dyssynergia, and stretching or scissoring and suprapubic bladder tapping have been reported as approaches that temporarily interrupt the dyssynergia and allow voiding (162).

Pharmacological Treatment Options
Alpha-Blocking Agents
In men with a suprasacral SCI and an intact sacral micturition reflex, reflex voiding into a condom catheter is sometimes used. However, upper-tract damage or elevated postvoid residuals can occur secondary to detrusor-sphincter dyssynergia. Alpha-adrenergic blocking agents have been shown to be effective at improving bladder emptying in patients with sphincter-detrusor dyssynergia and prostate outlet obstruction (163,164). In those with prostate outlet obstruction, α-adrenergic blocking agents are effective because the prostate smooth muscle is mediated by α-adrenergic stimulation. Placebo-controlled studies have shown both a clinically and statistically significant improvement in voiding in subjects taking phenoxybenzamine, prazosin, and more recently, terazosin (165).

Alpha-blocking agents may improve voiding in patients with sphincter dyssynergia secondary to an SCI due to several factors. After denervation, a supersensitivity of the urethra to α-adrenergic stimulation can occur. In addition, there may be a conversion of the usual β-receptors to α-receptors (24,26). Scott and Morrow (166) found that phenoxybenzamine worked well at decreasing residual urine volume in patients with suprasacral SCI and autonomic dysreflexia, but had variable effect on those without dysreflexia. An added benefit of α-blockers is their ability to blunt autonomic dysreflexia (167). When deciding which α-blocker to use, it is important to know that the manufacturer of phenoxybenzamine has indicated a dose-related incidence of gastrointestinal tumors in rats. There have been no cases of gastrointestinal tumors linked to phenoxybenzamine in humans in more than 30 years of use (168); however, the potential medicolegal issues of long-term use of phenoxybenzamine in young SCI patients should be considered.

The SCI Consortium bladder panel made several recommendations regarding α-blockers in those with SCI. They recommend to consider avoiding α-blockers in individuals who have symptomatic hypotension. They recommend that individuals are advised of the potential for complications of α-blockers, such as orthostatic hypotension. In addition, when first starting on an α-blocker, it is recommended to instruct the individual to take α-blockers at night, when supine. These instructions are particularly important for individuals with high-level spinal cord injuries because of the potential for orthostatic hypotension. It was also recommended to use phosphodiesterase inhibitors with caution in individuals with a high-level SCI who are on α-blockers. Particular caution should be used if α-blockers and PDE5 inhibitors are prescribed together (106).

Other Oral Sphincter Relaxing Agents
Three drugs that have been used for striated external sphincter relaxation are baclofen, diazepam, and dantrolene. In our
experience, these agents are not as effective as \( \alpha \)-blocking agents and should not be used as the drugs of choice for external sphincter relaxation. However, the author has noted on urodynamic studies that when baclofen is being tapered in individuals with SCI, they have an increase in detrusor-sphincter dyssynergia. Baclofen functions as an agonist for the inhibitory neurotransmitter gamma-aminobutyric acid (GABA), which blocks excitatory synaptic transmission, resulting in external sphincter relaxation. Diazepam is believed to cause external sphincter relaxation by increasing GABA inhibitory transmission in the spinal cord. Dantrolene acts peripherally by decreasing calcium release from the sarcoplasmic reticulum, thereby inhibiting excitation-contraction of the striated skeletal muscle fibers (169).

**Botulinum Toxin**

Botulinum toxin injections into the sphincter mechanism have also been used as a treatment for detrusor-sphincter dyssynergia. When botulinum toxin is injected into the urethral sphincter or bladder wall, it inhibits acetylcholine release at the neuromuscular junction, which in turn blocks neuromuscular contraction and relaxes muscles that are either spastic or overactive. It can, therefore, relax sphincter spasticity in those with detrusor-sphincter dyssynergia. Because over time, reinnervation of the neuromuscular junction occurs, botulinum toxin frequently loses its effectiveness after 3 to 6 months. Therefore, reinjections usually are necessary. There is no limit to the number of reinjections that may be required. In one study, botulinum toxin was injected into 24 individuals with detrusor-sphincter dyssynergia, 21 had significantly reduced urethral pressures with a concomitant decrease in postvoid residual volumes (170).

Botulinum toxin injections to treat detrusor-sphincter dyssynergia are especially useful in SCI individuals who have symptomatic hypotension, have adverse effects to \( \alpha \)-blockers or difficulty with compliance at taking medications. The SCI Consortium bladder panel recommended to (a) consider the use of botulinum toxin injections into the sphincter to help improve voiding in individuals with SCI with detrusor-sphincter dyssynergia, (b) monitor and inform individuals after botulinum toxin injections that onset is delayed up to 1 week and that the drug may lose its effectiveness in 3 to 6 months (106).

They also recommended to consider avoiding the injection of botulinum toxin into the sphincter of SCI individuals who (a) have a neuromuscular disease, (b) have a known allergy to or previous adverse effect from botulinum toxin, (c) are currently on an aminoglycoside, (d) have insufficient hand skills or caregiver assistance, (e) are unable to maintain a condom catheter, and/or (f) are female (unable to wear an external condom catheter) (106).

As with all treatments, potential side effects and contraindications must be weighed against potential benefits when using pharmacologic agents to improve emptying (167,169).

**Surgical Treatment Options**

**Urethral Stent**

Another method of sphincter dyssynergia treatment is the stainless steel woven mesh stent (e.g., Urolume Endourethral Wallstent, American Medical Systems), which holds the sphincter mechanism open. With an experienced team, this can be performed under local anesthesia. Because the sphincter is not cut, the procedure is potentially reversible with removal of the stent. The stent becomes covered by epithelium in 3 to 6 months, preventing calcium encrustations. A multicenter study of 153 men with SCI revealed a significant decrease in voiding pressures and postvoid residual urine volumes up to 2 years. Hydronephrosis resolved in 22 of 28 patients (78.6%). There was no loss of erectile function. Complications included mild postoperative hematuria in 10 patients (7.1%), penile edema in 2 patients, incrustation of the stent in 3 patients, stent removal (usually caused by stent migration) in 10 patients, and subsequent operation for bladder neck obstruction in 13 patients. Long-term follow-up studies are under way. This device has recently been approved by the FDA for the treatment of urethral strictures (171). In a 12-year follow-up study, five of the seven SCI men with urethral strictures developed bladder neck dyssynergia of varying degrees, as shown on VCUG; all were successfully treated with bladder neck incision. There were no problems with stent migration, urethral erosion, erectile dysfunction, or autonomic dysreflexia (172).

The SCI Consortium bladder panel recommended to consider avoiding a urethral stent in SCI individuals who (a) have insufficient hand skills or caregiver, (b) assistance to manage a condom catheter, (c) are unable to maintain a condom catheter, (d) are female, and (e) have urethral abnormalities (106). They are also advised that the individual be told of the potential for complications of urethral stents, such as (a) stone encrustation, (b) stent migration, (c) persistence of autonomic dysreflexia, (d) possible need for removal or replacement, (e) difficulty with removal, (f) possible urethral stricture after removal of stent, (g) urethral trauma, (h) tissue growth into the stent blocking urine flow, and (i) urethral pain (106).

**Transurethral Sphincterotomy**

Transurethral sphincterotomy (TURS) is a well-established treatment for SCI men with sphincter-detrusor dyssynergia. Indications include vesicoureteral reflux, high residuals with severe autonomic dysreflexia or recurrent UTIs, upper-tract changes with sustained high intravesical pressures, and poor compliance or side effects from medications being used to relax the outlet. Perkash reported a more than 90% success rate at relief of dysreflexic symptoms, decrease in residual urine, decrease in infected urine, and significant radiologic improvement. He stressed the importance of extending the incision to the bladder neck (173). The SCI Consortium Bladder panel recommended that one should consider TURS to treat detrusor-sphincter dyssynergia in males with SCI who want to use reflex voiding and who (a) have insufficient hand skills or caregiver assistance to perform IC, (b) have a repeated history
of autonomic dysreflexia with a noncompliant bladder, (c) experience difficult catheterization due to false passages in the urethra or secondary bladder neck obstruction, (d) have inadequate bladder drainage with severe bladder wall changes, (e) drop in renal function, (f) vesicoureteral reflex and/or (g) stone disease, (h) have prostate-rectal reflex with the potential for repeated epididymoorchitis, (i) experience failure with or intolerance to anticholinergic medications for IC, and/or (j) experience failure with or intolerance to α-blockers with reflex voiding (106).

It was recommended to consider avoiding TURS in males with a small retractable penis unable to hold an external collecting device unless a penile implant is planned following TURS.

The SCI Consortium Bladder panel recommended to advise males with SCI of the potential for complications of a sphincterotomy, such as (a) significant intraoperative and perioperative bleeding, (b) clot retention, (c) prolonged drainage with a large diameter catheter, (d) urethral stricture, (e) erectile dysfunction, (f) ejaculatory dysfunction, and/or (g) reoperation in 30% to 60% of cases.

The major concerns of most SCI patients are that the procedure is irreversible, it is a surgical procedure, and they will have to wear a leg bag. The traditional electrocautery method of performing a sphincterotomy has been reported to be at times causing bleeding requiring blood transfusion varying from 5% to 26%. This risk of bleeding has been largely eliminated with the use of Nd:YAG contact laser sphincterotomy (106).

Longitudinal studies have shown a 30% to 60% sphincterotomy failure rate. This has been attributed to a variety of causes, such as poor patient selection (i.e., those with detrusor areflexia or bladder contractions <30 cm H₂O), recurrent detrusor-sphincter dyssynergia, failure to recognize the need for a concomitant procedure (such as bladder neck incision or prostate resection), or new-onset detrusor hypocontractility (174).

**Supportive Treatment Options**

Supportive treatment options for the outlet are the same as those for retention caused by the bladder—specifically, IC or indwelling catheters. Occasionally, a person has so much sphincter spasticity that it is difficult to pass a catheter. Instillation of lidocaine jelly down the urethra 5 minutes before catheterization, administration of α-adrenergic blockers, or use of a coudé catheter often facilitates catheterization.

**Pediatric Bladder Management Considerations**

Children with incomplete emptying and sphincter dyssynergia caused by a nonneurogenic learned disorder have been treated successfully with biofeedback. This has involved looking at or listening to their sphincter EMG patterns during voiding. Sugar and Firlit (175) reported that all ten of their patients aged 6 to 16 years converted to a synergistic voiding pattern within 48 hours of therapy.

The same pharmacologic principles discussed under the general management section apply for children. The age of the child and decreased dosages need to be considered.

The same surgical procedures discussed under management can be used in children. In the past, children with vesicoureteral reflux were treated with urinary diversion, but because of long-term complications of urinary diversion and the excellent success of IC and ureteral reimplantation, this rarely, if ever, is used today. Surgical procedures for children with severe incontinence include an anterior fascial sling around the urethra and artificial urinary sphincter. There have been reports of 90% long-term success rates with the use of the artificial sphincter in children (176).

Clean IC has been shown to be effective treatment for children with failure to empty. In those with incontinence caused by the bladder, an anticholinergic medication often is also required (oxybutynin, 1.0 mg/yr of age twice daily). Parents often can learn IC in 1 day. It is thought that parents and children adjust to this program if it is started when the child is a newborn. There have been no reported cases of urethral injury, epididymitis, or UTIs requiring hospitalizations caused by this procedure. Children usually can begin performing their own IC at age 5 years. A child’s bladder capacity can be calculated using the formula: bladder capacity = 2 + (age) (in ounces).

In those who are unable to have a successful IC program due to problems such as significant urinary incontinence due to poor sphincter tone, small bladder capacity, or difficulty catheterizing the urethra, an alternative is to create a continent urinary diversion with either the Mitrofanoff principle (appendicovesicostomy) or a Monti tube (ileovesicostomy) can be performed. A concomitant bladder augmentation and urethral sling is frequently needed. In one 10-year follow-up study, 95% (51/54) of patients were continent, 90% (36/40) reported satisfaction, and 10% (4/40) reported dissatisfaction. Complications included bladder calculi (15%; 8/54), stomal stenosis (9%; 5/54), stomal bleeding (5%; 3/54), small bowel obstruction (2%; 1/54), and superficial wound dehiscence (2%; 1/54). This 10-year experience underscores that while complications do occur, there is successful long-term outcomes and durability of continent urinary diversions in children with neurogenic bladders (177).

**Complications of Voiding Dysfunctions**

**Urinary Tract Infections**

**Definition of a Urinary Tract Infection**

There is confusion in the literature and health care professionals regarding the definition of a true bladder UTI versus bacterial colonization in a person with a neurogenic bladder. It is important that a person being treated has a true UTI. Those being treated who do not have a true UTI will likely have unnecessary urine cultures, risk side effects of
antibiotics, and potentially develop resistant organisms, which will be difficult to treat if they develop a true UTI. In children and young able-bodied individuals, bacteria in the urine usually signify a bladder infection. However, no matter what type of bladder management, bacteriuria is common and usually represents colonization, not a bladder infection. The literature sometimes calls this an asymptomatic UTI. It is important to point out to patients and families that when an article says that a person will develop a bladder infection within a few days of putting in a catheter, or performing IC, the article is referring to colonization not a true bladder infection. To help clarify this misunderstanding, the National Institute on Disabilities and Rehabilitation Research had a consensus conference on UTIs in those with SCI. They came up with a very useful definition to help clarify the difference between a true UTI and bacterial colonization. To have a true UTI an individual with an SCI should meet three criteria. These are (a) the presence of bacteria in the urine, (b) an increase in white blood cells in the urine (pyuria) and (c) the onset of new symptoms (178). The following will give a further discussion and definition of each of these criteria.

Bacteriuria

Traditionally, a UTI was defined as more than 100,000 organisms in a midstream urine sample (179). However, studies show that symptomatic patients often have fewer than 100,000 organisms per milliliter. It has been found that 30% of able-bodied women with acute dysuria had less than 10,000 coliforms per milliliter, and many had less than 200/mL (180,181).

Bacteriuria is a common finding in patients with voiding dysfunctions. Lloyd et al. followed 181 new SCI patients discharged from an acute SCI center initially with sterile urine and on a variety of bladder management programs for 1 year. At 1 year, 66.7% to 100% had at least one episode of bacteriuria, depending on their bladder management program (182). Maynard and Diokno (183) reported on 50 new SCI inpatients on IC and found that 88% had one or more episodes of bacteriuria (i.e., any bacteria present). Elderly patients often compose a large part of the patient population in a rehabilitation hospital. Asymptomatic bacteriuria has been found to be present in 10% to 25% of community-dwelling and 25% to 40% of nursing home patients older than 65 years of age (184).

In SCI patients, Rhame and Perkash (185) reported that any specimen with more than 1,000 coliforms per milliliter was significant. Donovan et al. (186) thought that the appearance of any count of the same organism for two consecutive days was significant. The National Institute on Disability and Rehabilitation Research UTI consensus conference based the definition of significant bacteria on the method of urine collection and colony count. Significant bacteria did not necessarily mean an infection but, rather, confidence that the bacteria cultured were from the bladder and not contamination. For those on IC, “significant” meant greater than 10^2 colony-forming units (cfu)/mL; for those who did not use catheterization, greater than 10^4 cfu/mL; and for those who used an indwelling catheter, any detectable pathogens (178).

Pyuria

Elevated white blood cells in the urine (pyuria) generally signifies tissue invasion. Therefore, this is a reasonable reliable sign that a person has a UTI, particularly if he or she does not have a neurogenic bladder. Stamm et al. (181) found that 96% of able-bodied patients with symptomatic infections had greater than 10 leukocytes/mm^3. Deresinski and Perkash (187) reported that 79% of 70 SCI patients with symptoms and bacteriuria also had pyuria; however, 46% of asymptomatic patients also had significant pyuria. A person with a neurogenic bladder may have elevated WBCs but may not have a symptomatic “true” bladder infection. Possible causes include irritation from an indwelling catheter rubbing against the bladder wall, an intermittent catheter passing up the urethra, a bladder stone, or upper-tract problem such as a kidney stone. Anderson and Hsieh-ma (188) found that Gram-negative bacteria caused significant pyuria, but that this was not true of Staphylococcus epidermidis or Streptococcus faecalis, even in high numbers. Therefore, in cases of Gram-positive bacteria and a new onset of symptoms but no significant pyuria, a bladder infection should be considered.

Sign and Symptoms

The third component in the definition of a UTI is the new onset of signs and symptoms. It is important to emphasize the “new onset” to patients when asking about symptoms since many have chronic symptoms such as spasticity or pain. Signs and symptoms of a UTI involving the lower tract result from tissue invasion of the bladder wall with accompanying inflammation and generation of leukocytes from the bladder mucosal. In general, these may include cloudy urine, dysuria, frequency, urinary incontinence, or hematuria. Unless a person has had acute retention or urologic instrumentation, which can cause a bacteremia, it is very unlikely that a significant fever is due to a bladder infection. If a person complains of a high fever, a kidney infection (pyelonephritis) or nonurinary cause (such as an infected pressure ulcer) should be considered. Because many SCI patients have decreased or no bladder sensation, they may present with less specific signs and symptoms, such as weakness or malaise, increased abdominal or lower-extremity spasticity, new onset of urinary incontinence, occasionally retention from increased sphincter-detrusor dyssynergia, or autonomic dysreflexia in those with a lesion above T6.

Patients with acute upper-tract involvement may present with any of the above signs and symptoms. They also usually will have fever and chills and an elevated serum white blood cell count. Those with sensation usually report costal vertebral angle tenderness. It should be noted in the elderly that signs and symptoms may be much more subtle and patients may present simply with confusion or lethargy. UTIs also should be considered in the differential diagnosis of new cognitive changes in a head-injured patient.
Treatment of Asymptomatic Urinary Tract Infections

Guidelines for treatment have been difficult to establish for asymptomatic bacteriuria (colonization) because of articles that combine symptomatic UTIs and asymptomatic UTIs together and controversy whether or not colonization itself is a risk factor for UTIs and upper-tract damage. Ideally, the urine should be sterile; however, the side effects of antibiotics and development of resistant organisms need to be taken into account. Kass et al. followed 225 children on IC for 10 years and reported that in the absence of vesicoureteral reflux, bacteriuria proved innocuous, with only 2.6% of subjects developing fresh renal damage. In high-grade reflux, however, 60% developed pyelonephritis (189).

Lewis et al. studied 52 acute SCI patients during their initial hospitalization. Seventy-eight percent of patients had greater than 100,000 organisms, but only 13% had symptoms and required antimicrobial therapy over 6 months. Of interest is that 35% of cultures changed weekly from positive to negative, negative to positive, or one organism to another, necessitating a short course of antibiotics (190).

An accurate characterization of voiding dysfunction—such as voiding pressure, bladder compliance, and postvoid residuals, along with accurate characterization of level and completeness of injury—often is lacking in various studies discussing UTIs in SCI patients. Currently, there is a general consensus not to treat asymptomatic bacteriuria in those with neurogenic bladders. There is a strong consensus that asymptomatic bacteriuria in a patient with an indwelling Foley catheter should not be treated. Attempts should be made to eradicate asymptomatic bacteriuria before urologic instrumentation, in hydronephrosis particularly if this is due to partial obstruction since further obstruction could lead to sepsis, or in the presence of urea-splitting organisms. The authors have noted that some individuals are colonized with urea-splitting organisms such as Proteus. The authors will do a careful evaluation of the upper-tract (CT scan) and lower-tract cystoscopy and give two or three courses of antibiotics. If the organism persists, the author does not continue to treat the organism and monitors for possible stone formation 6 months to 1 year later. Use of prophylactic antibiotics is sometimes used for high-grade reflux. This should be strongly considered if individuals are getting episodes of pyelonephritis. However, it is best to control the reflux by preventing bladder overdistention, reducing intravesical pressure and/or change in bladder management (such as switching to an indwelling catheter).

Treatment of Symptomatic Urinary Tract Infections

Once a urine culture has been obtained, empiric oral antibiotic treatment can be started for patients with minimal symptoms while waiting for the culture results. Patients usually do well with a 7-day course of antibiotics. In those with high fevers (signifying pyelonephritis), dehydration, or autonomic dyssynergia, more aggressive therapy and upper-tract evaluation should be instituted. It is our opinion that these patients should be hospitalized, closely monitored, hydrated, and given broad-spectrum antibiotics (e.g., gentamicin and ampicillin) while waiting for the culture results and for defervescence of the fever. It is important to have an indwelling Foley catheter in place during intravenous or oral fluid hydration to keep the bladder decompressed. We believe that it also is beneficial to give an anticholinergic medication while the Foley catheter is in place; this will decrease the intrinsic pressure within the detrusor, allowing relaxation of the ureterovesical junction and improving drainage of the kidneys. Tempkin et al. (191) showed on renal scans that there was improved drainage of the upper tracts in SCI patients given anticholinergics. Anticholinergic medications will also help prevent autonomic dysreflexia in those with injuries at T6 and above. Autonomic dysreflexia may occur from uninhibited bladder contractions that may be provoked by the indwelling catheter. Patients with significant fever are likely to have upper-tract involvement (i.e., pyelonephritis) and, therefore, be continued on antibiotics longer than patients with bladder infections.

In addition, these patients with high fevers should undergo a urologic evaluation for cause of urosepsis. Acutely, this should consist of a plain abdominal radiograph to rule out an obvious stone, followed by a renal ultrasound. If there is a question of a stone, hydrenephrosis, or persistent fever, a CT scan will be helpful to further define the anatomy. Once the patient has been treated, it is often necessary that he or she undergoes a cystogram to evaluate for reflux, a cystoscopy to evaluate the bladder outlet and bladder, and urodynamics to evaluate voiding function.

Complications of Urinary Tract Infections

In addition to acute lower UTIs (i.e., cystitis) and acute upper UTIs (i.e., pyelonephritis), the physician should be aware of other potential problems. Those from lower UTIs include epididymitis, prostatic or scrotal abscess, sepsis, or an ascending infection to the upper tracts. Complications that may occur from upper-tract infections include chronic pyelonephritis, renal scarring, progressive renal deterioration, renal calculi if there is a urea-splitting organism such as Proteus, papillary necrosis, renal or retroperitoneal abscess, or bacteremia and sepsis.

Role of Prophylactic Antibiotics

There is controversy over the usefulness of prophylactic antibiotics. Anderson (192) reported a statistically significant difference in bacteriuria in SCI inpatients on a combination of oral nitrofurantoin and neomycin/polymyxin B solution compared with controls. Merritt et al. (193) reported a statistically significant decrease in bacteriuria with methenamine salt or co-trimoxazole compared with controls at 3 to 9 months, but not at greater than 15 months. Maynard and Diokno (183) reported that antibiotic prophylaxis significantly reduces the probability of a laboratory infection but not the probability of a clinical infection. Kuhlemeier et al. (194) evaluated vitamin C and a number of antimicrobial agents as prophylactic agents and found no beneficial effect in SCI patients compared with controls. These studies seem to show that prophylactic agents do not have a long-term effect in decreasing bacteriuria.
Bladder Stones
Bladder calculi are the second most common cause of morbidity (following UTIs) in those with SCI (195). The reason for this is the common use of indwelling catheters following SCI and the close association of bladder stones and indwelling catheters (196–198). The 1995 National SCI database found that in women, indwelling catheters are the second most common type of management at discharge (35%) and most common type of management in those at 5 and 10 years out from injury (46.5% and 55.1%, respectively). In men at discharge, indwelling catheters were the third most common type of bladder management (20.5%) and second most common type of management at 5 and 10 years out from injury (27.1% and 36.4%, respectively) (195).

A prospective study had a 35% incidence of bladder stones in those with indwelling catheters (197). In another large retrospective study of 500 patients, the absolute annual risk of stone formation in patients with a catheter was 4% compared with 0.2% for those on intermittent self-catheterization. However, having formed a stone, the risk of forming a subsequent stone quadrupled to 16% per year. Bladder stones were no more likely to form in patients with suprapubic catheters compared to those with indwelling urethral catheter type (198). We have noted that if a person has bladder stones, they are likely to have a recurrent stone. Conversely, those who do not have a stone are less likely to have a bladder stone on a subsequent cystoscopic examination.

The majority of bladder stones are composed of struvite caused by a urease-producing organism. These organisms lay down a biofilm on the surface of the catheter that then forms into stones. Complications from indwelling catheters include recurrent UTIs, hematuria, and blockage of an indwelling catheter with potential life-threatening autonomic dysreflexia (195,197). Prevention of stones is best accomplished by catheter changes every 2 to 4 weeks, and treatment of stone-forming organisms such as Proteus. In those with frequent recurrent bladder stones, daily instillations of renacidin (citric acid, glucono-delta-lactone, and magnesium carbonate) irrigation (15 to 30 mL left in the bladder for 5 to 10 minutes, two to three times a day) may be tried. Instillations may cause a problem in those with reflux or they may cause autonomic dysreflexia with bladder distention. In individuals identified as frequent stone formers, weekly catheter changes have also been found to reduce bladder stones (199). We have noted that weekly catheter changes also decrease catheter blockage in those with recurrent bladder stones.

The role of high fluid intake at preventing bladder stones is not known. The standard treatment of bladder stones involves cystoscopy and the breaking up and removal of the stones under direct visualization. Unfortunately, abdominal x-ray is not a reliable method to detect bladder stones in individuals with SCI. Only 21% of bladder stones found during cystoscopy were detected by the x-ray (200). Individuals with stones on their catheter should, therefore, be referred for cystoscopy to prevent catheter clogging. Finding encrustation of the catheter when it is removed is a reliable sign that there are retained stones in the bladder. Eighty-six percent of those with catheter encrustation were found to have stones on cystoscopy (197).

Bladder Cancer
Individuals with SCI who have indwelling catheters or bladder augmentations have been reported to have a higher risk for bladder cancer than their able-bodied counterparts (201,202). Among those with indwelling catheters, this increased risk may only be for squamous cell cancer, which is an extremely rare form of bladder cancer (129) (see above discussion of bladder cancer in the section “Indwelling Urethral Catheterization”). Possible causes include chronic irritation from UTIs, stasis of urine, bladder stones, or exposure of the bowel mucosal to urine following a bladder augmentation or urinary diversion (201,202). There have been a few reports of adenocarcinoma of the bladder in patients who have had an augmentation cystoplasty for more than 10 years (110). Hematuria was frequently a presenting sign in those with bladder cancer (202). These studies suggest that an evaluation of hematuria is important, particularly in the absence of a UTI, and yearly cystoscopy should be performed in people who have had indwelling catheters for more than 10 years. Before 10 years, many centers perform yearly cystoscopy on those with an indwelling catheter, primarily to rule out bladder stones. The role of other screening methods, such as cell cytology, is under investigation.

Vesicoureteral Reflux
Price and Kottke (203) reported in an 8-year study that vesicoureteral reflux was one of the factors frequently associated with renal deterioration after SCI. Fellows and Silver (204) found that there was a definite association between the degree of reflux and renal damage. Vesicoureteral reflux in children has been associated with a congenital shortening or absence of the submucosal ureter, absence of ureteral muscle in the submucosal segment, or a paraureteral diverticulum of the bladder (205). In people with neurogenic voiding dysfunctions, high intravesical pressures are thought to be a major cause of reflux. Recurrent cystitis and anatomic changes in the oblique course of the intravesical ureter caused by bladder thickening and trabeculation are believed to be other causes of reflux. Renal deterioration from reflux is thought to be secondary to recurrent pyelonephritis, resulting in renal scars as well as back-pressure hydronephrosis.

The mainstay of treatment in those with reflux and voiding dysfunction is to lower intravesical pressures and eradicate infections. Ureteral reimplants are technically difficult to perform in a trabeculated bladder and have not been uniformly successful.
Renal Calculi
Approximately 8% of patients with SCI develop renal calculi (i.e., staghorn calculi or struvite stones) (206). Kuhlemeier et al. (207) found that renal calculi were the single most important cause of renal deterioration. Without treatment, a patient with a staghorn calculus has a 50% chance of losing the involved kidney (208). DeVivo and Fine (206) evaluated 25 SCI patients who developed calculi and 100 SCI patients who did not have calculi and found that those with calculi were more likely to have neurologically complete quadriplegia, have Klebsiella or Serratia infections, prior history of bladder calculi, and high serum calcium values. Patients who present with persistent Proteus infections also should be monitored for renal calculi. Urea-splitting organisms form alkaline urine that in turn causes supersaturation and crystallization of magnesium ammonium phosphate.

Previously, a surgical pyelolithotomy or nephrolithotomy was performed to remove these stones (Fig. 51-7). Newer techniques, including percutaneous nephrolithotomy and extracorporeal shock wave lithotripsy, have largely replaced open surgical procedures (209). All of these procedures need to be combined with sterilization of the urinary tract of urea-splitting organisms. DeVivo and Fine (206) reported a 72% recurrence rate within 2 years of the first kidney stone. Investigations are under way on the use of acetohydroxamic acid as a prophylactic agent; limitations of this agent are reported side effects and high cost (210).

Hydronephrosis
Ureteral dilation for any reason results in inefficient propulsion of the urine bolus caused by inability of the walls to coapt completely, as well as in decreased intraluminal pressure caused by the increased ureteral diameter. Over time, this may result in further distention of the ureter with eventual hydronephrosis (7,211). There are several causes for ureteral dilation. It can occur transiently from a brisk diuresis effectively overloading the ureters, not allowing enough time for individual boluses to travel down the ureter. Another cause may be a mechanical obstruction, such as a stone or stricture. Those with poor bladder wall compliance, sphincter-detrusor dyssynergia, or outlet obstruction may develop a functional obstruction caused by high intravesical pressures. The elevated intravesical pressure increases the tension within the bladder wall, which in turn constricts the submucosal ureter, increasing the hydrostatic force within the bladder. Ureteral dilation will occur if ureteral peristalsis is unable to overcome these increased pressures (212).

McGuire et al. (91) reported that 81% of myelodysplastic children with leak point pressures greater than 40 cm H\textsubscript{2}O developed upper urinary tract changes, whereas only 11% with leak point pressures below 40 cm H\textsubscript{2}O developed upper-tract changes. Hydrostatic forces in the ureter and kidneys also may be increased by vesicoureteral reflux blocking the downward egress of urine (212). Teague and Boyarsky have identified another potential cause of ureteral dilation. They found that Citrobacter sp. and Escherichia coli from human urine cultures injected into the lumen of dog ureters produced marked suppression of peristalsis and ureteral dilation lasting up to 2 hours (213).

Renal Deterioration
Renal failure previously was the leading cause of death following SCI. The death rate from renal causes was reported in the 1960s at between 37% and 76%. Careful monitoring of the upper and lower urinary tract combined with effective bladder management programs has markedly reduced the incidence of renal failure. Price and Kortke (203) followed 280 patients for 8 years and reported 78% had good function, 13% mild deterioration, 4% moderate deterioration, and 5% severe deterioration. Factors most frequently associated with renal deterioration were vesicoureteral reflux, renal calculi, recurrent pyelonephritis, and recurrent decubitus ulcers. Kuhlemeier et al. evaluated 519 SCI patients with renal scans for up to 10 years. They found that factors associated with a statistically significant decreased ERPF were tetraplegia, renal stones, female patients older than 30 years of age, and a history of chills and fever presumably caused by acute UTIs. Renal calculi were the most important cause. Factors not found to be statistically significant included years since injury, presence of severe decubitus, bladder calculi, bacteriuria without reflux, and completeness of injury (207).
Renal Amyloidosis

Amyloidosis has rarely been reported in those with SCI. Of particular concern in those with SCI has been renal involvement (214–218). The clinical findings vary with the organ involved. If the liver and spleen are involved, these organs will usually be palpable. If the kidneys are involved, albuminuria has been reported as a consistent sign. Edema of hypoproteinemia, hyaline and granular casts, and azotemia are late findings of renal involvement (215). It has been hypothesized that amyloidosis is due to the triad of three inflammatory processes, specifically pyelonephritis, decubiti, and osteomyelitis. It has been suggested that amyloidosis is associated with the total mass of the inflammation rather than a single focus (214).

It is interesting to note that most of the publications are from 30 or more years ago. Improved methods to treat pyelonephritis, pressure ulcers, and osteomyelitis have probably contributed to this being a very infrequent occurrence in a person with SCI. There is some controversy on the reversibility of amyloidosis. It has been reported that after the treatment of pressure sores with skin grafts, serum albumin improved and patients clinically felt better (216). It has also been found that albumin improved and there was remission of the nephritic syndrome. Renal biopsies did not show any decrease in amyloid (217). Another study reported that amyloid did regress after treatment of the inflammatory condition (218). These studies help to emphasize the importance of attempting to prevent and to effectively treat secondary inflammatory conditions in those with SCI.

Autonomic Dysreflexia

Another potential problem that can occur in those with injuries at thoracic level 6 (T6) and above from urologic causes is autonomic dysreflexia. Of most concern is a sudden severe elevation in blood pressure that occurs until the cause of the autonomic dysreflexia is removed. Other symptoms may include sweating, flushing, “goose bumps,” headache, and bradycardia (219). “Silent” dysreflexia may also occur, in which there are no symptoms of dysreflexia despite a significant rise in blood pressure (89).

Because autonomic dysreflexia can occur with any noxious stimuli, various urologic problems, such as bladder distention, UTIs, epididymitis, and urologic instrumentation, may provoke its onset. The ejaculation reflex, whether it is induced by vibratory stimulation or electroejaculation, will usually cause autonomic dysreflexia in those with injuries at or above T6 (219). The onset of uninhibited bladder contractions has been noted in urodynamics studies frequently to cause a rise in the blood pressure (89).

When a person presents with autonomic dysreflexia, one should first sit the person up and loosen any restrictive clothing. It is then important to evaluate for any potential urologic causes, such as a blocked catheter or kinked catheter tubing. If a catheter is blocked and going to be changed, and lidocaine jelly is readily available, consideration should be given to instilling the lidocaine into the urethra before the catheter change. This helps to decrease sensory afferent input that could increase autonomic dysreflexia. A national guideline on the acute management of autonomic dysreflexia for individuals with SCI presenting to health care facilities has been published (219). Lidocaine can also be instilled into the bladder before urologic procedures or during an acute episode of autonomic dysreflexia when a person is felt to be having a UTI (97).

THE GASTROINTESTINAL SYSTEM

Traditionally, urinary and gastrointestinal system impairments are grouped together for discussion because of similarities in anatomic location, function, and neurologic innervation. However, there are limits to further extension of this analogy and a few important distinctions. The colon is only the terminal segment of the gastrointestinal tract and is greatly influenced by activity up the alimental pathway. Neurogenic bowel refers primarily to functional changes in the colon and pelvic floor, although a variety of neurologic conditions affect various other gastrointestinal organs as well (220). The sequelae of neurogenic bladder dysfunction can frequently be life threatening, whereas neurogenic bowel dysfunction is seldom life threatening but certainly life limiting, affecting the quality of life. Neurogenic bowel dysfunction can prevent full continence, voluntary defecation, and self-management of clean-up; cause dread and fear of incontinence; and constrict full participation in life (221).

SCI is the condition that is most frequently associated with neurogenic bowel dysfunction, although there are many other neurologic conditions contributing to gastrointestinal dysfunction as well. Furthermore, gastrointestinal involvement is not limited to the colon after SCI; all segments are affected (220). Oral hygiene is affected in patients with high tetraplegia. SCI can contribute to dysphagia as a result of local protrusion of osteophytes discs or cervical fixation hardware. Gastric ulcers are more common after SCI and peak during the first few weeks after injury. The prevalence of gallstones after SCI can reach 30%, yet in spite of this high prevalence, acute cholecystitis does not occur with greater frequency or at increased risk for mortality (222,223). The transverse portion of the duodenum (third portion) crosses the midline within an angle created between the aorta posteriorly and the superior mesenteric artery, which arises off the anterior wall of the aorta superior to the transverse duodenum. Rapid weight loss causes loss of the fat pad that maintains this angle and may result in constriction of the third portion of the duodenum. Superior mesenteric artery syndrome produces gastric distention, bloating, and vomiting and further weight loss that results from the superior mesenteric artery descending and pinching off the distal duodenum: this is termed the nutcracker effect. Dilation of the stomach and the proximal duodenum are present radiographically.

While infrequent, nontraumatic pancreatitis sometimes occurs after acute SCI, this may be due to spincter of Oddi spasm, thickening of pancreatic secretions, and vagal dominant innervation (220,224). Serum amylase, though nonspecific, is
the most useful tool to prompt further studies and initiation of supportive care (NPO, hydration and optimization of oxygenation). Unfortunately, stool guaiac is so nonspecific as to be irrelevant in diagnosis of pancreatitis, and hypocalcemia (speculated to be the consequence of ionized extracellular calcium becoming linked to fats in the peripancreatic inflammatory phlegmon also known as saponification) is a late finding in very severe pancreatitis—not a useful diagnostic tool either.

Anatomic and Physiological Considerations

Colon

The large intestine serves to desiccate and form fecal matter as it is transported from the ileum to the rectum. There the feces can be willfully evacuated at the right time and place. In addition to its role as a transport and storage organ, the colon functions to absorb water and electrolytes. The colon normally reduces the approximately 1,500 mL of small intestinal content that is introduced into the cecum to about 150 mL/d.

Colonic propulsion of the fecal contents requires motility, a term frequently misused and often misunderstood. Unfortunately, a variety of experimental measurements have been associated with motility and attributed to functional fecal transport (225). Motility and fecal transport are measured differently. For motility, colonic wall contractions can be quantified manometrically (226,227), and/or electronically via electrode recordings of the patterns and intensity of smooth muscle activation (228). It is not possible to deduce the movement of material through the colon by the measurement of electrical activity, however. Therefore, unqualified terms such as increased motility or decreased motility say little about success with fecal transport. Alterations in colonic transport may be caused by events that occur outside of the colon, such as distal obstruction or proximal dilation mediated by the intestinal colic inhibitory reflex. Indices of motility may be abnormal in situations with low transport. Therefore, interventions directed solely at correcting the abnormal indices of colonic motility may not be effective.

The most clinically useful index of overall colonic function is colonic transit time, the amount of time required for contents to pass from the cecum to the outside. Several techniques exist for measuring transit time. A practical economic method uses radio-opaque markers that are swallowed and followed through the intestinal tract with serial abdominal radiographs (229,230). Swallowed markers are counted on serial radiographs (229). Although even such objective measures have limitations, as expected, the transit of material through the colon does depend on colonic motility. After SCI, in the absence of regular spontaneous or willful defecation, transit and motility are dependent on regular scheduled bowel care. Rectal distention in itself can reflexively inhibit peristalsis and affect alimentary function as proximal as the stomach, preventing gastric emptying.

Anorectum

The anal canal provides the primary mechanism for fecal continence and is the barrier that must be traversed for evacuation to occur. The internal anal sphincter (IAS) is the specialized thickening of the circular smooth muscle layer of the rectum that maintains a continuous state of maximal contraction. It is responsible for the majority of resting tone in the anal canal (231).

Normal internal anal canal resting pressures range from 50 to 100 cm H₂O. This resting tone is not altered after SCI (232). The external anal sphincter (EAS) is a striated muscle that is continuous with the pelvic floor and is innervated by the pudendal nerves bilaterally. The EAS, along with muscles of the pelvic floor, displays the unusual property of continuous electrical activity in both the waking and sleeping states (233). Presumably, this special property allows the EAS and pelvic floor to maintain continence without conscious will or during sleep. Although the EAS plays a small role in the resting state, contraction of the EAS can double anal canal pressure for short periods of time. External sphincter function is felt to be important during events that are an acute threat to continence, such as coughing, acute rectal contraction, or assuming an upright position.

The puborectalis muscle originates from behind the pubic symphysis and extends posteriorly to loop around the rectum just proximal to the anal canal. The puborectalis tugs the rectum anteriorly and creates an angle between the rectum and the anal canal, the anorectal angle. This kink in the fecal pathway aids in continence (234,235). Conversely, failure of the puborectalis to relax appropriately, with persistence of an acute anorectal angle, has been associated with inability to defeate (236). The IAS, the EAS, and the puborectalis work together synergistically to maintain continence and are collectively named the anal sphincter mechanism.

The mucosa at the proximal end of the rectum and anal canal (anorectal junction) is rich in sensory receptors. This allows for socially critical judgments about the phases of the matter therein: liquid, solid, and gaseous material. The rectoanal inhibitory or sampling reflex allows a sample of the rectal contents to come into contact with this sensory zone. The rectoanal inhibitory reflex consists of a transient relaxation of the IAS stimulated by a rise in rectal pressure. A simultaneous increase in EAS tone, the guarding reflex, occurs to preserve continence while sensory receptors of the anorectal junction appraise the contents. The rectoanal inhibitory reflex occurs during sleep and throughout the day, usually at a subconscious level.

Regulation of Gut Function

Fortunately, many of the organs, tissues, and cells of the gastrointestinal system operate autonomously. Regulatory influences include the parasympathetic and sympathetic nervous systems, the enteric system within the gut, and endocrine, myogenic, and intraluminal cues. Preliminary knowledge is available but further investigation is required. Rudimentary explanations are offered so that the reader may exploit these systems therapeutically in neurogenic bowel management. The variety of these regulatory influences act in concert and interact in feedback loops. They can be explained from the gut lumen out.
### Intraluminal Content

The intraluminal contents can affect gastrointestinal function through alteration of the physical characteristics of stool, bacterial action, and the effects of various substances on mucosal receptors. The physical characteristics of the fecal mass have an important effect on colonic motility and defecation efficiency. The typical Western diet, deficient in fiber, results in small, hard, often scybalous stool that is difficult for the colon to propel and the rectum to evacuate. The effect of stool consistency on rectal evacuation has been well studied in able-bodied subjects (237,238). Dry, hard stool is not compressible; breaks apart; and contributes to high friction against the colonic wall. These stools produce difficulty with evacuation.

Addition of vegetable fiber to the diet increases stool volume and moisture content, which increases plasticity and decreases transit time (239). With increased dietary fiber, fecal microbial mass also increases (240). Bacterial fermentation of indigestible fiber within the colon generates short-chain fatty acids. These short-chain fatty acids are passively absorbed by colonic mucosa and may be oxidized as an important energy source (241). Bacteria also are important in the metabolism of bile salts. Bile salts stimulate colonic motility and may play a role in the hypermotility of patients with the irritable bowel syndrome (242).

Furthermore, the luminal content also can affect gastrointestinal function by the stimulation of specific mucosal receptors. Five types of gastrointestinal sensory receptors are known to exist. These respond to mechanical, chemical, osmotic, thermal, or painful stimuli.

### Intrinsic or Enteric Neural Regulation

The enteric intrinsic nervous system is embedded in the wall of the gut and runs the length of the digestive tube, from the pharynx to the anus. Meissner’s plexus is distributed throughout the submucosa and transmits local sensory and motor signals to Auerbach’s plexus, autonomic ganglia, and the spinal cord. This intrinsic nervous system includes Auerbach’s plexus (intramuscular myenteric) distributed between the longitudinal and circular muscle layers. The enteric nervous system is organized into three levels:

- **Afferent neurons**, which gather sensory information and relay it to an array of interneurons
- **Interneurons**, which process information locally, integrate the incoming sensory information, and coordinate local motor and secretory responses
- **Efferent neurons**, which exert their influence on target cells, such as secretory, absorptive, or muscle cells

The enteric nervous system integrates sensory information from the contents and coordinates local and distant secretion and peristalsis. Ablation of all enteric nervous activity by the neurotoxin tetrodotoxin causes the colon to contract, the rectum to undergo tonic and phasic contractions, and the IAS to contract tonically (243). Thus, it appears that the major effect of the enteric nervous system on the lower gastrointestinal tract is to provide an inhibitory influence.

### Extrinsic or Extraintestinal Neural Regulation

Extrinsic neural influences provide overall coordination of intestinal reflexes as well as integration of the gastrointestinal tract with the whole organism. Sensory receptors in the gut may send afferents directly to the CNS, to prevertebral sympathetic ganglia, or to interneurons within the gut wall. The sensory information may be processed within the CNS, the prevertebral ganglia, or the enteric nervous system (Fig. 51-8). In fact, the enteric nervous system contains a greater number of neurons than the spinal cord (244). This is no surprise when the complexity and variety of gastrointestinal functions controlled by the enteric nervous system are considered. The important regulatory role of the enteric nervous system also is apparent in the organization of visceral efferents from the CNS. All efferents, except those to striated muscle in the
The gastrointestinal tract receives both parasympathetic and sympathetic nervous system innervation. The function of the parasympathetic efferents is complex. Separate groups of vagal preganglionic fibers may innervate inhibitory or excitatory neurons in the same organ. Therefore, whole-nerve stimulation may produce excitatory as well as inhibitory effects in the same organ. For example, transaction of the vagus at the level of the lower esophagus (i.e., truncal vagotomy) has been used as treatment for peptic ulcer disease. Gastric acid secretion is diminished to a variable extent, and gastric motility is significantly impaired; however, small bowel and colonic motility are largely unaffected. Injury to the inferior splanchnic nerve (parasympathetic S2-4) results in impaired defecation and constipation (245). The function of the sympathetic nerves is generally to cause inhibition of motor and secretory activity and contraction of gastrointestinal sphincters. Sympathetic stimulation leads to adynamic ileus and decreased bowel activity. Although surgical sympathectomy has little clinical effect on bowel function in humans, diarrhea has been reported in animal models (246).

The striated muscle of the pharynx and the EAS is innervated by α-motor neurons directly from the CNS. The cell bodies of the α-motor neurons to the EAS are located in the anterior horn cells of spinal segments S2 through S4. Their axons are carried in the pudendal nerve. Injuries to the pudendal nerves or the sacral cord produce flaccid paralysis of the pelvic floor along with the external sphincter.

As can be inferred from the previous discussion, the regulatory mechanisms for gastrointestinal function are highly interdependent and frequently redundant. Loss of any individual component does not necessarily result in an identifiable syndrome. In fact, the inherent automaticity of the gut often allows it to function quite well in the absence of all extrinsic control. The effects of loss of extrinsic neural control (i.e., neurogenic bowel) are most apparent at the two ends of the gastrointestinal tract, where complex voluntary behaviors occur. The primary function lost with neurogenic bowel is voluntary control over defecation.

GASTROINTESTINAL HORMONES

The gastrointestinal hormones influence activity in three systems throughout the alimentary canal. This influence is communicated by blood-borne chemical messengers (i.e., endocrine action) and secondarily transmitted by locally released messengers that traverse the interstitial space to reach a target cell (i.e., paracrine action). A third chemical messenger system, even more specific in its location of activity, is the neurocrine system. These neuropeptides are active only at the nerve terminals where they are released. The central role of neuropeptides in both neural and endocrine function illustrates the overlap and inseparability of these two systems.

Our understanding of the neuroendocrine potential of the colon has progressed much over the last decade (247). The colon and rectum harbor peptides and amines that come from enterochromaffin cells in rectal crypts and enkephalinergic neurons in the myenteric plexus (247). The colon may be a source of enterogastrone, which is secreted in response to intracolonic nutrients and reduces gastric acid secretion from the stomach. Locally acting neuropeptides, including intestinal peptide, substance P, and GRP/bombesin, stimulate secretion of ions. Conversely, somatostatin, opioid peptide, and neuropeptide stimulate absorption.

Transit and Motility

Studies of colonic transit time in people with SCI generally have shown a prolongation of total colonic transit time that is predominantly caused by slowing through the rectosigmoid segment (248,249). Interpretation of this finding is difficult because transit time studies cannot differentiate slow transit through the rectosigmoid because of altered colonic motility from slow transit because of infrequent or inefficient evacuation of the rectum. In our SCI population, bowel care was performed at a mean interval of approximately 2 days; the interval between bowel movements in a neurologically intact control group was 1 day. Because any movement in the rectosigmoid colon between bowel-care days would not be expected, the difference between SCI and controls may simply be related to the frequency of bowel care and bowel movements. Furthermore, Kellow et al. have demonstrated that distention of the rectum slows intestinal transit in both the fasting and fed states. Inhibition of contraction in one portion of the intestinal tract by distention of another (i.e., intestinointestinal, intestinocolic, and colocolic reflexes) was first reported 50 years ago (250). Interestingly, these reflexes are mediated in the prevertebral ganglia, not the spinal cord, so they likely would be unaffected by SCI (251).

Attempts to describe the effect of SCI on colonic motility by measurement of electrical or pressure events have yielded inconsistent results. Connell et al. (252), measuring intraluminal pressure waves, concluded that people with injury levels above T9 had decreased rectosigmoid motor activities, whereas those with lesions below T9 had increased motor activity. However, Aaronson et al. (253) found an increase in rectal myoelectric activity in a disparate group of six people with SCI when compared with neurologically intact controls. Glick et al. (254) found no change in myoelectric activity when nine people with SCI were compared with controls. Interpretation of these studies is plagued by many of the previously mentioned methodological problems. The study groups were small and uncontrolled for important variables such as time after injury, age, level and completeness of injury, and presence of gastrointestinal symptoms. Even if a consistent pattern of motility disorder could be found, it could as easily be caused by the inability to evacuate as by a primary motility problem.

Fecal Continence

Fecal continence is achieved with a combination of structures that work together to prevent inadvertent passage of stool.
These barriers are removable to allow bowel evacuation to occur. The current state, continence or evacuation, depends on the balance between those forces that favor expulsion of stool and those that resist it (Fig. 51-9). Expulsive forces include intra-abdominal pressure, colorectal contraction, elastic forces, and gravity. Resistive forces include the anorectal angle, anal canal tone contributed by the IAS and EAS, and friction. Stool consistency is a pivotal factor that can shift the balance in either direction. Small, hard stool is evacuated less completely and with more difficulty than soft, bulky stool (237,238), whereas even subjects with normal continence mechanisms may be incontinent when faced with large volumes of liquid stool as in during an acute diarrheal illness.

**CLINICAL EVALUATION**

As with any patient, the evaluation begins with a complete history and physical examination. A request for symptoms guides the evaluation. Symptoms typically are vague; clues to their origin may be derived from a careful search for a relationship to exacerbating or remitting factors. The effect of body position, time of day, eating, bowel care, medications, and urinary function should be elicited (255). The presence of associated symptoms, such as autonomic dysreflexia, abdominal wall spasticity, fever, and weight change, should be noted. A gastrointestinal review of symptoms further clarifies function. A careful query of premorbid bowel function should be made, because neurologic lesions may alter the way that preexisting problems are manifest. Recognition of the impact of the symptom on the patient’s ability to carry out important life activities is essential; neurogenic bowel dysfunction is notorious for detrimental impact on quality of life (256). Finally, the components of the bowel program should be systematically evaluated. A history of the type of diet—with special emphasis on fluid and fiber intake, use of laxatives, stool softeners, fiber supplements, and medications with anticholinergic properties—should be obtained. The frequency, duration, and technique of bowel care are elicited, as well as problems with stool consistency, lack of stool in the rectum at the time of stimulation, incontinence, and bleeding.

The physical examination is not only an assessment of the colon and pelvic floor but also includes a survey for associated pathology, sensory/motor impairments, and activity limitations that could affect bowel care (220). Abdominal examination allows for motor assessment of the abdominal wall, percussion for tympany, palpation of the colon for impaction, and elicitation of symptoms (255). Pelvic floor innervation and function should be appraised with neurologic and rectal examination. The anocutaneous reflex is recognized with EAS contraction on anal skin contact. The bulbocavernosus reflex is elicited with penile head squeeze or clitoral pressure and the recognition of an increase in anal sphincter tone. Both these reflexes suggest an intact sensory (S3-4) to motor (S2) reflex arch (221). The rectal examination should assess for sensation, voluntary contraction, puborectalis tone, masses the rectal vault, and stool consistency. The examination is an opportunity for education and explanation of options for various techniques of bowel care. The Consortium for Spinal Cord Injury Medicine has published recommendations on “Neurogenic Bowel Management in Adults With SCI.” These give recommendations for assessment, management monitoring, and education concerning neurogenic bowel. A summary of these recommendations is provided in the appendix of this chapter. The full text can be downloaded at PVA.org or reviewed in *The Journal of Spinal Cord Medicine* (255).

**RECTODYNAMICS**

To understand the effects of SCI on bowel evacuation, we designed a test to measure the opposing expulsive and resistive forces at work in the anorectum. We call the test rectodynamics because of its similarity to urodynamics, the study of bladder emptying (Fig. 51-10). Rectodynamics is performed by placing a triple-lumen catheter through the anus so that one pressure-measuring sideport is in the rectum and another is in the high-pressure zone of the anal canal. The third lumen is used to fill a balloon located in the rectum. A concentric EMG needle is placed in the EAS. Rectal and anal pressures, along with external sphincter pressure, are measured at rest and during stimulation of the anorectum by digital stimulation, the Valsalva maneuver, rapid rectal distention (i.e., air is rapidly injected and removed from rectal balloon to elicit the rectoanal inhibitory reflex), and slow, continuous filling of the rectal balloon with saline.
The typical features of rectodynamic study of an asymptomatic person with upper motor neuron SCI are illustrated in Fig. 51-11A. During continuous filling of the rectal balloon, the intermittent rises in rectal pressure (i.e., rectal contractions) are invariably linked with decreases in anal canal pressure. When the rectal threshold pressure is reached (20 to 30 cm H₂O), the external sphincter EMG becomes silent, and anal pressure decreases toward zero. The rectal pressure rises slowly until the rectal reservoir capacity is overcome (usually at 150 to 300 mL) and then rises rapidly until anal pressure is overcome and spontaneous evacuation of the balloon occurs.

### BOWEL EVACUATION MECHANICS

In neurologically intact people, the sequence of defecation begins with the perception of rectal fullness. This often occurs in response to a giant migratory contraction that pushes stool into the rectum. Rectal distention may be perceived at volumes as low as 10 mL by stretch receptors in the rectal wall, puborectalis muscle (257), and the pelvic floor. Continued rectal distention triggers rectal contraction and the rectoanal inhibitory reflex (258). Frenckner, while studying a small group of SCI patients, found that distention of the rectum with small volumes caused increased activity of the EAS (i.e., guarding reflex), whereas distention with large amounts caused the EAS to totally cease its electrical activity (232). This cessation of EAS activity rarely is seen in neurologically intact subjects. We have found that the volume at which the cessation of EAS activity occurs is variable and probably related to rectal capacity but that the intrarectal pressure at which it occurs is quite constant (20 to 30 cm H₂O; see Fig. 51-11A). There is a subgroup of patients with upper motor neuron lesions in whom EAS activity is aberrantly linked to rectal pressure (see Fig. 51-11B). In these patients, the spikes in rectal pressure that occur during dynamic rectal filling are met by increases in anal pressure. There is a failure to “turn off” EAS sphincter activity despite the presence of rectal pressures greater than 100 cm H₂O. We call this pattern anorectal dysynergia by analogy with detrusor-sphincter dyssynergia of the bladder. These SCI patients usually have significant difficulty with bowel evacuation.

If the choice is made to defecate, the sitting position is assumed. Sitting makes the angle between the rectum and anal canal less acute (259). In the able bodied, a rise in intra-abdominal and perhaps intrarectal pressure is generated by closing the glottis to retain a full breath in the chest. The abdominal muscles are contracted and the diaphragm is pulled down. Neurologically intact people often are able to increase intrarectal pressure by 100 cm H₂O or more with a Valsalva maneuver. Persons with upper motor neuron SCI often initiate bowel evacuation with digital stimulation. Digital rectal stimulation is done by introducing the entire length of the gloved and lubricated finger into the rectum and dilating the distal rectum by moving in a circular funnel pattern. This key unlocks the sphincter mechanism, stretches the puborectalis, and provides a stimulus for peristalsis (220,255,260). Digital rectal stimulation results in the transient loss of approximately 75% of resting anal tone and presumably further straightens the anorectal angle. Expulsive force is added by performance of a Valsalva maneuver as well as the addition of external abdominal compression. The ability to elevate intra-abdominal pressure with the use of these techniques is closely related to the level and completeness of SCI. Patients with C3 to C6 level injuries rarely can generate intra-abdominal pressures greater than 10 cm H₂O.

The most common rectodynamic finding in patients with symptomatic difficulty with bowel evacuation is a failure to generate expulsive forces that are sufficient to overcome the residual tone of the normal, fully relaxed sphincter mechanism. These patients typically have long-standing cervical cord injuries, often with associated megacolon and megarectum. In these patients, the relationship between rectal pressure and anal relaxation remains intact, although it can take rectal volumes greater than 500 mL to generate enough rectal stretch to trigger anal relaxation. Rectal volumes of 100 to 150 mL usually are sufficient to turn off the anal sphincter mechanism.

Once evacuation is initiated, the entire left colon may empty by mass peristaltic action, or the fecal bolus may be passed bit by bit. The presence of an anocolic reflex (i.e., stool passing through the anus causes colonic contraction) has been postulated but not proven. Stool consistency is probably the major determinant of the pattern of defecation. Stool consistency is of even greater importance in SCI patients, in whom the balance between expulsive and resistive forces is so tenuous.
Neurogenic Bowel Function

Supraspinal bowel dysfunction occurs with lesions rostral to the pons. Voluntary defecation depends on an accurate perception of the need to defecate as well as the necessary motor function to position over a toilet and to initiate the complex motor activity of bowel evacuation. A failure to perceive rectal fullness is common in the elderly and may manifest as overflow incontinence around a fecal impaction (261). Although disorders of continence and bowel evacuation are not uncommon after stroke, few distinct syndromes have been described. One syndrome, frontal lobe incontinence, is caused by lesions involving the anterior cingulate gyri (262). Patients with frontal lobe incontinence are said to have no awareness of bladder or rectal fullness. A pontine defecation center, analogous to the pontine micturition center, has been postulated but not proven (263). Patients with Parkinson’s disease and multiple sclerosis often have disordered bowel evacuation; however, it may be difficult to factor out the effects of immobility, diet, and medications in these cases.

Upper motor neuron bowel dysfunction occurs with lesions between the pons and the sacral spinal cord. This is the type of dysfunction most common in patients with C1 through T10 SCI. Associated deficits in function include sensory deficit for perception of fullness, a spastic pelvic floor that retains stool,
and less effective peristaltic patterns (221). The gastrocolonic response is preserved after SCI. This is the increase in peristalsis in the large and small intestine in response to a meal. The gastrocolonic response has been demonstrated to be preserved after complete SCI, although it may be somewhat less robust as compared with able-bodied controls (264). A bowel program is a unique plan designed to effectively eliminate stool and prevent incontinence. The components are fluid intake, diet, exercise, medications, and scheduled bowel care (255).

Bowel care is the procedure for assisted defecation. Through empirical manipulation of diet, hydration, medications, and stimulation techniques, a balance is maintained between continence and evacuation.

Using the rectodynamics technique, we have identified two patterns of bowel evacuation mechanics that are associated with an inability to undergo satisfactory reflex evacuation: anorectal dyssynergia and insufficient expulsive forces. In both situations, spontaneous evacuation of the saline-filled balloon rarely occurs. In asymptomatic patients, spontaneous evacuation at volumes of 150 to 300 mL is the rule. Patients with insufficient expulsive forces tend to be older and to have had their SCI longer than patients with anorectal dyssynergia. It is attractive to postulate that both patterns are different stages of the same process. Chronically obstructed evacuation from dysynergia may result in an end-stage dilated and decompensated rectum that is unable to generate a sufficient expulsive force. Against this theory is the finding that many patients with the insufficient expulsive force pattern retain a normal linkage between rectal and anal pressures, albeit at much higher volumes. Because all patients with insufficient expulsive force have cervical injuries, it is tempting to implicate a specific neurophysiologic cause, such as an overabundance of sympathetic input, although it is difficult to explain why the majority of patients with complete cervical injuries do not manifest this pattern. A more likely association with cervical injuries is that with denervation of the abdominal wall, only minimal increases in intra-abdominal pressure can be generated. Although these studies do not reveal the underlying pathophysiologic process of neurogenic bowel, they do provide a practical basis for thinking about the management of patients with neurogenic bowel.

**Gastrointestinal Problems After Spinal Cord Injury**

It has long been clear to clinicians that gastrointestinal symptoms are very common in the chronic phase after SCI. Colonic and pelvic floor function are entwined in etiquette and governed by social norms. The need for scheduled bowel care and fears of unanticipated evacuation are life limiting (256). Glickman and Kamm evaluated 115 consecutive outpatients with chronic SCI. Eighty-nine were male, the mean duration of injury was 62 months, and the level of injury was categorized as 48% cervical, 47% thoracic, and 5% lumbar. In this group, 95% required at least one therapeutic method to initiate defecation, half were dependent on others for toiletting, and 49% took more than 30 minutes to complete bowel care (265). In a prospective study carried out by the Palo Alto Veterans Affairs Medical Center Spinal Cord Injury Service, chronic gastrointestinal problems that were severe enough to alter lifestyle or require chronic treatment occurred in 27% of subjects. The limited way in which people with SCI can manifest symptoms resulted in symptoms that were characteristically nonspecific. The most common problems were difficult bowel evacuation (20%) and poorly localized abdominal pain (14%). The term *pain* was used in its most general sense and was defined as a distressing sensation in a particular part of the body. Presumably, the pain sensation was carried to the brain by means of autonomic afferents (266).

Difficult bowel evacuation was said to be present in patients who required more than 60 min/d for their bowel care or needed manual disimpaction more than once a week. When managing persons with SCI, the term *constipation* is avoided because it is and has varied implied meanings. Specific descriptions are preferred, such as prolonged defection, insufficient results, or abdominal fullness (256). The impact of problems with bowel evacuation on quality of life is formidable. In patients with difficult bowel evacuation, bowel care routinely occupies a significant part of the day, greatly restricted diets are adopted to minimize symptoms, and emergency trips to a physician for disimpaction are common. Episodes of dysreflexia, rectal bleeding, and incontinence from overtreatment with laxatives also occur frequently.

There are a few surgical options to improve success in neurogenic bowel management should conventional bowel care be ineffective (267). The appendicoceceostomy stoma is constructed by bringing the appendix through the abdominal wall. This stoma is irrigated with a 100 to 500 mL of saline to trigger defection. This procedure has been used most successfully in patient populations with spina bifida (268). A few carefully selected patients with SCI who have benefited from the catheterizable stoma have been reported (269). Colostomy is an elective procedure that can improve independence in bowel care and enhance quality of life (270–272). The decision for surgical adaptation requires interdisciplinary evaluation (255) to explore implications of surgical risk (270), body image, and functional outcome. Interdisciplinary trials of various appliances and methods for self-management can lead to optimization of stoma location and choice of equipment. Colostomy has been performed with good success in the subgroup of these patients with the most severe disability (270). In patients with severe difficulty with evacuation, unresponsive to alterations in the bowel program, colostomy can shorten the time needed for bowel care and enhance self-efficacy and quality of life (271,273).

Stone and collaborators prospectively reviewed gastrointestinal complications in a cohort of patients with acute and chronic SCI (270). Incontinence is more common in patients with lower motor neuron than upper motor neuron bowel except during diarrheal illness or as a result of overtreatment for difficult bowel evacuation. Autonomic dysreflexia arising from the gastrointestinal tract occurred in almost one half of the subjects with lesions above T6 at some time but was
rarely a significant problem by itself. Fecal impaction was the most common precipitating cause of autonomic dysreflexia. Other less common causes were massive abdominal distention and routine digital rectal stimulation. Rectal bleeding during bowel care occurred in 74% of subjects. The majority of these patients were treated with hemorrhoidal bands or expectant management. Twelve percent eventually required operative hemorrhoidectomy. Esophageal reflux symptoms were present in 24% of subjects, and 10% of patients had a proven peptic ulcer. These rates are similar to those seen in the nondisabled population (270).

Gastrointestinal problems have special significance for people with SCI. Some problems, such as autonomic dysreflexia and ventilatory insufficiency caused by massive abdominal distention, are life-threatening problems that are unique to SCI patients. Other problems are shared with the nondisabled population but are associated with disproportionate morbidity when they occur in people with SCI. Twenty-three percent of SCI patients required an admission to the hospital for evaluation or treatment of a chronic gastrointestinal problem. Absence of specific symptoms made emergency operations the rule rather than the exception for problems such as peptic ulcers and gallstones.

Stone et al. (270) further noted that chronic gastrointestinal problems usually do not develop for several years after injury. This finding is particularly significant because it suggests that these problems potentially are preventable. The situation may be analogous to that of the urinary tract in years past, when improvement in chronic care protocols has greatly reduced morbidity. For example, if chronic rectal overdistention eventually impairs the ability to evacuate, techniques to reduce rectal distention, such as more frequent bowel care, may be of benefit. In this way, decompensation of the rectum may be analogous to detrusor decompensation from chronic bladder overdistention.

Medical and Rehabilitation Management

The approach of evaluation and intervention to improve gastrointestinal function after SCI should be interdisciplinary and address all domains of disablement (221). The World Health Organization has recently redefined disablement domains that address the organ (impairment), person (activity limitation), and the person as related to the environment (barriers to participation). A spectrum of inquiry that spans organ function and task capabilities of the person and needs for full community integration will lead to the most effective treatment (221).

Utilization of the entire interdisciplinary team for assessment and goal setting integrates bowel management into the comprehensive rehabilitation plan. Overall bowel management is outlined in the bowel program, which includes the following components: diet, fluid intake, medications, physical activity, and a schedule for bowel care. Bowel care is the procedure for assisted defecation with one or more of the following components: positioning, assistive devices, rectal simulation or trigger for defecation, and supportive maneuvers (abdominal massage) (221,255).

All components of the bowel program should be designed for maximal performance. People with SCI, particularly those with tetraplegia, have very little capability to overcome resistive forces and expel stool with bowel care. Stool consistency should be titrated with diet and medications (see Fig. 51-9). Adjustments in the bowel program that may be trivial in the nondisabled population can shift the balance significantly. Stool consistency is a pivotal variable. A diet containing 15 to 30 g of dietary fiber per day should be encouraged (221,255,271). Wheat bran and psyllium make the stool more pliable by increasing the fecal water content (274). Other helpful interventions include upright bowel care to utilize gravity, abdominal binders, daily osmotic laxatives, increased frequency of bowel care, more frequent digital stimulation, and stronger triggering medications.

Avoidance of colorectal overdistention is desirable because rectal distention is known to decrease intestinal transit by the colonic reflex. Frequent bowel care (i.e., every 1 to 2 days) will avoid colonic distention with stool and may also enhance colonic transit by the hypothetical anocolic reflex. A preliminary intervention to promote bowel evacuation is movement of the sigmoid contents into the rectum. This may be accomplished by timed bowel care after a large meal to take advantage of the gastrocolonic response or by giving an oral laxative the day before. SCI persons with upper motor neuron bowel have SCI lesions above the conus and retain bulbocavernous and anocutaneous reflexes.

The basic bowel care procedure for upper motor neuron neurogenic bowel has been broken down into steps (Table 51-4) and presented in patient education booklets (255) and a film, Accidents Stink!: Bowel Care 202 (275). Rectal minienemas or suppositories and digital stimulation trigger and facilitate efficient bowel evacuation. A suppository or the contents of a minienema are first inserted as high as possible into the rectal

<table>
<thead>
<tr>
<th>TABLE 51.4</th>
<th>Steps of Bowel Care: Reflexic Neurogenic Bowel</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Getting ready and washing hands: Empty bladder.</td>
<td></td>
</tr>
<tr>
<td>2. Setting up and positioning: Transfer to a toilet or commode. If you don’t sit up, lie on your left side.</td>
<td></td>
</tr>
<tr>
<td>3. Checking for stool: Remove any stool that would interfere with inserting a suppository or minienema.</td>
<td></td>
</tr>
<tr>
<td>4. Inserting stimulant medication: Using a gloved and lubricated finger or assistive device, place the medication right next to the rectal wall.</td>
<td></td>
</tr>
<tr>
<td>5. Waiting: Wait about 5–15 min for the stimulant to work.</td>
<td></td>
</tr>
<tr>
<td>6. Starting and repeating digital rectal stimulation: To keep stool coming, repeat digital rectal stimulation every 5–10 min as needed, until all stool has passed.</td>
<td></td>
</tr>
<tr>
<td>7. Recognizing when bowel care is completed: You’ll know that stool flow has stopped if (a) no stool has come out after two digital stimulations at least 10 minutes apart, (b) mucus is coming out without stool, or (c) the rectum is completely closed around the stimulating finger.</td>
<td></td>
</tr>
<tr>
<td>8. Cleaning up—Wash and dry the perianal area.</td>
<td></td>
</tr>
</tbody>
</table>
vault and positioned against the mucosal lining. Suppositories containing bisacodyl are frequently used (276,277). They stimulate the sensory nerve endings, resulting in local and conal mediated reflex increases in peristalsis, frequently signaled by flatus. It generally takes 15 to 60 minutes for passage of the first flatus, which is followed shortly thereafter by stool flow (221). A significant reduction in the duration of bowel care can be achieved by using polyethylene glycol–based (PGB) bisacodyl suppositories rather than hydrogenated vegetable oil–based (HVB) bisacodyl (276,277). The average time to first flatus was 10 minutes with the PGB suppository compared with 37 minutes with the HVB suppository. The average total bowel-care time was 46 minutes with the PGB suppository compared with 85 minutes with the HVB suppository (278).

Suppositories containing sodium bicarbonate and potassium bicarbonate are sometimes used. With this suppository, there is a chemical reaction that releases carbon dioxide. The colonic expansion from this reaction often initiates a rectal-anal reflex. However, use of this type of suppository usually requires significant anal sphincter tone to prevent the carbon dioxide from being expelled before causing expansion within the rectal vault. Glycerine suppositories are often used as a transition from bisacodyl suppositories to no suppositories and digital stimulation. Glycerine draws water into the stool to soften it, provides a mild irritant stimulus to trigger defecation, and lubricates the rectum to reduce resistance for expulsion.

Following an adequate time for the suppository to take effect, digital stimulation is initiated. If a person is having difficulty with autonomic dysreflexia during digital stimulation, pretreatment with topical application of viscous 2% lidocaine may be beneficial. If no stool is present in the rectal vault, stimulation may be repeated a few minutes later. If stool still is not present in the rectal vault, a small-volume saline enema may be given to lubricate the rectum and create acute rectal distention. If this is still unsuccessful, a small-volume bisacodyl enema may be helpful to stimulate colorectal peristalsis.

Patients with lower motor neuron bowel typically have SCI lesions affecting the conus or cauda equina. Anal canal tone is reduced, anocutaneous bulbocavernous reflexes are absent, and the pelvic floor may passively descend. Bowel care for lower motor neuron SCI typically consists of digital rectal stimulation and manual evacuation after the morning and evening meals. A study of bowel-care patterns of persons with SCI revealed that the average frequency of bowel care in persons with lower motor neuron bowel was twice per day, and the frequency of those with upper motor neuron bowels averaged three times per week (278).

Oral medications may also be used to potentiate the success of bowel care. Oral dioctyl sodium sulfoisuccinate functions as a stool softener by having a detergent effect and increasing fluid accumulation in the stool. Oral senna or bisacodyl preparations are sometimes helpful 6 to 12 hours before bowel care. Their action is limited to the colon and felt to stimulate Auerbach’s plexus (221,271).

In patients with difficult bowel evacuation, the first step is optimization of stool consistency with fluids, fiber, oral osmotic agents, or glycerin suppositories. If possible, anticholinergic agents such as oxybutynin are discontinued. If rectodynamic studies show anorectal dysynergia, applying viscous 2% lidocaine topically and providing a prolonged anal stretch occasionally are of benefit. If the rectodynamic study shows that a megarectum already exists, large-volume enemas (300 to 500 mL) may be necessary to provide the necessary rectal distention to raise rectal pressures and shut down anal sphincter activity. In those prone to autonomic dysreflexia, careful monitoring of blood pressure is necessary.

In patients who have failed all attempts to modify their bowel program, colostomy may provide remarkable relief (221,255,266,275). We currently reserve colostomy for patients with severe long-term disability, excessively long bowel care with poor results, and exhaustion of all nonsurgical options. Transit time studies are performed before colostomy. If markers accumulate in the rectosigmoid, a sigmoid colostomy is performed. If the entire colon is acontractile, an ileostomy is performed.

Traumatic superficial mucosal erosion is by far the most common cause of bright red rectal bleeding after SCI. It usually is manifest as streaks of blood on the glove or stool. This should be distinguished from hemorrhoidal bleeding (i.e., bleeding from high-pressure vessels within hemorrhoids), which usually manifests as blood dripping into the commode or passage of clots. Traumatic erosion usually is treated conservatively. Hemorrhoids need only be treated when they cause a symptom such as bleeding, mucous soilage, or difficulty with anal hygiene. They should be kept clean, wiped with a solution or witch hazel, and supported on an air-filled cushion. A minimal evaluation for bright red rectal bleeding includes digital examination, anoscopy, and proctoscopy or flexible sigmoidoscopy.

Bright red rectal bleeding, when it occurs after SCI, tends to arise from circumferential mucosal excoriation rather than from a discrete site on a hemorrhoid. Therefore, treatment of the entire circumference of the anal canal may be necessary to eliminate bleeding. The traditional treatment has been rubber-band ligation of bleeding hemorrhoids. Patients must be aware that the likelihood of recurrent bleeding is high, that multiple applications may be necessary, and that complications, although rare, do occur. Transient episodes of autonomic dysreflexia have occurred at the time of banding, although all have been brief and self-limited (266).

Traditionally, in patients with irreducible hemorrhoidal prolapse (Grade 4), blood loss sufficient to cause anemia, or failure of rubber-band ligation, operative excisional hemorrhoidectomy is performed (279,280). A conventional surgical technique is used. To avoid manipulation of the fresh suture lines, however, we perform a full mechanical bowel preparation before surgery. No bowel care is performed for 4 days after operative hemorrhoidectomy.

A newer technique called Procedure for Prolapse and Hemorrhoids (PPH) may be an effective alternative. It involves using a surgical stapler to remove a circumferential ring of prolapsing mucosa just above the anorectal junction. In able-bodied individuals, it has been found to be faster and a much easier way to remove bleeding tissue and is associated with substantially less pain, shorter hospital stay, quicker return to normal...
activities, and voiding difficulties than a conventional excisional hemorrhiodectomy. This technique may be especially useful for those with Grade 3 (rectal prolapse is manually reducible) or Grade 4 (rectal prolapse irreducible) hemorrhoids (281).

The frequency with which rectal bleeding occurs makes screening for colorectal cancer by testing for occult blood of little value. We screen patients older than 45 years with flexible sigmoidoscopy. If polyps or tumors are seen, a full colonoscopy is performed (266).

APPENDIX: SUMMARY OF NEUROGENIC BOWEL MANAGEMENT RECOMMENDATIONS

From The Consortium for Spinal Cord Injury Medicine recommendations “Neurogenic Bowel Management in Adults With SCI” Reproduced with Permission from The Paralyzed Veterans of America. (255) (Please use the wording requested from PVA.)

The recommendations for assessment, management monitoring, and education concerning the neurogenic bowel are summarized below. The subsequent text contains the scientific evidence and supporting rationale for each recommendation.

ASSESSMENT OF THE NEUROGENIC BOWEL

Assessment of Impairment and Disability

1. A systematic, comprehensive evaluation of bowel function, impairment, and possible problems should be completed at the onset of SCI and at least annually throughout the continuum of care.

2. The patient history should include the following elements:
   • Premorbid gastrointestinal function and medical conditions
   • Current bowel program, including patient satisfaction
   • Current symptoms, including abdominal distention, respiratory compromise, early satiety, nausea, evacuation difficulty, unplanned evacuations, rectal bleeding, diarrhea, constipation, and pain
   • Defecation or bowel care (assisted defecation procedure) frequency, and duration and characteristics of stool
   • Medication use and potential effect on bowel program

3. A physical examination should be done at the onset of SCI and annually thereafter. The examination should include:
   • Complete abdominal assessment including palpation along the course of the colon
   • Rectal examination
   • Assessment of anal sphincter tone
   • Elicitation of anocutaneous and bulbocavernous reflexes to determine if the patient has upper motor neuron (UMN) or lower motor neuron (LMN) bowel
   • Stool testing for occult blood beginning at age 50

Assessment of Function (Disability)

4. An assessment of knowledge, cognition, function, and performance should be conducted to determine the ability of the individual to complete bowel care or to direct a caregiver to complete the procedure safely and effectively. The assessment should include the following elements:
   • Ability to learn
   • Ability to direct others
   • Sitting tolerance and angle
   • Sitting balance
   • Upper-extremity strength and proprioception
   • Hand and arm function
   • Spasticity
   • Transfer skills
   • Actual and potential risks to skin
   • Anthropometric characteristics
   • Home accessibility and equipment needs

MANAGEMENT OF THE NEUROGENIC BOWEL

Designing A Bowel Program

5. The bowel program should provide predictable and effective elimination and reduce evacuation problems and gastrointestinal complaints. Bowel programs should be revised as needed throughout the continuum of care.

6. Within established parameters of safety and effectiveness, the design of the bowel program should take into account attendant care, personal goals, life schedules, role obligations of the individual, and self-rated quality of life.

7. Bowel programs should be initiated during acute care and continued throughout life, unless full recovery of bowel function returns. Differences in bowel programs for reflexic and areflexic bowels include type of rectal stimulant, consistency of stool, and frequency of bowel care. To establish a bowel program:
   • Encourage appropriate fluids, diet, and activity
   • Choose an appropriate rectal stimulant
   • Provide rectal stimulation initially to trigger defecation daily
   • Select optimal scheduling and positioning
   • Select appropriate assistive techniques
   • Evaluate medications that promote or inhibit bowel function

8. A consistent schedule for defecation should be established based on factors that influence elimination, preinjury patterns of elimination, and anticipated life demands.

9. Prescribe mechanical and/or chemical rectal stimulation to predictably and effectively evacuate stool.

10. The use of assistive techniques should be individualized and their effectiveness in aiding evacuation should be evaluated. Push-ups, abdominal massage, Valsalva maneuver,
deep breathing, ingestion of warm fluids, and a seated or forward-leaning position are some of the techniques used to aid in bowel emptying.

**Nutrition**

11. Individuals with SCI should not be placed uniformly on high-fiber diets. A diet history should be taken to determine the individual’s usual fiber intake. The effects of current fiber intake on consistency of stool and frequency of evacuation should be evaluated. A diet containing no less than 15 g of fiber daily is needed initially. Increases in fiber intake should be done gradually, from a wide variety of sources. Symptoms of intolerance should be monitored, and reductions in fiber are recommended, if they occur.

12. The amount of fluid needed to promote optimal stool consistency must be balanced with the amount needed for bladder management. In general, fluid intake should be approximately 500 mL/d greater than the standard guidelines used to estimate the needs of the general public (National Research Council, 1989). Standard guidelines indicate that adult fluid needs can be estimated by either of the following formulas:

\[
1 \text{ mL fluid/Kcal of energy needs} + 500 \text{ mL/d}
\]

or

\[
40 \text{ mL/kg body weight} + 500 \text{ mL/d}
\]

**Managing The Neurogenic Bowel at Home or in the Community**

13. Appropriate adaptive equipment for bowel care should be prescribed based on the individual’s functional status and discharge environment.

14. Careful measures should be taken to avoid pressure ulcers and falls related to the use of bowel care equipment.

15. Adequate social and emotional support should be available to help individuals manage actual or potential disabilities and handicaps associated with neurogenic bowel.

16. All aspects of the bowel management program should be designed to be easily replicated in the individual’s home and community setting.

**Monitoring Program Effectiveness**

17. The following variables should be monitored during and documented after every bowel care procedure during hospitalization or when developing or revising a bowel program in any community setting:

- Date and time of day
- Time from rectal stimulation until defecation is completed
- Total time for completion of bowel care
- Mechanical stimulation techniques
- Pharmacological stimulation
- Position/assistive techniques

18. When a bowel program is not effective (i.e., if constipation, GI symptoms or complications, or unplanned or delayed evacuations occur) and a consistent schedule has been adhered to, changes in the following components should be considered:

- Diet
- Fluid intake
- Level of activity
- Frequency of bowel care
- Position/assistive techniques
- Type of rectal stimulant
- Oral medications

19. In the absence of adverse reactions and indicators for potential medical complications, the bowel care regimen should be maintained for three to five bowel care cycles prior to considering possible modifications. Only one element should be changed at a time.

20. When evaluating individuals complaining of bowel management difficulties, adherence to treatment recommendations should be assessed.

21. Colorectal cancer must be ruled out in individuals with SCI over the age of 50 years with a positive fecal occult blood test or with a change in bowel function that does not respond to corrective interventions.

**REFERENCES**


Pressure Ulcers

SCOPe OF THE PROBLEM

Definition

Tissue necrosis and breakdown is referred to by many terms including decubitus ulcers, bedsores, ischemic sores, and pressure sores. The term “pressure ulcer” is the most accurate nomenclature to describe both the cause and nature of soft tissue wounds, caused primarily by excessive applied pressure. This term will be used throughout this chapter. A critically important precursor to appropriate care is an accurate identification and classification of a soft tissue wound as a pressure ulcer. This definition has recently been updated by the National Pressure Ulcer Advisory Panel (NPUAP) to be a localized injury to the skin and/or underlying tissue usually over a bony prominence, as a result of pressure, or pressure in combination with shear and/or friction (1). This change takes into account an increasing body of evidence that pressure ulcers are generally multifactorial in origin. Furthermore, pressure ulcers are frequently associated with deep tissue injury that can be difficult to detect on initial examination, particularly in individuals with darker skin tones. The NPUAP has, therefore, also updated the guidelines for staging in order to provide clinicians with a standardized and reliable framework for classification (see “Assessment” below). Appropriate and effective care of pressure ulcers is essential in order to minimize further medical complications that can include amyloidosis, osteomyelitis, septicemia, and even death.

Economic Implications

The presence of one or more chronic pressure ulcer has a significant negative impact, both on the individual’s health status and on the overall health care system. The total economic impact is hard to quantify. The most recent figures in the United States date from 1999 when the treatment costs were estimated to be in excess of $1.3 billion annually (2). More recently, in 2004, Bennett et al. (3) found that the annual cost of treating pressure ulcers in the United Kingdom was approximately 4% of the total health care expenditure (1.4 to 2.1 billion GBP), equivalent to around $4 trillion per annum in February 2008. This major trend for increasing costs may continue as the overall population at risk for chronic wounds increases with changing demographics, in particular an increasing aged population with reduced mobility. Cost containment will require continuous improvements in clinical care for early identification, prevention, and treatment of pressure ulcers. Lyder et al. (4) found that implementing comprehensive prevention programs can significantly reduce the incidence but that costs are not significantly reduced. Prevention intervention strategies should, therefore, be based on risk stratification.

Incidence and Prevalence

Pressure ulcers are a very common cause of morbidity and even mortality. No central registry exists and many pressure ulcers may go unreported, leading to higher rates of both morbidity and mortality (5). Incidence and prevalence vary widely, both between studies and across different health care settings. For example, in the United States the prevalence of pressure ulcers in acute care hospitals has been reported to be between 3% and 69%, with annual incidence rates of 1% to 29% (6–11). More recent studies in Swedish and Canadian acute health care settings have found the overall prevalence rates of 29% to 33% (12,13).

In the U.S. long-term care facilities and nursing homes, prevalence rates from 1% to 35% have historically been noted, with annual incidence rates of 10% to 38% (14–16). Brandeis et al. reported that pressure ulcer development was associated with a 4.5 times greater risk of death for elderly nursing home residents than for those who did not have pressure ulcers but had the same risk factors (17). The prevalence of pressure ulcers among the outpatient population has not been well-defined, although studies have reported similarly varying rates (18).

These figures should be viewed cautiously; some variations, particularly in smaller studies with lower reported incidence rates, may be due to factors such as pressure ulcer definition, study population characteristics, and data source, for example, retrospective chart review or self-report (19). For example, in some studies over 45% of reported pressure ulcers were classified as Stage 1, whereas in others Stage I pressure ulcers were not included. Although Stage I pressure ulcers may not appear clinically severe, and are generally readily treatable, it has been found that these ulcers are an important warning of potential future deterioration (20), and they should therefore always be taken seriously.

Although there is paucity of recent, reliable, and widely available data for U.S. health care facilities, there is a growing assumption, particularly reflected in regulatory and quality management guidelines, that the development of any pressure ulcer is evidence of negligent care by a health care provider or health care system (21). Increasing public awareness has provided the perception that pressure ulcers are preventable...
and readily treatable. The Centers for Medicare & Medicaid Services (CMS) guidelines mandate that any resident who enters a nursing home without a pressure ulcer must not develop pressure ulcers unless the individual’s clinical condition is such that they were “demonstrably unavoidable” (22). Any resident who has a pressure ulcer must receive necessary treatment to promote healing, prevent infection, and prevent new ulcers from developing. Both federal and state surveyors ensure compliance with all aspects of the mandate. Penalties for noncompliance and/or violations can be severe, including fines, loss of financial support, and even immediate closure of the facility. This environment dictates that all clinical practitioners be highly educated in all aspects of pressure ulcer care for their patients.

ANATOMY AND PATHOPHYSIOLOGY: NORMAL SKIN

The skin is the largest organ in the integumentary system, the external covering of the body. It has multiple functions, including protection of the body from trauma, dehydration, and microorganisms; thermoregulation, excretion of waste products through perspiration, and sensory perception. The skin also produces vitamin D through exposure to ultraviolet light. Therefore, it is vital to our health and damage to the skin may have potentially serious consequences.

There are two main layers to the skin: the epidermis and the dermis. There are subcutaneous tissues underlying the dermal layer. The dermis is about four times the thickness of the epidermis and consists of collagen fibers, blood vessels, lymphatic vessels, sweat glands, hair roots, and sensory nerves. Embryologically, this layer is derived from the mesoderm. The main function of the dermis is to support the epidermis, and provide strength and elasticity to the skin. There are two main layers in the dermis, the superficial papillary dermis and the deeper reticular dermis. The papillary dermis consists of loose connective tissues, capillaries, and elastic fibers, while the reticular dermis consists of dense connective tissues, elastic fibers, larger blood vessels, fibroblasts, mast cells, sensory nerves, and lymphatic vessels. This layer is essential for blood supply and sensory perception. The subcutaneous tissues lie between the dermis and the underlying muscle and bone. They consist of adipose tissues and their main function is to provide insulation and storage of nutrients.

The blood supply of the skin comes from the underlying structures, supplying the various structures in the skin along the way, and is crucial for thermoregulation. The blood supply from the bone travels into the skin, supplying the muscles, fascia, nerve, and connective tissues. The blood vessels in the skin anastomose and form a horizontal network of blood vessels in the dermis. Blood flow to the skin is very high and allows effective heat dissipation from the skin through vasodilatation of the blood vessels. Conversely, vasoconstriction in the skin will lead to diminished blood flow and thus reduced heat loss and a cool skin surface.

WOUND HEALING

Normal physiological wound healing is a complex, dynamic process that can be defined as comprising four main phases: hemostasis, inflammation, proliferation, and tissue remodeling (23,24) (Fig. 52-1). Inflammation can be further divided into early and late inflammation. Similarly, proliferation consists of both granulation tissue and extracellular matrix (ECM) formation. The entire sequence of events exhibits substantial overlap of the various phases, with each phase characterized by different predominant cell types.

Immediately following injury, the healing response begins with the activation of the coagulation cascade and the creation of a blood clot (25), producing hemostasis and preventing further blood loss (26). Various chemotactic factors, including platelet-derived growth factor (PDGF), stimulate the migration of blood neutrophils and monocytes, helping to initiate the second phase of healing, the inflammatory response (27).

Inflammation is responsible for containing, neutralizing, diluting, or walling off the injurious agent or process (28). Acute inflammation lasts only minutes to days depending on the extent of injury as leukocytes, predominately neutrophils (28), invade the wound and clear it of debris (25,27). As the neutrophils diminish, the wound moves to the chronic inflammation phase with increased activity of macrophages and lymphocytes. Macrophages produce several biologically active enzymes and cytokines, including collagenases that phagocytose cell debris, senescent neutrophils, and devitalized tissue in the wound region. This debrides the wound and promotes wound healing. Interleukins and tumor necrosis factor (TNF) concurrently promote angiogenesis and stimulate fibroblasts activity (28), leading to the wound transitioning to the proliferative phase.

Granulation tissue formed during the proliferative phase of healing provides a temporary weak tissue layer in the wound. Granulation tissue is so called because of the pink, soft granular surface appearance (27–29). New blood vessels are produced by budding or sprouting from pre-existing vessels in a process known as neovascularization or angiogenesis (28).
Angiogenesis is essential for granulation tissue formation and improves the overall tissue perfusion (30). Another key aspect of the proliferation phase is reepithelialization: Migration and proliferation of keratinocytes from the wound edges leads to the formation of a thin epithelial layer over the developing granulation tissue (27). In a full thickness, nonsutured wound the rate of reepithelialization is thus inversely related to the wound size. Wounds left uncovered and open to the air develop a layer of eschar, or scab, over the granulation tissue. Although this layer of dried wound exudate and dead cells protects the wound from environmental contamination, it also prevents keratinocyte migration so that reepithelialization cannot occur. Reepithelialization may thus be accelerated by preserving an optimally moist environment.

Collagen repair starts in the proliferative phase and extends into the final remodeling phase. Remodeling of the ECM occurs as granulation tissue is reduced via apoptosis (31). Collagen is constantly being synthesized and degraded in order to gain strength through the reorientation of collagen fibers, as wound tissue approaches full strength (25). Collagen deposition in normal wound healing reaches a maximum by 21 days after the wound is created (27); however, it may take up to 2 years or more for a scar to completely form. If the remodeling process is disrupted, the equilibrium between ECM deposition and degradation is lost, leading to the development of fibrosis and excessive scarring (32). Scar tissue contains fewer cells than normal skin and never achieves the same tensile strength as the original skin it has replaced (27,29).

**TABLE 52.1** Risk Factors in Pressure Ulcer Development

<table>
<thead>
<tr>
<th>Extrinsic</th>
<th>Intrinsic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Applied pressure</td>
<td>Muscle atrophy</td>
</tr>
<tr>
<td>Surface shear</td>
<td>Impaired nutritional status</td>
</tr>
<tr>
<td>Friction</td>
<td>Anemia</td>
</tr>
<tr>
<td>Local microenvironment</td>
<td>Impaired vascular status</td>
</tr>
<tr>
<td>Psychosocial</td>
<td>Impaired mobility</td>
</tr>
<tr>
<td></td>
<td>Impaired sensation</td>
</tr>
</tbody>
</table>

**PRESSURE ULCER PATHOPHYSIOLOGY**

The term “pressure ulcer” may be misleading in that it implies that pressure is the only risk factor. In fact, the characteristic tissue breakdown and a subsequent failure to heal are well recognized to be multifactorial processes. For example, there is increasing evidence that shear forces are critically important. Clinical evidence regarding risk factors has been mostly correlative rather than causative (33), although animal models and in-vivo studies are providing further details. It can be useful to consider risk factors broadly as either extrinsic, that is, external to the patient’s body, or intrinsic, that is, within the patient’s body (Table 52-1). Psychosocial factors are complex and probably can be classified as both intrinsic and extrinsic.
Thorough consideration of the risk factors is of vital importance because they contribute to the development of treatment and rehabilitation strategies. In general, extrinsic risk factors respond more rapidly to clinical interventions, such as changing the seating system used, but this is not always feasible or easily optimized, for example, intraoperative pressure relief. Changes in clinical status can alter intrinsic factors and increase the risk of pressure ulcer development. For example, urinary incontinence will alter the microenvironment of the skin surface and make it more susceptible to maceration and breakdown.

Given these risk factors, it is not surprising that certain populations are especially prone to pressure ulcer development. These include the elderly with reduced mobility, hospitalized individuals with complex medical conditions, those with cognitive impairments, and those with motor and sensory impairment such as spinal cord injury.

Pressure

Applied isostatic pressure compresses the soft tissues equally along all axes and causes little or no tissue damage. However, in most real-life situations, applied pressure is not isostatic and the soft tissues will compress along the axes of loading. As applied pressure increases the local capillaries, deeper vasculature will become progressively occluded, impeding blood flow. Lymphatic circulation will be impaired, reducing local drainage of intracellular waste materials and increasing the risk of edema.

The soft tissues can be defined as a composite viscoelastic material, in other words, the response to applied loads varies with the duration of application. Relatively high loads can be tolerated for a short period of time, with complete elastic recovery on removal of the load. Over longer periods of applied load, the viscous nature of the soft tissues predominates, leading to irreversible tissue damage and increased full recovery time for reversible damage. Specific times and magnitudes of critical applied loads vary individually, dependant on the resilience of the local tissues, which is, in turn, dependant on intrinsic factors (34) such as those described in this chapter.

Removal of applied loads after a period of vascular occlusion results in a transient increase in regional blood flow. The duration of this phenomenon, known as reactive hyperemia, is directly related to the preceding duration of occlusion. During the hyperemic phase, tissue damage can occur due to oxygen-free radical activity, which is incompletely buffered in tissue under oxidative stress (35–37). This has led to the proposal that repetitive ischemia/reperfusion injury is a significant pathway to the development of chronic pressure-related wounds. Animal studies have found that repetitive ischemia/reperfusion causes more severe tissue damage than a single, prolonged ischemic insult (38,39). The response to cyclic loading varies for metabolic and myogenic components of skin blood flow, suggesting that optimization of alternating pressure support surfaces requires personalized evaluation of intrinsic blood flow (40).

Many clinicians have observed that patients may appear to have a minor wound or even only some localized discoloration, which then rapidly breaks down to reveal extensive subcutaneous damage, even down to and including the bone. Such wounds are the result of deep tissue injury, that is, tissue breakdown that has its origin in the soft tissues nearer to the bone rather than to the skin surface. Theoretical models using finite element analysis have indicated that peak pressures under applied load occur at the bone-muscle interface (41) and that the magnitude is strongly dependant on the local geometrical and biomechanical characteristics (42). Recent human studies have shown that internal tissue loads are significantly higher in individuals with paraplegia compared to age-matched able-bodied individuals (43), and that postural changes significantly affect these loads. These methods have yet to be applied to other at-risk populations, such as the elderly; however, similarities in tissue biomechanical properties would be predicted.

Shear and Friction

Shear force is defined generally as a force or a component of force acting parallel to the plane. For a patient, shear forces can originate from nonuniform intrinsic pressure distributions, surface shear, or tangential forces at the body/support interface and cause transverse tissue distortion, that is, shear stresses and strains. Although it has been found that shear force alone does not appear to induce tissue breakdown (44), shear in combination with normal applied pressure will occlude blood vessels at much lower applied forces than normal applied pressure alone (45). Shear strains within the tissue can also disrupt tissue layers. Shear is thus widely considered to be the most important risk factor because blood flow occlusion and tissue damage occurs much more rapidly when shear stresses and strains are present. In recognition of this, the NPUAP convened the “Shear Force Initiative” in 2005 (46). To date, this group has not issued specific guidelines and the direct measurement of shear forces in clinical practice remains challenging.

If shear forces are overcome, slip will occur between the two surfaces. This friction may result in both epidermal abrasion and subepidermal blistering (47,48). Friction alone may not cause pressure ulcers, but in combination with other risk factors such as a compromised microenvironment, it would appear to heighten the risk of rapid tissue damage.

Immobility

Impaired mobility is commonly due to paralysis or other concurrent medical conditions, leading to an increased risk of pressure ulcer development, especially in the hospital setting (49,50). Motor paralysis will directly affect a person’s ability to respond unconsciously to potential noxious stimuli, for example, fidgeting while sitting or turning while asleep. Reduced mobility also profoundly alters the individual’s ability to consciously perform postural maneuvers necessary to relieve prolonged applied pressure, from weight shifting while sitting to walking. The loss or reduction of mobility may be further
complicated by sensory impairment, leading to the absence or alteration of normal perception of environmental stimuli such as pain or temperature. Patients with impaired sensation and/or proprioception are at an increased risk of pressure ulcer development because the individual cannot sense the warning signals that precede tissue damage.

**Nutrition**

Impaired nutritional status is another significant risk factor for the development of pressure ulcers (51,52). Moreover, the malnourished patient will also have an impaired response to healing. Protein depletion may lead to decreased perfusion and impaired immune response. Low serum protein level may be associated with interstitial edema, leading to decreased tissue elasticity, hence reducing tissue transportation of oxygen and nutrients from the blood to the skin (53). Increased tissue pressure as a result of interstitial edema may decrease blood flow to the skin, thus increasing the risk of pressure ulcer development (54). Furthermore, the presence of an exuding pressure ulcer may cause protein loss and the patient may then develop negative nitrogen balance.

Nutritional status can be measured by various means: food intake, body weight, body mass index (BMI), serum total protein, albumin, and serum prealbumin. Serum total protein level less than 6.4 g/dL has been associated with the development of pressure ulcers (10). Serum albumin level less than 3.5 mg/dL has also been associated with pressure ulcer development (55). Serum prealbumin has a shorter half-life than serum albumin, and therefore, is a more sensitive indicator for measuring nutritional status. The Mini Nutritional Assessment (MNA) and Mini Nutritional Assessment Screening Form have been found to be advantageous in the screening of nutritional status in elderly people with pressure ulcers (56,57). The severity of a pressure ulcer can be directly related to the degree of hypoalbuminemia (58). Fluid balance must also be considered in conjunction with nutritional status since dehydration will decrease the cellular nutrient delivery.

**Local Microenvironment**

The local environment at the skin-support interface, termed the microenvironment, has a localized effect on the risk to pressure ulcer development. This umbrella term covers several factors including temperature, moisture, and acidity.

Skin temperature provides an indicator of tissue perfusion. It has been found that immobile patients do not recover preloading skin temperatures after a period of prolonged loading (59). However, elevated body temperatures will raise metabolic activity increasing the need for cellular oxygen. Animal studies have found that deep tissue damage occurs at 10° above room temperature whereas even higher temperatures caused both cutaneous and subdermal damage (60). In some patient populations, such as those with SCI, dysfunctional temperature regulation may produce abnormal variations in microenvironmental temperature.

Moisture from sweat or incontinence will both superhydrate the epidermis and alter the acidity. Skin maceration will both soften the stratum corneum and increase the friction coefficient of the epidermis. The net effect is rapid destruction of the epidermis and reduced antibacterial functionality (61). Both sweat lactate and urea concentrations become elevated with prolonged ischemia (62,63), implying that sweat analysis may be a useful predictive indicator of early tissue damage and pressure ulcer development.

**Anemia**

Anemia has been identified as a risk factor in both the general (64) and the spinal cord literature (65,66). Hemoglobin level below 12.0 to 14.0 g/dL has been associated with the development of pressure ulcers (67). The literature suggests a correlational relationship, but presumably anemia will lead to impaired delivery of oxygen and nutrients to the local tissue, affecting both the development and the healing of pressure ulcers.

**Impaired Vascular Status**

Impaired vascularity of the soft tissues is increasingly recognized as a risk factor for the development of pressure ulcers. Tissue vascularity can be impaired as a result of different pathophysiology, for instance, the impairment of microcirculation due to diabetes mellitus (68), and ischemia to deep tissues as a result of compression or pressure. In fact, there is increasing evidence that deep tissue injury is a significant risk factor in the development of pressure ulcers, especially in conjunction with other risk factors (69). This concept is further discussed elsewhere in this chapter.

**Muscle Atrophy**

Muscle atrophy is a common consequence of aging and neuromuscular conditions, such as spinal cord injury and stroke. It may be due to muscle denervation, or simply due to disuse. Individuals with muscle atrophy are particularly at high risk of developing pressure ulcers because it leads to an increase in the interface pressure between the weight-bearing skin areas and the support surfaces. For instance, the gluteal muscles provide cushioning effect between the ischial tuberosity and the seating surface. Gluteal muscle atrophy following spinal cord injury will increase the interface pressure between the ischial skin area and the support surfaces such as the wheelchair cushion or the mattress. When this problem is compounded by other risk factors such as immobility, it puts the individual at high risk of pressure ulcer development.

**Psychosocial/Lifestyle**

Some psychosocial factors and lifestyle can be readily addressed by clinicians; however, there is often a need to address non-clinical problems (70). For example, smoking cessation and substance abuse programs can improve the overall physical and mental health. Other issues, such as familial support, access to specialized health care, and independence of community mobility impact social integration. Effective management requires a multidisciplinary team approach.

---

**PRESSURE ULCERS**

---

**CHAPTER 92 | PRESSURE ULCERS**
### RISK FACTOR MEASUREMENT

Many of the risk factors described in the previous section can be evaluated using standard clinical procedures, such as blood work to determine the nutritional status. Quantitative measurement of specific pressure ulcer risk factors, such as interface pressures and blood flow, require more specialized approaches.

#### Interface Pressure Mapping

Interface pressure mapping has become more widely used as an assessment tool in wheelchair and seating clinics (71,72). Patients with complex seating needs and those requiring repeated evaluations routinely receive interface pressure assessments as part of the review of both posture and function. Several commercially available systems are reliable, accurate, repeatable, and user-friendly. All these systems include a pressure sensor mat placed between the patient and the support surfaces. This mat is thin and flexible in order to minimize disruption of the interface being measured and contains an array of sensing elements (sensels) that are rapidly scanned to complete a detailed image of the regional pressure distribution. The resulting pressure map is displayed in real-time on a computer screen and provides clinical information, such as location of high pressure points. In addition, real-time pressure maps can be used as biofeedback for patients, who can see the effects of various cushions and weight-shifting maneuvers.

#### Blood Flow Measurement

The measurement of interface pressures alone is not sufficient: Maintenance of tissue health requires an adequate supply of nutrients from the blood. Blood flow measurement is thus important for quantitative identification of pressure ulcer risk. Various noninvasive techniques have been used to measure blood flow under applied load.

Transcutaneous blood gas electrodes measure the partial pressure of oxygen and/or carbon dioxide in the skin. Localized heating ensures maximal vasodilation so that capillary blood gas tension is equal to arterial tension. Multisensor systems allow the simultaneous measurement of several at-risk locations. Transcutaneous blood gas measurement has been used in studies to determine both pressure ulcer risk (63,73,74) and healing (75,76).

Laser Doppler flowmetry provides noninvasive measurement of skin blood flow at a depth of about 1 mm via laser and fiber optics technology, and has been used to assess the effects of both prolonged and alternating loading (40,77).

Tissue reflectance spectroscopy utilizes the absorption spectrum of visible light to quantify the blood content and oxygenation of the skin. Sensors have been designed that are suitable for measurements at the support interface (78); however, use is currently mainly in research settings (79,80).

#### Soft Tissue Imaging

The use of soft tissue imaging to identify pressure ulcer risk and/or measure pressure ulcers has mostly been limited to research studies. Both CT and MRI can show deep tissue damage (81,82) but are limited in clinical applicability. High frequency ultrasound has been reported to show specific progression of pressure ulcer development (83); however, it would appear that a significant level of expertise is required for reliable interpretation.

### ASSESSMENT

Scales have been developed to provide clinicians with structured tools for pressure ulcer risk assessment and wound healing evaluation. These scales assign ordinal values to multiple qualitative and quantitative measures. Thresholds are defined that determine a positive or negative outcome, that is, high risk/low risk or healing/nonhealing.

#### Pressure Ulcer Risk Assessment Scales

Risk assessment scales are designed to be broadly applicable to patients with varying diagnoses in any health care setting. The goal is to provide a readily useable tool to identify patients needing preventative interventions so that limited resources can be most effectively allocated. Even the most streamlined tool requires some degree of observer training.

The three most widely used scales are the Braden (84), Norton (85), and Waterlow (86) scales. Of these, the Braden scale has been most widely used and validated in several different settings (87–89), although it is not always recommended as the optimal tool for clinical management, particularly in acute care settings (90,91). There are marked differences between these scales (92) and it should be borne in mind that subscales cannot necessarily be used independently with the same validity. Modifications of scales for specific settings may improve the predictive value at a particular location (93), but testing and observer training is required for widespread implementation.

The use of a pressure ulcer risk assessment scale should not override all clinical judgement; specific patient factors may affect risk in ways not captured by a scale and must always be considered in an individual plan of care.

#### Staging

The intent of staging is to provide an objective description of the severity of the pressure ulcers. Up until 2007, there had been multiple proposed systems, but the most accepted one was the four-stage system proposed by the NPUAP in 1989 (94). The NPUAP system was based on the layers of the tissues involved. It was derived from the staging system proposed by Shea in 1975 (95). Subsequently, it was revised in 1997 to specifically address the description of Stage I ulcers so that variables such as skin tone, temperature, and tissue consistency would be addressed. The NPUAP staging system has been endorsed by the Wound Ostomy & Continence Nurses Society (WOCN) and the Agency for Health Care Policy & Research (AHCPR) (96). The NPUAP staging system was not meant to be used for other skin wounds, such as diabetic foot ulcer, or arterial ulcers. It must be noted that the staging system is not based on the...
natural history of pressure ulcer development. The four stages described are not necessarily based on histopathology or the progression of the pressure ulcers. This widely adopted staging system was revised in 2007. In order to understand the changes in the new system and the rationale behind the changes, it is important to compare the two systems (Table 52-2).

As the pressure ulcer improves, the depth and size may decrease. However, it is not appropriate to "reverse stage" (97) the ulcer, because the staging system was not designed to indicate the healing process. This would lead to the assumption that full-thickness pressure ulcers heal by replacing the same structural layers of tissues that were lost during the development of the pressure ulcer. As full-thickness pressure ulcers heal and become shallower in depth, they do not replace lost muscle, subcutaneous fat, and dermis before they reepithelialize. Rather, the defect is filled with granulation tissue of a rather different composition (endothelial cells, collagen, fibroblasts, and ECM). Therefore, it would be inaccurate and misleading to describe a healing Stage IV ulcer to become Stage III, then Stage II as it starts to heal and become less deep. It is appropriate to describe a "healing Stage IV ulcer" with the accurate size, depth, and description of the other wound characteristics, as one should do with the comprehensive assessment of any pressure ulcer.

**Pressure Ulcer Healing Evaluation Scales**

Several wound evaluation tools have been developed that combine qualitative assessment of wound status with wound size measurement. These include the Pressure Ulcer Scale for Healing (PUSH) (98,99), the Pressure Sore Status Tool (PSST) (100), and the Leg Ulcer Measurement Tool (LUMT) (101). Observer training is necessary for reliable implementation of these tools. Qualitative measures may have high interobserver variability, due to subjective biases.

The PUSH tool is the most extensively validated tool for assessment of pressure ulcer healing in clinical settings (102,103). Wound size, defined as surface area, is the predominant factor in this tool. Careful wound size measurement is essential to maintain the overall accuracy, particularly for repeated measures over extended time periods. A recent survey found that the PUSH tool is generally considered to be reliable and easy to use, although respondents were less positive regarding its usefulness (104).

**Wound Measurement**

In order to evaluate and monitor the progress of a pressure ulcer over time, accurate wound measurement is essential. Linear measurements of length (maximum length measured in the axis from head to toe), width (maximum width perpendicular to the length), and depth are commonly used in clinical practice. However, they may not be the ideal measurement method for pressure ulcers that can be rather large, deep, and irregular in size and shape (105). Needless to say, linear measurements may be operator-dependent and have low inter-rater reliability. Therefore, alternative wound measurement methods that provide high accuracy and good inter-rater reliability are needed. Digital wound measurement devices can potentially satisfy these criteria.

Two digital imaging techniques that have been introduced for clinical use are planimetry and digital image analysis technique, photogrammetry. Planimetry measures the surface area of a pressure ulcer by tracing its boundary, with the traced area digitally calculated by a computerized device (e.g., The Visitrak system, Smith & Nephew, Largo, FL, USA), whereas photogrammetry utilizes digital photography of the pressure ulcer to measure its surface area (e.g., The VeV MD, Verge Videometer Measurement Documentation system, Vistamedical, Manitoba, Canada). Both digital techniques were shown to have higher accuracy than linear measurement (106,107). However, their clinical applicability may be limited by their cost and the increased time needed to produce the measurements. Furthermore, they still do not measure wound volume, which is the most accurate measurement for pressure ulcers. All of these factors should be considered when deciding which measurement method is to be used.

**PRINCIPLES OF TREATMENT**

The treatment of pressure ulcers is a complex process. It starts with the prevention of their development. This includes the identification, and subsequently the correction of risk factors, as discussed above. Then specific treatment of a pressure ulcer aims to correct the local wound environment, in order to provide the most optimal condition for healing. This includes the selection of appropriate dressings, management of the wound bed, use of specific modalities to promote wound healing, as well as the specific methods of pressure relief. As a general rule, conservative management is the first line of treatment, with surgical intervention being considered after adequate conservative treatment fails in Stage III or IV pressure ulcers. There are multiple, comprehensive, practical guidelines on the management of pressure ulcers, including the guidelines by the Wound, Ostomy and Continence Nurses Society (108), the Consortium for Spinal Cord Medicine (109), the AHCPR (96), the Registered Nurses Association of Ontario (110), and the Royal College of Nursing (111), to name but a few. The goal of this section is to provide the principles and scientific basis of the overall management of pressure ulcers.

**Medical Treatment of Underlying Systemic Factors**

The management of the intrinsic and extrinsic risk factors is crucial in both the prevention and treatment of pressure ulcers. Correction of some of the stated risk factors is self-evident and will not be discussed in this section. We will, therefore, focus on specific risk factors that require further discussion.

**Nutrition**

The management of nutrition has been widely recognized as an essential part of pressure ulcer prevention and management,
**TABLE 52.2** Comparison of NPUAP Staging Systems (1989 and 2007)

<table>
<thead>
<tr>
<th>NPUAP Staging System (1989)</th>
<th>2007 Revision</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Suspected Deep Tissue Injury:</strong></td>
<td>Purple or maroon localized area of discolored intact skin or blood-filled blister due to damage of underlying soft tissue from pressure and/or shear. The area may be preceded by tissue that is painful, firm, mushy, boggy, warmer, or cooler as compared to adjacent tissue.</td>
</tr>
<tr>
<td><strong>Further description:</strong></td>
<td>Deep tissue injury may be difficult to detect in individuals with dark skin tones. Evolution may include a thin blister over a dark wound bed. The wound may further evolve and become covered by thin eschar. Evolution may be rapid, exposing additional layers of tissue even with optimal treatment.</td>
</tr>
<tr>
<td><strong>Stage I:</strong></td>
<td>Intact skin with nonblanchable redness of a localized area usually over a bony prominence. Darkly pigmented skin may not have visible blanching; its color may differ from the surrounding area.</td>
</tr>
<tr>
<td><strong>Stage II:</strong></td>
<td>Partial thickness loss of dermis presenting as a shallow open ulcer with a red pink wound bed, without slough. May also present as an intact or open/ruptured serum-filled blister.</td>
</tr>
<tr>
<td><strong>Stage III:</strong></td>
<td>Full thickness tissue loss. Subcutaneous fat may be visible but bone, tendon, or muscle are not exposed. Slough may be present but does not obscure the depth of tissue loss. May include undermining and tunneling.</td>
</tr>
<tr>
<td><strong>Stage IV:</strong></td>
<td>Full thickness tissue loss with exposed bone, tendon, or muscle. Slough or eschar may be present on some parts of the wound bed. Often include undermining and tunneling.</td>
</tr>
</tbody>
</table>

*Stage I:* An observable pressure-related alteration of intact skin whose indicators as compared to the adjacent or opposite area on the body may include changes in one or more of the following: skin temperature (warmth or coolness), tissue consistency (firm or boggy feel), and/or sensation (pain, itching). The ulcer appears as a defined area of persistent redness in lightly pigmented skin, whereas in darker skin tones, the ulcer may appear with persistent red, blue, or purple hues.

*Stage II:* Partial thickness skin loss involving epidermis, dermis, or both. The ulcer is superficial and presents clinically as an abrasion, blister, or shallow crater.

*Stage III:* Full thickness skin loss involving damage to, or necrosis of, subcutaneous tissue that may extend down to, but not through, underlying fascia. The ulcer presents clinically as a deep crater with or without undermining of adjacent tissue.

*Stage IV:* Full thickness skin loss with extensive destruction, tissue necrosis, or damage to muscle, bone, or supporting structures (e.g., tendon, joint capsule). Undermining and sinus tracts also may be associated with Stage IV pressure ulcers.
with a focus on protein and micronutrients such as vitamins and amino acids. Dietary protein is very important in the healing of pressure ulcers in malnourished patients (112). It has been suggested that the need for protein is often underestimated in those with wounds (113). The AHCPR guidelines recommend that individuals without a pressure ulcer should take 1.0 to 1.25 g/kg of body weight/day of protein, while those with a pressure ulcer should take 1.25 to 1.5 g/kg/day. The Consortium of Spinal Cord Medicine (109) guidelines recommend 1.2 to 1.5 g/kg/day for those with a Stage II pressure ulcer, and 1.5 to 2.0 g/kg/day for those with a Stage III or Stage IV pressure ulcer. It must be noted that the nutritional needs should be adjusted according to the presence of any medical comorbidity.

The role of micronutrients such as vitamin A, C, E, zinc, and arginine in the treatment of pressure ulcers is less clear. Arginine is a nonessential amino acid. It has been shown to enhance the wound strength and collagen deposition in both the human and rodent models, with artificial incisional wounds (114). L-Arginine is the only substrate for nitric oxide synthesis. In diabetes mellitus, there is a decrease in nitric oxide at the wound site, and this is accompanied by impaired healing. In the diabetic rat model, the supplementation of L-arginine has been shown to improve wound healing with increased nitric oxide synthesis (115). Though some research trials have demonstrated that the use of vitamin C, zinc, and arginine may be helpful in the treatment of pressure ulcers (116,117), further randomized, controlled trials with larger sample size are needed to confirm these findings (118,119).

**Impaired Mobility**

Impaired mobility is a significant problem in those with neuromuscular weakness and in postsurgical patients whose mobility may be temporarily impaired. It is a constant risk factor as the patient has unrelieved pressure from immobility whether s/he is sitting or lying down. Pressure relief for both the prevention and treatment of pressure ulcers is of critical importance to this at-risk population. When at-risk individuals are in bed, many guidelines recommend turning every 2 hours, even when a turning mattress is being used. It is important to pay attention to specific pressure-bearing areas, such as the sacrum, ischial tuberosity, trochanter, and the heels. Turning from side to side will relieve pressure over the sacrum and ischial tuberosities, but the patient should not be turned all the way to the side or the trochanteric areas will be under high pressure. The heels can be suspended by placing pillows under the lower extremities, or by the use of orthoses that take pressure off the heels. Caregivers and patients should be educated about the importance of turning. Research is needed to determine the optimal turning frequency (120). For individuals in the sitting position, current guidelines recommend pressure relief every 15 minutes.

## Table 52.2 Comparison of NPUAP Staging Systems (1989 and 2007) (Continued)

<table>
<thead>
<tr>
<th>NPUAP Staging System (1989)</th>
<th>2007 Revision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unstageable: Full thickness tissue loss in which the base of the ulcer is covered by slough (yellow, tan, gray, green, or brown) and/or eschar (tan, brown, or black) in the wound bed.</td>
<td></td>
</tr>
<tr>
<td>Further description: Until enough slough and/or eschar is removed to expose the base of the wound, the true depth, and therefore stage, cannot be determined. Stable (dry, adherent, intact without erythema or fluctuance) eschar on the heels serves as “the body’s natural (biological) cover” and should not be removed.</td>
<td></td>
</tr>
</tbody>
</table>

*Bruising indicates suspected deep tissue injury.

**Wound Care**

The purpose of local wound care is to provide the wound with the most optimal environment for healing. Specific factors that need to be taken into consideration are:

- **Level of moisture:** The provision of a moist healing environment has been accepted to be the standard of care in pressure ulcer management. It is believed that wound healing is optimized at an appropriately moist environment, while a dry or excessively moist environment will be detrimental to wound healing. Depending on the pre-existing level of moisture in the wound bed, various dressings can be used to correct the level of moisture in the wound (121). The different types of dressings are discussed under General Dressing Guidelines.

- **Debridement of necrotic tissues:** The removal of necrotic tissues, eschar, and slough is a well-accepted practice in wound bed preparation for healing. These devitalized tissues may support the proliferation of pathogens. Though debridement is widely practiced, there has not been any substantial research study on this topic. There are four methods by which debridement can be achieved (122). In the order of their onset of action, with the most rapid onset being first, they are: sharp debridement with scalpel or scissors, nonselective mechanical debridement through the use of irrigation or a wet-to-dry dressing, chemical debridement through the use of an enzyme such as collagenase without damage to the granulation tissues, and...
autolytic debridement. Whenever possible, sharp debridement is the method of choice as it provides the most effective way to remove any necrotic tissue from the wound, though it must be noted that the NPUAP advises against the debridement of eschar in the heel (1.97). Sharp debridement requires the skills of an experienced clinician and sometimes the necrotic materials may be too adherent at the wound base for successful sharp debridement. In such cases, mechanical and chemical debridement would be more appropriate, though it takes longer to achieve the desired effect.

Wound cleansing: wound cleansing facilitates the removal of necrotic materials, exudates any metabolic wastes away from the wound, thus promoting wound healing (96). It also may decrease the bacterial load in the wound tissue; this is important because a bacterial count of greater than $10^5$ may be associated with the development of wound infection (123). Wound irrigation at a pressure of 4 to 15 psi (pounds per square inch) is recommended (124). Below 4 psi, the irrigation pressure may not be effective but above 15 psi, the pressure may potentially drive the wound bacteria into the blood, increasing the risk of bacteremia. Irrigation devices that deliver a pressure above 8 psi, such as a 35 mL syringe with a 19-gauge needle or angiocatheter, are thought to be more effective. The use of a bulb syringe will produce a lower irrigation pressure, and is thus less effective. Hydrotherapy with whirlpool or pulsatile lavage is effective for wound cleansing and will be discussed under ADJUVANT THERAPIES. Antiseptic solutions are not recommended for wound cleansing because of their cytotoxic properties, which may impair wound healing. Normal saline is the irrigant of choice.

Protection of wounds: this can be achieved by the use of an appropriate dressing to protect the wound from external factors such as further trauma, and bacterial or chemical exposure.

GENERAL DRESSING GUIDELINES

Dressings provide a physical barrier that protect the wounds, help optimize the moisture level in the local wound microenvironment, assist in autolytic debridement, and can also be used as a medium to deliver topical medication to the wound. The following factors need to be considered when selecting a dressing:

- Does the dressing provide sufficient physical barrier for protection from the environment?
- Is there necrotic material in the wound? If so, is autolytic, enzymatic, or mechanical debridement most appropriate?
- What is the moisture level in the wound? Is drainage control or moisture retention the goal?
- Is there a suspicion of bacterial overload in the wound bed, for example, strong odor, excessive drainage? All pressure ulcers will have some degree of colonization but this does not necessarily mean that the wound is infected. Is an antimicrobial topical agent necessary?
- The frequency of change and ease of application of the dressing.
- The availability and cost of the dressing to the patient.

It is important to recognize that the pressure ulcer healing process is dynamic and therefore, the wound tissue condition needs to be assessed continuously with appropriate adjustment of the dressing choice.

The huge number of commercially available dressings can be helpful and yet overwhelming to the clinician. It is more helpful to understand the principles and mechanisms of action of the different types of dressings, than to know the various products by name. Table 52-3 describes the classes of the available dressings. Topical agents such as enzymatic debridement agents and antimicrobial agents may be used in conjunction with some of these dressings. Though growth factor gel, for example, bercaplermin, has been used for the treatment of diabetic foot ulcers, it has not been indicated for pressure ulcer treatment.

Adjuvant Therapeutic Modalities

There are many therapeutic modalities used in the treatment of pressure ulcers, including electrical stimulation (ES), hyperbaric oxygen, infrared, ultraviolet, low-energy laser irradiation, and ultrasound. The AHCPR Guidelines (96) supported only the use of hydrotherapy for the cleansing of the ulcers and ES for nonhealing Stage III and IV pressure ulcers. This section will examine these recommended modalities, and some newer, but commonly used therapeutic modalities.

Hydrotherapy—Whirlpool and Pulsatile Lavage Therapy

Hydrotherapy is particularly helpful for cleansing and mechanical debridement of Stage III and IV wounds. Conventionally, hydrotherapy for pressure ulcer management has been carried out through the use of a whirlpool. Despite its clinical efficacy (125), whirlpool therapy has been found to have some clinical and practical limitations, including cross-contamination between patients, pseudomonas infections (126), and potential skin infections for caregivers who are in contact with the contaminated water.

Pulsatile lavage therapy is a different form of hydrotherapy that encompasses all of the advantages of whirlpool therapy, but without the potential adverse effects. A portable device is used that delivers pulsed jet streams of water at a known, preset pressure, which is compliant with the pressure range (4–15 psi) recommended for wound cleansing (127). The potential for cross-contamination between patients is eliminated because the device is for single-patient use, and is carried out by a caregiver in the patient’s room. Furthermore, pulsatile lavage therapy is less labor intensive than whirlpool hydrotherapy because it does not involve any patient transfer and there is no decontamination of equipment afterwards. Pulsatile lavage appears to be a preferred alternative to whirlpool therapy in the management of pressure ulcers. The clinical efficacy of pulsatile lavage in the treatment of pressure ulcers is under investigation.

Electrical Stimulation

ES has been used for the treatment of chronic wounds for many years (128) and has been specifically recommended for
the treatment of severe (Grade III or IV) pressure ulcers by the AHCPR Clinical Practice Guidelines. There have been many clinical reports of the technique. Three separate meta-analyses (129–131) have attempted to consolidate the many varying clinical reports on the success of this technique (132–134). All these reports considered ES to be an effective modality; however, the specific treatment and stimulation paradigms employed were found to be highly variable. The mechanisms by which ES promotes wound healing are not fully understood, leading to a need to optimize delivery of treatment.

Negative Pressure Wound Therapy

Negative pressure wound therapy (NPWT) is based on the theory that the negative pressure facilitates drainage of wound exudates and enhances wound healing through a number of mechanisms (135). More specifically, NPWT is proposed to decrease the bacterial load and edema while concurrently promoting an improved local circulation and increasing granulation. NPWT devices consist of a suction pump with foam and occlusive dressing to create negative pressure on the wound being treated. Despite the lack of official guidelines on the use of NPWT, it has been widely used in clinical practice and there have been consensus reports (136,137). The primary recommendations are that NPWT is indicated when the following clinical criteria are met:

- Anatomical surfaces that allow a tight seal
- Adequately prepared wounds, for example, debrided, free of eschar, and necrotic materials
- Wound drainage
- Patient compliance

NPWT is contraindicated when wounds are dry, there is uncontrolled pain, untreated infection, malnutrition, or poor hemostasis.

Effective NPWT will produce a response within 2 to 4 weeks. Discontinuation is recommended if there is less than 30% wound size reduction after 4 weeks.

Despite its popularity, none of the major clinical reviews and guidelines have found sufficient scientific evidence to support the use of NPWT for wound healing (138–141). There were too few randomized controlled trials. Small sample size and poor study design limit the validity of existing studies. Further research is needed to determine the utility of NPWT as a treatment modality for appropriate nonhealing pressure ulcers.

<table>
<thead>
<tr>
<th>Class of Dressing</th>
<th>Pressure Ulcer Indication</th>
<th>Mechanisms of Action</th>
<th>Potential Adverse Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gauze</strong></td>
<td>Dry dressing</td>
<td>Scabbed over wounds</td>
<td>Physical barrier</td>
</tr>
<tr>
<td></td>
<td>Wet-to-dry</td>
<td>Stage III, IV pressure ulcers with necrotic materials</td>
<td>Mechanical debridement, absorbent</td>
</tr>
<tr>
<td></td>
<td>Wet-to-moist</td>
<td>Stage III, IV pressure ulcers</td>
<td>Provide moist healing environment</td>
</tr>
<tr>
<td><strong>Oclusive</strong>—transparent film dressings that are semi-permeable</td>
<td>Stage I, II pressure ulcers with no drainage</td>
<td>Physical barrier, moisture retention, promote autolytic debridement</td>
<td></td>
</tr>
<tr>
<td><strong>Gel</strong>—hydrophilic polymer that comes in sheet, granules, or liquid gel forms</td>
<td>Stage II, III, IV pressure ulcers with no to little drainage</td>
<td>Provide moist healing environment, promote autolytic debridement</td>
<td></td>
</tr>
<tr>
<td><strong>Hydrocolloid</strong>—dressings containing gel-forming agents (e.g., sodium carboxymethylcellulose and gelatin), often combined with elastomers and adhesives, and applied to a carrier</td>
<td>Stage II (III) pressure ulcers with little drainage</td>
<td>Occlusive and adhesive wafer dressing that forms a gel-like substance with wound exudate, promoting moist healing and autolytic debridement</td>
<td></td>
</tr>
<tr>
<td><strong>Foam</strong>—polyurethane dressing that comes in sheets or filers</td>
<td>Stage II, III, IV pressure ulcers with little to great drainage</td>
<td>Provide moist healing environment, promote autolytic debridement</td>
<td></td>
</tr>
<tr>
<td><strong>Alginites</strong>—highly absorbent, biodegradable dressings derived from seaweed</td>
<td>Stage III, IV pressure ulcers with moderate-to-great drainage</td>
<td>Absorb excessive exudates, promote autolytic debridement</td>
<td></td>
</tr>
<tr>
<td><strong>Matrix</strong>—collagen matrix that provides three-dimensional scaffolding, attracts host cells and tissue remodeling</td>
<td>Stage III, IV pressure ulcers</td>
<td>Promote granulation and epithelialization into the matrix</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Class of Dressing</th>
<th>Pressure Ulcer Indication</th>
<th>Mechanisms of Action</th>
<th>Potential Adverse Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gauze</strong></td>
<td>Dry dressing</td>
<td>Scabbed over wounds</td>
<td>Physical barrier</td>
</tr>
<tr>
<td></td>
<td>Wet-to-dry</td>
<td>Stage III, IV pressure ulcers with necrotic materials</td>
<td>Mechanical debridement, absorbent</td>
</tr>
<tr>
<td></td>
<td>Wet-to-moist</td>
<td>Stage III, IV pressure ulcers</td>
<td>Provide moist healing environment</td>
</tr>
<tr>
<td><strong>Oclusive</strong>—transparent film dressings that are semi-permeable</td>
<td>Stage I, II pressure ulcers with no drainage</td>
<td>Physical barrier, moisture retention, promote autolytic debridement</td>
<td></td>
</tr>
<tr>
<td><strong>Gel</strong>—hydrophilic polymer that comes in sheet, granules, or liquid gel forms</td>
<td>Stage II, III, IV pressure ulcers with no to little drainage</td>
<td>Provide moist healing environment, promote autolytic debridement</td>
<td></td>
</tr>
<tr>
<td><strong>Hydrocolloid</strong>—dressings containing gel-forming agents (e.g., sodium carboxymethylcellulose and gelatin), often combined with elastomers and adhesives, and applied to a carrier</td>
<td>Stage II (III) pressure ulcers with little drainage</td>
<td>Occlusive and adhesive wafer dressing that forms a gel-like substance with wound exudate, promoting moist healing and autolytic debridement</td>
<td></td>
</tr>
<tr>
<td><strong>Foam</strong>—polyurethane dressing that comes in sheets or filers</td>
<td>Stage II, III, IV pressure ulcers with little to great drainage</td>
<td>Provide moist healing environment, promote autolytic debridement</td>
<td></td>
</tr>
<tr>
<td><strong>Alginites</strong>—highly absorbent, biodegradable dressings derived from seaweed</td>
<td>Stage III, IV pressure ulcers with moderate-to-great drainage</td>
<td>Absorb excessive exudates, promote autolytic debridement</td>
<td></td>
</tr>
<tr>
<td><strong>Matrix</strong>—collagen matrix that provides three-dimensional scaffolding, attracts host cells and tissue remodeling</td>
<td>Stage III, IV pressure ulcers</td>
<td>Promote granulation and epithelialization into the matrix</td>
<td></td>
</tr>
</tbody>
</table>

TABLE 52.3 Classes of Available Dressings
Therapeutic Ultrasound
Therapeutic ultrasound is a deep heating modality that is commonly used for pressure ulcer healing. Its deep heating property is theorized to improve the vascularity of the wound tissues, thus improving healing. However, there is only limited evidence from clinical trials (142).

Electromagnetic Therapy
There is growing interest in the use of electromagnetic therapy for the treatment of pressure ulcers. This modality has been shown to increase the blood flow, collagen formation, and also granulocyte infiltration in both in vitro and animal models to induce healing (143). However, clinical trial evidence is again lacking.

PROVISION OF APPROPRIATE SEATING SYSTEMS AND MATTRESSES
Appropriate seating systems and mattresses for at-risk patients must primarily reduce the risk of pressure ulcer development. Depending on the specific patient condition and health care setting, other factors must also be considered such as postural control, optimization of independent function and, when applicable, availability of assistance and care. There are many different types of support surfaces available. General guidelines for prescription are of value, but it remains important to assess each patient individually. The provision of support surfaces should be reviewed regularly for at-risk patients, both in the short-term acute care setting and in longer term care. Changes in overall health and other risk factors frequently change a patient’s requirements for effective pressure relief.

Seating Systems
Along with the provision of an appropriate wheelchair, an effective cushion is essential in order to provide a seating system that will maximize the function and minimize the risk for pressure ulcer development. There are a wide and continuously increasing range of pressure relief cushions available. The main classes of pressure relief cushions available are summarized in Table 52-4. Some cushions will combine two or more classes of material in order to achieve improved pressure relief properties, in combination with postural stability. USA TechGuide is an online resource maintained by the United Spinal Association that provides complete and constantly updated reviews of currently available pressure relief cushions (144). Donut-rings should never be used as pressure relief cushions because, although pressure is relieved at the center of the ring, pressure distribution around the perimeter leads to high regional pressures and blood flow occlusion, thus exacerbating the overall risk.

Pressure Relief Mattresses
Pressure relief mattresses use many of the same materials as pressure relief cushions. Static and dynamic support mattresses are routinely available. Static support mattresses are appropriate for patients without a pressure ulcer or who can be positioned so that no load is applied to the pressure ulcer at any time. Dynamic mattresses include low air loss mattresses and air-fluidized beds, and are frequently only used for patients with existing pressure ulcers, due primarily to high usage costs. However, all immobile patients are at an increased risk of pressure ulcer development and the recommended 2-hourly turning schedule for pressure relief sometimes poses a staffing challenge. If loading over an existing pressure ulcer cannot be avoided at all times then a dynamic support mattress should be used. The medico-legal repercussions of a patient developing a nosocomial pressure ulcer have led to an increased consideration of use of dynamic support mattresses in many health care facilities (145–147), and it may also be cost-effective (148).

SPECIALIZED CLINICS
Specialized clinics for patients at-risk for pressure ulcers can include both wheelchair clinics and wound care clinics. The personnel involved in these clinics will vary slightly but a common focus is the prevention and treatment of pressure ulcers.

The wheelchair clinic team generally includes a physiatrist, a physical and/or occupational therapist, and a medical engineer. The physiatrist plays a critically important role in ensuring that the overall risk factor status is considered and that comorbidities, such as nutritional status, are addressed. It is valuable for at least one of the team members to have completed an appropriate Rehabilitation Engineering and Assistive Technology Society of North America (RESNA) accreditation program (149). Certification as an Assistive Technology Practitioner (ATP) is for professionals who assess patients and assist in system selection and user training. In the wheelchair clinic, the therapist or ATP activities will include

<table>
<thead>
<tr>
<th>TABLE 52.4 Cushion Material Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Material</td>
</tr>
<tr>
<td>----------</td>
</tr>
<tr>
<td>Foam</td>
</tr>
<tr>
<td>Viscoelastic foam</td>
</tr>
<tr>
<td>Gel</td>
</tr>
<tr>
<td>Flotation (water or air)</td>
</tr>
</tbody>
</table>
SUGICAL MANAGEMENT OF PRESSURE ULCERS

Surgical procedures may be used to treat pressure ulcers. They include: direct closure, skin grafting, skin flaps, musculoscutaneous flaps, and free flaps (96). Though there are no clear criteria or clinical practice guideline for the indication of surgical intervention, patients with nonhealing Stage III or IV pressure ulcers despite optimal conservative care, and those with underlying osteomyelitis requiring debridement of the infected bone, may be candidates for surgery. Surgical removal of a bony prominence such as ischiectomy is contraindicated due to poor pressure distribution and postural control postoperatively (154). Risk factors that impair wound healing should be controlled before surgery is considered, including malnutrition, anemia, spasticity, urinary tract infection, incontinence, smoking, and psychosocial issues. It is imperative to provide postoperative pressure relief to the surgical site with appropriate positioning of the patients and pressure-relieving mattresses. As the long-term outcomes of pressure ulcer surgery remain unclear (155,156), careful selection of surgical candidates is of utmost importance.

NEW CONCEPTS IN PRESSURE ULCER CARE

Minimizing Bedrest for Pressure Ulcer Treatment

Prolonged bedrest can lead to many complications, thus effective clinical management without bedrest for patients with pressure ulcers is desirable. Prone trolleys allow patients to gain some mobility but provide only limited function and mobility in the community environment. Standing wheelchairs have recently become an increasingly popular option although they remain expensive. These systems can provide full, extended pressure relief over pelvic region wounds, thus allowing remobilization of individuals with existing pressure ulcers. Standing wheelchairs also have multiple systemic health benefits including increases in lower extremity range of motion, upper extremity strength, bladder, and other systemic functions together with decreased spasm and pain (157). USA Tech-Guide currently lists over 15 power-standing wheelchairs and 11 manual-standing wheelchairs (158,159).

Dynamic wheelchairs and cushions provide an alternative approach to a full standing wheelchair. A dynamic configuration wheelchair varies individual system components in order to relieve pressure over areas of tissue breakdown (160). Dynamic wheelchair cushions provide a means to provide automatic weight-shifting due to cyclic inflation and deflation of air cells in high-load regions and may provide benefits similar to those achieved using an advanced wheelchair system.

Further research is needed to determine the relative merits of advanced wheelchairs and wheelchair cushions for both pressure ulcer prevention and treatment.

Gluteal Neuromuscular Electrical Stimulation System

Regular use of ES for functional applications can produce changes in regularly stimulated muscles that may increase the health of the muscle and surrounding soft tissues (161,162). Bogie et al. have investigated the use of a 4-channel gluteal electrical stimulation system (GSTIM), specifically designed to decrease risk factors associated with pressure ulcer development for individuals with SCI (163). It was found that subjects who received a GSTIM system showed statistically significant decreases in ischial region pressure over time, together with increased gluteal cross-sectional area and enhanced regional blood flow (164). Dynamic weight shifting produced by the GSTIM system varies seated posture and pressure distributions at the seating interface and augments the efficacy of conventional pressure relief maneuvers.

SUMMARY

Pressure ulcers remain a major complication for many rehabilitation patients with impaired mobility. Our understanding of the pathophysiology of tissue breakdown and the impact of concurrent systemic risk factors has increased. Advances in therapeutic modalities and standardization of clinical care contribute to the success of rehabilitation practice.

REFERENCES


CHAPTER 52 | PRESSURE ULCERS


PART V

Special Populations
Physiatric involvement in the field of sports medicine has greatly increased over the past two decades. More and more physical medicine and rehabilitation specialists are serving as team physicians at the high school, collegiate, Olympic, and professional levels. In addition, there has been increased involvement in the American College of Sports Medicine as participants and presenters. The Physiatric Association of Sports, Spine and Occupational Rehabilitation continues to promote and actively participate in educational programs dealing with sport medicine. It is encouraging that more physiatrists are publishing articles related to sports medicine, thus increasing Physical Medicine and Rehabilitation’s contribution to the literature. Validation of the effectiveness of non-operative treatment of various sports injuries will continue to substantiate the role of physiatrists as leaders in the area of sports medicine.

The primary purpose of this chapter is to review selected concepts pertaining to the treatment of sports injuries. The evaluation and treatment of some common injuries are also presented. Overuse injuries are discussed in a separate chapter in this textbook. The reader should understand that the area of sports medicine is broad; therefore, for those interested in treating sports injuries, additional reading of current textbooks and journals is essential. In addition, clinical experience gained through preparticipation physical examinations, sports clinics, and on-field coverage is invaluable to anyone who desires to care for injured athletes.

Since the previous edition of this book was published in 2005, a team physician consensus statement was developed. In a collaborative effort by six professional organizations—American Academy of Family Physicians, American Academy of Orthopedic Surgeons, American Academy of Sports Medicine, American Medical Society for Sports medicine, American Orthopedic Society for Sports Medicine, and American Osteopathic Academy of Sports Medicine—a consensus statement was devised that outlined the most important elements of sports medicine and their specific issues in injury and illness prevention (1). This chapter addresses every topic discussed in that consensus statement, and thus it is completely up to date in its information at the time of this meeting. Much of the previous edition was left intact, and new contributions have been included on the subjects of cardiovascular issues in sports and on anterior cruciate ligament (ACL) injury prevention.

TEAM APPROACH TO THE INJURED ATHLETE

Various health care professionals are involved in the care of injured athletes. Medical specialties traditionally involved in the care of athletes include physiatry, orthopedic surgery, internal medicine, family practice, emergency medicine, and pediatrics. Additional important team members include podiatrists, athletic trainers, therapists, nutritionists, and strength and conditioning specialists. The concept of working within a team approach is inherent in physiatry. During residency training and in practice, physiatrists continually coordinate care by interacting with physical therapists, occupational and speech therapists, psychologists, and other pertinent individuals. Although the vast majority of athlete symptoms are musculoskeletal in nature, a variety of problems can arise, ranging from skin disorders to cardiac symptoms. In addition, most acute injuries do not require operative intervention. A well-trained physiatrist would appear to be the most logical choice to direct the sports medicine team, addressing the majority of the nonoperative injuries in conjunction with athletic trainers, therapists, and others, and referring other conditions as needed to the most appropriate specialist in a timely manner.

BASIC TISSUE PHYSIOLOGY

Skeletal Muscle

Muscle constitutes the largest tissue mass in the body, comprising 40% to 45% of the total body weight. Muscles originate either from bone or dense connective tissue and are connected to these tissues by either the muscle itself or associated tendons. The muscle-tendon unit will cross one or more joints to its site of insertion, typically on bone. Those muscles that cross one joint are usually located close to bone and are generally involved in postural activities. One-joint muscles possess a slower contraction speed with increased force production. Examples of one-joint muscles include the vastus medialis/lateralis/intermedius of the quadriceps group and the soleus. In contrast to the one-joint muscles, those muscles that cross two or more joints are located more superficially and exhibit faster contraction speeds with relatively reduced force production. Examples of two-joint muscles include the biceps brachii and the gastrocnemius. Muscles can generate and attenuate force on a limb and stabilize a joint when firing synergistically with larger muscles.
Muscle that is torn has a limited ability to regenerate and often heals by dense scar formation. It has been reported that aging muscle demonstrates a decrease in size, number of muscle fibers (type II), and the number of motor units (alpha motor neuron and the muscle fibers it innervates). There is, however, no decrease in metabolic potential for aerobic and anaerobic activities with aging (2). Petrofsky and Lind (3) demonstrated little change in mean handgrip in men aged 22 to 62 years. Vandervoort and McComas found little change in dorsi- flexion and plantar flexion torques between men aged 20 to 60 years, with a marked reduction in strength from age 60 to 100 years (4). Frontera (5) strength-trained men between 60 and 70 years and showed 107% and 226% strength increase for knee flexors and extensors. Fiatarone et al. (6) strength-trained nonagenarians, which resulted in a 174% increase in quadriceps strength and a 48% increase in gait speed. These studies demonstrate that muscular strength can be retained until very late in life, thereby supporting the concept that muscle loss with aging is primarily due to disuse more than age-related muscle fiber loss.

Tendons
As previously mentioned, a tendon is the structure that connects muscle to bone. Tendons consist of dense, regularly arranged collagen fibers meshed with elastin and a proteoglycan/ glycosaminoglycan ground substance. The primary function of the tendon is to transmit the force generated in muscle to the bone. This process allows for the generation of movement of the extremities. The tendon is surrounded by a glistening, synovial-like membrane called the epitenon. The epitenon is continuous on its inner surface with the endotenon, a thin layer of connective tissue that contains lymphatics, blood vessels, and nerves. In some tendons, the epitenon is surrounded by a loose connective tissue called the paratenon. The paratenon functions as an elastic sheath that permits free movement of the tendon against surrounding tissue. The blood supply of tendons is variable and is compromised at sites of friction, torsion, or compression. Tendons with a blood supply that is frequently compromised include the supraspinatus, Achilles, and tibialis posterior. Tendons subjected to large, repetitive stresses can tear, become inflamed, or degenerate.

The aging process is associated with an increase in collagen cross-linkages, causing the fascia, ligaments, and tendons to be less distensible. Activity promotes connective tissue hypertrophy, whereas inactivity leads to atrophy. Physical activity also enhances the rate of collagen turnover, which shortens its life span. This turnover retards the effect of maturation of collagen cross-linkages and makes collagen more distensible (7).

Cartilage
Cartilage requires mechanical loading and unloading to remain healthy. This helps maintain nutrition to the tissues. Cartilage rarely has a direct blood supply and derives most of its nutrition via diffusion. If deprived of bearing weight, the cartilage undergoes degeneration (8). With repetitive overuse, cartilage can break down over time and lead to symptoms such as pain.

Ligaments
Ligaments are composed of approximately 70% collagen, along with ground substance and a larger percentage of elastin than tendons. They typically attach bone to bone and are important stabilizers of joints. Ligaments contain minimal blood supply, with most supplied by periarticular arterial plexuses. For example, the ACL, the main stabilizer of the knee against anterior translation of the tibia, has minimal intrasubstance arterial supply. The ACL does, however, receive some nutrition from the synovial fluid as well. The poor blood supply to ligaments explains the poor healing response in complete/partial injuries. The principles of collagen cross-linkages described in the tendon also apply to ligaments. Ligaments function to stabilize joints in concert with normal osseous alignment and musculo-tendinous stability. Ligaments also enhance proprioception via neurosensory receptors.

Bone
Bone is a dynamic tissue that remodels in response to external stress. Bone remodeling/healing represents the spectrum of bone’s response to stress (Wolff’s Law). Bone is a hard tissue consisting of cells in a matrix of ground substance and collagen fibers. The fibers are impregnated with mineral substance, chiefly calcium phosphate and carbonate. In addition to failure of bone secondary to repetitive stress (i.e., stress fractures), bony malalignment will also predispose the athlete to injury. Examples of malalignment include valgus alignment of the knee predisposing to patellar maltracking or a cavus foot predisposing to Achilles tendinitis.

Basic Principles of Nonoperative Functional Rehabilitation
Several basic principles can be applied to almost any acute sports injury, as outlined in Table 53-1. These rehabilitation phases provide a stepwise approach to treat and assess the progress of an athlete with an acute injury. Immobilization is avoided as much as possible because of its multiple detrimental effects on tissue healing (e.g., scar formation, contracture, and atrophy).

Phase I: Decrease Pain and Control Inflammation
The initial phase of treatment is to control the inflammatory reaction that occurs after an acute injury and its associated

<table>
<thead>
<tr>
<th>TABLE 53.1 Phases of Sports Rehabilitation</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. Resolving pain and inflammation</td>
</tr>
<tr>
<td>II. Restoring range of motion</td>
</tr>
<tr>
<td>III. Strengthening</td>
</tr>
<tr>
<td>IV. Proprioceptive training</td>
</tr>
<tr>
<td>V. Sports/task specific activities</td>
</tr>
</tbody>
</table>

Note: One should always remember to identify and address deficiencies along the entire kinetic chain.
pain, which inhibits muscle function. It should be noted that mediators involved in the inflammatory response are also important factors involved in the healing of soft-tissue injuries. Therefore, although the goal is to control the inflammatory response, eliminating it completely could be detrimental to tissue healing. The PRICE (protection, rest, ice, compression, elevation) approach is well known to those who care for athletes.

After an injury, the area can be protected either by splinting, bracing, or taping/wrapping. It is very rare that any ligamentous or musculoskeletal injuries would require casting, which is usually avoided. Crutch ambulation (usually weight bearing as tolerated) for lower extremity injuries can be very helpful until a normal, pain-free gait pattern can be reestablished. Bracing should be limited to that which protects the specific area while allowing for full motion at other areas. It is often possible to use the same brace to facilitate early protective motion as well as return to functional activities later in the rehabilitation process (e.g., a double upright hinged knee brace following a medial collateral knee sprain).

Rest should be prescribed carefully, as it is important that the athlete does not become deconditioned during the rehabilitation of an injury. Fatigue results in a decrease in neuromuscular functioning and joint control, thereby placing greater dependency on the static stabilizers of joints (i.e., the ligaments), placing them at greater jeopardy for injury. Therefore, the proper term is relative rest, which means that while the affected area is rested, the remainder of the body is exercised. In particular, cardiovascular conditioning must be maintained. This can be done by alternative exercises that allow for protection of the injured area while stressing the cardiovascular system at the same intensity, duration, and frequency as the athlete had previously trained. For example, a running athlete who suffers a lower extremity injury that must be unloaded can use deep-water vest running at the same intensity level as before the injury. This has allowed athletes to maintain their cardiovascular fitness level while their injuries heal, allowing for safe return to play at close to the preinjury level and possibly prevent additional injuries.

Ice controls the initial inflammatory response and facilitates pain control immediately following an injury. The affected areas should be generally iced 20 minutes four to five times a day, or more often if possible. Ice is used for its properties of vasoconstriction, which limits the edema as well the release of vasoactive and pain factors, such as bradykinins and leukotrienes. Ice can also decrease conduction along pain fibers and act as a counterirritant to assist in pain control and to reduce muscle spasm. There are various methods of icing that include ice pack, ice massage, ice immersion, and devices that combine both ice and compression.

Compression is also used in an effort to limit the edema in the injured area. Ace wrapping is often used but can be problematic because of the difficulties in getting uniform or gradient (from distal to proximal) compression. A compressive stockinette (e.g., Tubigrip) can be very helpful in this regard. A sleeve can be cut to whatever size is necessary and simply applied. For additional compression, it can be folded over onto itself. Care must be taken to avoid excessive pressure over bony protuberances or superficial nerves. Compressive braces can also be effective (e.g., air splints for ankle sprains). Finally, devices that combine icing with compression have been found to be very useful and effective in the postinjury as well as postoperative rehabilitation of athletic injuries. They can be used not only by therapist and athletic trainers but also at home by motivated athletes.

Elevation is yet another means to control postinjury swelling. The injured limb should be elevated above the level of the heart to optimally assist with venous and lymphatic drainage and therefore control edema. Keeping the lower extremities out of a dependent position is helpful as well in limiting the pooling of inflammatory and posttraumatic products.

Additionally, nonsteroidal antiinflammatory drugs (NSAIDs) for a short period of time, if not contraindicated, and electrical stimulation can assist with both inflammation and pain control. Whether they offer clear advantages over using just the above program is a matter of debate, but if available and not contraindicated, appear reasonable.

**Phase II: Restore Normal/Symmetric Range of Motion**

Pain and swelling can inhibit motion or produce altered motor patterns that, if established, often require retraining to restore proper motor control. An example is an athlete with an antalgic gait following a knee or ankle injury. This movement pattern must be discouraged while the area is gradually mobilized. Immobility will result in scar and contracture and therefore is not recommended. Range of motion (ROM) allows for controlled stress to a joint, which will stimulate proper collagen deposition. Motion provides sensory input to the central nervous system, which stimulates the proprioceptive system as well as modulates pain via the Gate theory. In the early phase, pain-free movement of a joint and stretching that prevents contractures are encouraged as the motion that results in stress on the injured area is avoided. As the pain and inflammation subside, more aggressive stretching and mobilization continue until symmetric (to the unaffected limb) motion is achieved with normal movement patterns.

**Phase III: Restore Normal/Symmetric Strength**

Strengthening a painful, inflamed limb that lacks normal ROM can result in further problems that can delay recovery from injury. Therefore, a stepwise approach toward strengthening must be used. In the early postinjury phases, pain-free isometric contractions performed several times throughout the day are encouraged in an effort to retard muscular atrophy. A simple method is to recommend 10-second contractions, with 10 repetitions, 10 times a day. These isometric contractions may need to be performed through multiple angles as the strengthening is specific to the manner and position in which a muscle is trained. As the injured area recovers and ROM is restored, isotonic strengthening can begin if possible. Currently, there is no significant role for isokinetic strengthening because of poor functional
carrying over. To that end, closed kinetic chain exercise should begin as soon as possible and progressed as able. Resistance training can be in the form of exercising against gravity, free weights, machines, and resistance tubing. The strengthening should be as functional as possible, attempting to match the demands of the sport. Resistance tubing strengthening is attractive because of its ease and simplicity; however, the greatest tension with resistance occurs at the end ROM, where the muscle is usually weakest and the joint is most vulnerable. Therefore, this should be reserved for later stages of the strengthening program. In addition, the use of plyometric exercise should be included as the athlete is preparing to return to sport, because such training may ready the athlete for explosive bursts that are often a necessary part of many high demand sports.

**Phase IV: Neuromuscular Control (Proprioceptive)**

Retraining

In order to dynamically control a joint during sport activity, there needs to be not only full ROM and normal strength but also adequate dynamic motor control. Specifically, the injured joint needs to be stabilized by synchronous activation of appropriate muscle groups so that the larger, more powerful muscles may safely produce the necessary force required in sports activity. Many injuries can result in proprioceptive loss that may predispose an athlete to repeat injury. As in other areas addressed in the rehabilitation process, the proprioceptive system needs to be progressively challenged in order for progress to be made. Simple proprioceptive training can include seated exercises with a wobble board for lower extremity injuries or loading exercises of the arm either on a table or wall. As the athlete recovers, and assuming that there is near full ROM and strength, the proprioceptive system is progressively challenged (e.g., balancing on a single leg while catching and throwing, balancing with eyes closed). Proprioceptive training requires a great deal of one-on-one work with a therapist or trainer and often creativity in developing ways to challenge the proprioceptive system that corresponds to the athlete’s sport (9).

**Phase V: Return to Sport Activities**

As the athlete completes these phases, the therapist or trainer must then begin the transition to return to sport. This occurs as the athlete successfully meets the challenges of the previous phases. The athlete then is put through activities that replicate the demands of the sport. For example, a basketball player will be given various drills that include running, cutting, and jumping (and landing) using optimal biomechanics. Once the athlete demonstrates that he or she can successfully negotiate the various drills and challenges that will occur in the sport in a controlled situation, then it should be clear to everyone on the team (including the athlete) that he or she can safely return to their sport.

**CARDIOVASCULAR ISSUES IN SPORTS**

An integral part of the preparticipation physical examination for athletes is the cardiovascular evaluation. The goal of this evaluation is to identify athletes who are at risk for sudden cardiac death during vigorous physical activity. By applying elements of the personal history, the family history, and the physical exam, the most important signs and symptoms of the most common cardiac reasons for sudden death can be obtained. These include hypertrophic cardiomyopathy (HCM), selected arrhythmias, coronary artery anomalies, ruptured aortic aneurysms, and commotio cordis (1). In particular, HCM has received much attention in the press and in the literature as it has taken the lives of several high profile athletes. It is the primary cause of sudden atrumatic death in athletes, responsible for nearly 35% of those deaths (10). We will briefly discuss HCM below because it is important to understand the pathophysiology in order to be able to recognize the signs and symptoms of this silent killer.

**Hypertrophic Cardiomyopathy**

HCM is a hereditary condition that occurs as commonly as 1:500 in the adult population (11). It is most commonly transmitted via an autosomal dominant pattern in which there is a genetic defect in the sarcomere contractile proteins (12). This leads to an increased ventricular muscle mass that is characteristically the hallmark of this condition. It is important to note that there is *not* an associated increase in actual ventricular cavity size; just a hypertrophy of the muscle tissue itself. This is important because it can often be difficult to differentiate HCM from a conditioned athlete’s heart. A well-conditioned athlete can sometimes develop a structural remodeling and an increased ventricular cavity to accommodate the larger ejection fraction that develops due to their increased efficiency for oxygen extraction at the tissue levels, while at the same time maintaining normal systolic and diastolic function (11). In contrast, in HCM, there is a severe net reduction in actual inner ventricular cavity size because of the much enlarged and hypertrrophied muscle mass.

In HCM, there is also evidence of reduced compliance secondary to the inability to adequately relax the hypertrophied muscle mass. This causes diastolic dysfunction, which in turn leads to an increased left ventricular end-diastolic pressure (LVEDP) and subsequent impedance to appropriate diastolic filling. This increase in LVEDP may then lead to the development of a pressure gradient between the left ventricle and the aorta. This gradient is clinically significant to the sports medicine physician because it can be affected with certain maneuvers aimed at increasing or decreasing left ventricular end-diastolic filling. Techniques such as Valsalva maneuver or changing positions from standing to squatting and back to standing are aimed at increasing or decreasing left ventricular end-diastolic filling. If a murmur is present, these techniques will affect the quality of the murmur differently depending on its etiology.

A systolic murmur can occasionally be heard in HCM. It is the result of the pressure gradient between the left ventricle and the aorta that was previously mentioned. Any maneuver that decreases left ventricular end-diastolic filling, such as Valsalva or rising from squatting to standing, will subsequently
lead to dynamic obstruction in the left ventricular outflow tract, thereby increasing the intensity of the HCM murmur. Conversely, the murmur is also ominous if it decreases in intensity when maneuvers to increase left ventricular end-diastolic filling are performed. Such maneuvers will alleviate dynamic obstruction in the left ventricle outflow tract, thus decreasing the murmur. Therefore, when a patient goes from standing to squatting and the murmur that the physician is hearing is decreased in intensity, this is a sign of abnormal hemodynamics and can cause the physician to be suspicious for HCM. Benign murmurs, on the other hand, will increase with squatting because blood return to the heart is increased; and they will decrease in intensity upon rising from squatting, as blood return to the heart diminishes (10).

The most common signs/symptoms of HCM are syncope or sudden cardiac death with exertion. Therefore, risk factors for HCM, as well as for all the other cardiovascular anomalies such as arrhythmias or ruptured aneurysms or coronary artery anomalies, must be understood and thoroughly investigated in the preparticipation physical examination.

**General Cardiac Evaluations**

The guideline recommendations for preparticipation athletic screening of the cardiovascular system include a thorough evaluation of the family history, the personal and present medical history, and the physical evaluation (13,14).

Importantly, one should ask if there is a history of premature or sudden cardiac death, or if there have been any deaths of unknown etiologies in young family members. As mentioned earlier, HCM is a familial disease, as is Marfan’s Syndrome, which can lead to ruptured aortic aneurysms. Thus, any history of prior sudden cardiac death, particularly in young family members, is imperative. It is also important to screen for history of heart disease in anyone in the family under the age of 50.

There are also elements of the personal history that are crucial to the cardiovascular evaluation of any athlete. Questions regarding medical history of heart murmurs can signify more than just a benign flow murmur. Likewise, systemic hypertension, particularly if it is not well controlled, can cause abnormal hemodynamics and add undue stress to the heart. The physician must also importantly inquire about exertional symptoms. Has the patient ever had exertional syncope or near syncope? Is there a history of exertional chest pain? Is there a history of dyspnea upon exertion that is out of proportion to the amount of exercise being performed? Have there been palpitations or irregular heartbeats? All of these could be symptoms of a heart that is struggling to meet the demands of the exercising musculoskeletal system (14,15).

The physical exam, as we have alluded to, is key to identifying possible at-risk athletes. Auscultation of a murmur warrants further evaluation and requires employment of the previously described tactics (Valsalva, squatting, standing) to be able to demonstrate the qualities of the murmur and to possibly define some pathology. Of course, any heart rhythm abnormality needs to be evaluated further. Pulses must be palpated for intensity and heart rhythm. Lateral and inferior migration of the point of maximal impact (PMI) of the heart on the chest wall may also be a sign of left ventricular hypertrophy. If any of the above-mentioned scenarios are uncovered, the athlete must be precluded from exercise until further workup has occurred with EKG, echocardiogram, and possibly a Holter monitor.

Interestingly, Corrado et al. performed a multiyear study in Italy, analyzing the efficacy of a nationally implemented standardized preparticipation cardiovascular evaluation for the detection of at-risk athletes for sudden death. Their study also included an EKG on top of all the previously described history and physical exam (16). This multiyear study concluded that there was decreased incidence of sudden cardiac deaths since the initiation of the national screening program secondary to better detection of the at-risk athletes and precluding them from participation. It prompted some here in the United States to ask questions about whether we need to have a national standardized preparticipation screening program, and if we do, should we add the EKG to the preparticipation evaluation as a routine screening test to complement the history and physical exam. However, there have been questions raised regarding true cost effectiveness and utility of the EKG as a part of the routine screening cardiac evaluation (17).

Needless to say, an aggressive and focused preparticipation evaluation must be an essential part of the armamentarium of the sports medicine physician. He or she must understand the pathophysiology of the most common causes of sudden cardiac deaths in athletes, as well as be able to identify which athletes are at risk for this based on their preparticipation evaluation. The physician must also serve as an educator to coaches, players, and parents about the warning signs of dangerous cardiac scenarios (1).

**CONCUSSION IN SPORTS**

A concussion results from trauma transmitted either directly or indirectly to the head, causing impairment of the brain’s normal function (Table 53-2). This impairment may last from seconds to days. At times, dysfunction or postconcussive symptoms (headache, dizziness, tinnitus, irritability, memory impairment, nausea/vomiting, fatigue, etc.) can last months to years. Concussion is a type of brain injury, which can be classified as minor, mild, moderate, or severe (18).

Documented signs and symptoms of a brain injury (concussion) include amnesia (retrograde/antegrade), loss of consciousness (LOC), headache, dizziness, nausea, attentional deficit, and blurred vision. Additional features have been mentioned and include confusion/disorientation, being “dazed” or having one’s “bell rung,” inability to relate games specifics (period, opponent, score, plays), dizziness, seeing flashing lights, tinnitus, diplopia, impaired concentration, slurred speech, inappropriate behavior (laughing/crying), irritability, altered taste or smell, impact seizure, imbalance, and decreased ability to play the sport (19).
Concussion is defined as a complex pathophysiological process affecting the brain, induced by traumatic biomechanical forces. Several common features that incorporate clinical, pathological, and biomechanical injury constructs may be used in defining the nature of a concussive head injury include:

1. Concussion may be caused by a direct blow to the head, face, neck, or elsewhere on the body, with an "impulsive" force transmitted to the head.

2. Concussion typically results in the rapid onset of short-lived impairment of neurological function that resolves spontaneously.

3. Concussion may result in neuropsychological changes but the acute clinical symptoms largely reflect a functional disturbance rather than structural injury.

4. Concussion results in a graded set of clinical syndromes that may or may not involve LOC. Resolution of the clinical and cognitive symptoms typically follows a sequential course.

5. Concussion is typically associated with grossly normal structural neuroimaging studies.

The evaluation of the athlete suspected of a concussion should include a history of past head and neck injuries (including orofacial), severity of impact (magnitude of force, linear, rotational), prior structural deficits if the athlete had a prior injury, and genetic phenotype if known (apolipoprotein). This should be combined with the detailed evaluation of the current episode. Although LOC is an appropriate concern, it appears insensitive for most concussions, and amnesia may be a better measure of severity.

If a history of multiple concussions is obtained, care should be taken to quantify severity of the precipitating events and duration of concussive symptoms. It appears that a history of concussion results predisposes to subsequent brain injuries.

Postconcussive symptoms include headaches, nausea, vomiting, drowsiness, numbness or tingling, balance impairment, vertigo, sleep impairment, light or noise sensitivities, difficulty concentrating or remembering, sadness, anxiety, dizziness, irritability, or fatigue. Although postconcussive symptoms are controversial, they are felt to arise from either the direct brain impairment or a psychological reaction to the brain injury itself.

Concern for concussion should go beyond the immediate injury. If an athlete were to return to competition without being fully recovered, he or she could sustain a second injury of greater severity. Schneider first described such a syndrome in 1973 (20). It was later coined "second-impact syndrome" by Saunders and Harbaugh in 1984 (21). Second-impact syndrome results from a person sustaining a second brain injury before symptoms of a prior concussion have cleared. This second trauma can be relatively minor in severity and may not involve the head. It is believed that the brain's vascular autoregulation becomes impaired from the first injury, leading to engorgement within the cranium. Increased intracranial pressure ensues, resulting in herniation of the medial temporal lobes through the tentorium or cerebellar tonsils through the foramen magnum. Clinically, the athlete suffers a second head injury and becomes dazed. Within 15 seconds to a couple of minutes, the athlete rapidly decompensates. He or she collapses, becomes semicomatose, and suffers from pupil dilation, loss of eye movements, and respiratory failure.

The precise incidence of second-impact syndrome is unknown, but from 1980 to 1993, only 35 presumed cases were reported to the National Center for Catastrophic Sports Injury Research in Chapel Hill, North Carolina. Seventeen of these cases were confirmed. The mortality rate is 50%, and the morbidity rate is 100% (22).

Impact seizures are an uncommon result of a mild head injury. They develop within seconds of the insult and are not associated with any structural brain injury or long-term risks. The seizures do not need treatment, and the athlete should not necessarily be eliminated from sport.

Concussions are more common in some sports, including American and Australian football, ice hockey, rugby, and soccer. The National Center for Catastrophic Sports Injury Research's database compiles statistics on baseball, ice hockey, tennis, basketball, lacrosse, track, cross-country skiing, volleyball, field hockey, soccer, water polo, football, softball, wrestling, gymnastics, and swimming. Estimated annual incidence of concussions is approximately 300,000 in the United States. Concussion rates per 1,000 player hours range from 0.25 to 23.0 (23).

Many different scales have been developed to evaluate concussed athletes for research purposes (Cantu, Colorado, American Academy of Neurology, Virginia Neurological Institute, Torg, etc.). Each includes parameters of consciousness and amnesia with varying importance to determine the grade or severity of injury. Currently, posttraumatic amnesia or working memory is felt to be of greater importance than LOC. People have tried abbreviated sidelines psychological batteries (24), balance (25) and cerebellar physical examination maneuvers, and assistance from coaches and athletic trainers for return-to-play evaluation. Recent research with more sensitive instruments demonstrates impairment in conceptual thinking (26), sustained attention and visuoperceptual processing (27), and reaction time (28,29). Currently, clinicians have no scientifically validated information on clinical management of concussions or treatment effects on long-term outcomes, though studies are currently on-going.

Minor traumatic brain injury (TBI) has not been discussed in the literature as such. An athlete sustains a momentary loss or alteration in consciousness and has memory impairments that may be unrecognized. Often athletes are evaluated (usually football players), who show signs and symptoms of...
a mild TBI when there has been no clear identifiable event. It is unclear whether their symptoms are the result of many “mini” concussions, one blow to the head they cannot recall, or repetitive acceleration-deceleration trauma to the head. These seemingly less severe concussions, in which no LOc occurs, may have greater long-term neurologic sequelae than a grade 2 concussion, in which a significant period of LOC exists. Further contributing to the difficulty in grading concussions is the underreporting of grade 1 concussions by the athlete. This occurs because of his or her fear of being withdrawn from the contest or because the athlete feels it is “part of the game.” Athletes commonly refer these milder concussions as getting the “bell rung.” One must question whether they are truly milder head injuries. This should be taken into consideration when evaluating an athlete. Studies are currently ongoing on this subject matter.

Mild head injury generally has no LOC, posttraumatic amnesia of less than 1 hour, and a Glasgow Coma Scale of 15 (30). The greatest number of concussions falls into this category of severity (including the “minor” TBI) (31). Because the Glasgow Coma Scale is not sensitive enough to be useful in the evaluation of a mild TBI, and LOC also appears to be too insensitive, more weight has been placed on amnesia.

Moderate TBIs generally have no LOC of less than 5 minutes, and posttraumatic amnesia from 1 to 24 hours, whereas severe TBIs have an LOC greater than 5 minutes and posttraumatic amnesia greater than 24 hours.

Paper-and-pencil testing (McGill ACE, SAC) and computerized evaluations (IMPACT, CogSport, etc.) are currently being researched (National Football League, National Hockey League, Australian football, Pennsylvania State University, University of Pittsburgh, etc.) in an attempt to develop a quick, simple, cost-effective, and accurate adjunct to the clinical evaluation for evidence-based decision making and return-to-play guidelines. Many of these are being implemented in the preseason as baseline tests in many high schools and universities around the country, with the goal being to retest an athlete after a concussion occurs. The athlete’s progress can be tracked based on his or her scores in these evaluations. This serves, as stated above, as an adjunct to the clinical evaluation for decision making regarding a return-to-play timeframe.

There had been no consensus regarding concussions or return to play until a conference was held in 2001 in Vienna (18) (Table 53-3). The consensus is that when an athlete is initially diagnosed with a concussion, he or she should not return to the current practice or game. He or she should have continued, intermittent monitoring for deterioration. There should be a medical evaluation if concern for a concussion was made by someone other than a physician. Return to play should be in a stepwise fashion, and if doubt is present, the athlete should be held from competition.

Return to play should include no postconcussive symptoms at rest or exertion, consideration or the concussion scales, and prior history of the athlete and his or her injuries. Some institutions have preparticipation neuropsychological batteries on athletes, which can be repeated and compared with themselves, as well as age-matched and position-matched controls. This can be a great assistance in the decision-making process.

### CERVICAL SPINE INJURIES IN SPORTS

Cervical spine–related injuries are relatively common in contact/collision sports, such as football and wrestling. These injuries can include trauma to the tissues of the neck, including the muscles, ligaments, discs, and bones, as well as the neurologic structures, such as the nerve roots and spinal cord.

### Cervical Disc

As with the general population, cervical disc disease is seen in the more mature athlete. Some sports such as football and wrestling appear to increase the risk of developing a cervical disc syndrome. Cervical disc disease most commonly affects C5 to C6, followed by C6 to C7 and then C4 to C5. This pattern roughly corresponds to the amount of normal motion that occurs at each level, with areas of greater motion (i.e., C5

---

**TABLE 53.3 Concussion Guidelines from the 2001 Vienna Concussion in Sports Guidelines (CISG)**

- Revised definition
- Strength and weakness of existing concussion guidelines (no singly system endorsed)
- Absence of scientifically validated return-to-play guidelines
- Defined signs and symptoms of acute concussion
- Cognitive (confusion, amnesia, LOC, disorientation)
- Symptoms (headache, dizziness, nausea, loss of balance, “dinged” or stunned, seeing stars of flashing lights, tinnitus or double vision, sleepiness, sleep disturbance, fatigue, slowness, etc.)
- Signs (LOC/impaired consciousness, impaired coordination, impact seizure, gait abnormality, slow to answer or follow direction, distracted-poor concentration, inappropriate behavior, vomiting, vacant stare, slurred speech, personality change, poor game/practice play, etc.)
- Neuropsychological testing is a cornerstone tool
- Neuroimaging not essential for uncomplicated concussion
- Research is ongoing and vital
- Rehabilitation program essential
- Return-to-play protocol (no activity until asymptomatic → light aerobic activity → sport-specific training → noncontact drills → full contact practice → game play → able to progress as long as asymptomatic)
- Prevention—rule changes and enforcement of rules are vital
- Education—promotion of fair and ethical play in sports

to C6) developing degeneration sooner. Some physicians have divided cervical disc disease into soft and hard disc diseases. The soft disc refers to an acute disc herniation, usually in a younger person with or without neurologic involvement, depending on the degree of mechanical or chemical irritation by the nucleus pulposus on the nerve root structures. The hard disc is the result of events occurring over time and the progressive decline in the cervical disc and surrounding structures resulting in spondylosis. This degeneration process is characterized by disc bulges, herniations, thickening of the posterior longitudinal ligament, thickening of ligamentum flavum, degeneration of the synovial facet joints, or osteophytes. These anatomic changes result in a greater likelihood of structural compromise or impaired spinal kinematics. These changes may result from the athletes’ sport or, as with the general population, chronic impaired loading through altered posture and inappropriate loading of the various spinal structures over time. Of specific note, radiographic changes correlate poorly with signs and symptoms that patients exhibit (Figs. 53-1 to 53-3) (32,33).

Cervical Cord Neuropraxia

Cervical cord neuropraxia results from an injury to the spinal cord presenting with complete transient paralysis, motor impairment, sensory changes, or a combination of these in at least two limbs. Transient quadriplegia is a rare cervical spine injury caused by axial loading of the neck in extension or flexion that results in nonpermanent complete sensory and motor impairment in all four limbs resolving in 10 minutes to 48 hours (34). One should remember that any neurologic impairment in more than one extremity is a central nervous system injury unless proven otherwise (sometimes referred to as burning hands syndrome). Obersteiner originally described this as a spinal cord concussion in 1879 (35). It has been estimated at 1.3 episodes per 10,000 athletes or an incidence of 6 h 10,000 athletes (36).

It appears that this is a neuropraxic injury from local cord anoxia and increased intracellular calcium (37). During hyperextension, a “pincer” effect can occur, causing the posterior inferior aspect of the superior vertebra and anterior superior aspect of the inferior vertebra to produce a resultant compression of the spinal cord. In this position, the diameter can narrow up to 30% (38). With associated spinal degeneration, even greater compromise can occur.

In the event of either a significant head injury in which a concomitant neck injury may be present or isolated cervical injury, the medical team should be appropriately prepared by having the correct equipment available and having practiced the evaluation and treatment of the athlete with this type of injury. This includes the possibility of spine-boarding the athlete and removal of the athlete’s protective gear (i.e., facemask removal). An athlete who is wearing protective headgear (e.g., a helmet) with a suspected spinal injury should not have the helmet removed. It is imperative that the sideline physician understands how to protect the spine and coordinate spine-boarding and transportation of an athlete with a suspected spine injury.

In the evaluation of an athlete with suspected cervical disc disease, the location, referral pattern, quality of the symptom, and associated neurologic symptoms—such as Lhermitte’s sign, balance/gait abnormalities, limb weakness, or bowel/dysfunction—must be obtained. This should occur to determine if additional structures have been damaged. A complete physical examination (including cervical tenderness and ROM), a detailed neurologic examination of the
upper and lower limbs (including power, sensation, reflex assessment, and a Spurling's maneuver), and evaluation for upper motor neuron signs (Hoffman, plantar responses, tone, etc.) are all necessary.

X-rays may be helpful in evaluating the athlete for a fracture, spinal instability (with the use of flexion/extension views), osseous degeneration, or destructive lesions from tumor or infection. Anteroposterior, lateral, open-mouth, and oblique views should be the initial evaluation. If pain or neurologic signs persist or ROM has not returned, further imaging should be considered. Magnetic resonance imaging (MRI) is probably the study of choice if not contraindicated. A fat-suppressed image should be included to evaluate for ligamentous injury, especially when instability is seen on radiographs or clinically suspected. The spinal diameter and relative space can be evaluated. Normal midsagittal diameter on plain x-rays should be 14 to 23 mm. If the canal is less than 13 mm, it is considered stenosis (39).

Torg et al. introduced a ratio to assist with measuring and stratifying persons at risk for cervical cord neuropraxia from spinal stenosis (40). A lateral cervical spine x-ray is obtained and measurements are made. The Torg ratio is the distance from the midpoint of the posterior aspect of the vertebral body to the nearest point on the corresponding spinolaminar line divided by the anterior to posterior width of the vertebral body. It is of limited utility in screening athletes for participation due to its low specificity and positive predictive value despite its high sensitivity.

Return to Play

Return-to-play considerations should initially include normal ROM or thorough imaging evaluation—if ROM is abnormal—and a normal neurologic examination (41,42). There are absolute and relative contraindications to returning an injured athlete to play (Table 53-4).

With traumatic lesions, some athletes may return to sport without limitations, depending on the injury, healing, and treatment. Atlantoaxial instability noted on lateral flexion-extension views, rotatory fixation from computerized tomography (CT), fused C1 to C2 segments, and acute spine fractures are an absolute contraindication for return to play. Healed nondisplaced Jefferson's fractures, odontoid fractures (type I and II), and healed lateral mass C2 fractures with pain-free full ROM and no abnormal neurologic signs are relative contraindications.

If lateral flexion-extension x-rays demonstrate greater than 3.5 mm of horizontal translation from one vertebral body to the adjacent in the middle and lower cervical spine, or greater than 11 degrees of angulation of adjacent vertebral bodies, return to play is absolutely contraindicated (42,43). Healed fractures of a stable displaced vertebral body compression fracture without an anteroposterior sagittal component or a stable fracture involving the posterior neural ring in individuals who have a normal examination have a relative contraindication. If the examination is abnormal, or if there is superimposed ligamentous laxity, then the contraindication is absolute.

Acute intervertebral disc injuries are an absolute contraindication, but when they have resolved and the patient has a normal exam, the athlete may return to play. Abnormal neurologic signs, loss of motion, an acute “soft” herniation superimposed onto “hard” preexisting abnormalities with/without congenital stenosis, and loss of the normal cervical lordosis are all absolute contraindications.
A one-level fusion with normal alignment, motion, and examination incurs no restriction. A two- to three-level fusion and normal examination represent a relative contraindication, whereas a fusion extending more than three levels represents an absolute contraindication to play (42).

**Proximal Nerve Pathology**

The most common cervical root and brachial plexus injury in an athlete is a “burner or stinger.” An athlete (most commonly a football player) sustains a trauma resulting in symptoms of pain and paresthesias down one limb (44). Symptoms usually last less than 1 minute and typically there are no abnormal neurologic signs on examination. Two common mechanisms of trauma have been described. First, the cervical spine may be side bent and extended, causing compression on an exiting nerve in the intervertebral foramen. Alternatively, the shoulder may be forcefully depressed while the head is driven in the opposite direction, causing traction on the nerve. A direct blow to the plexus can also result in dysfunction, but less commonly (45). The injury usually affects the upper cervical nerves of the plexus (C5 to C6), but can extend to involve additional portions of the plexus. The strength, sensation, and reflexes of the upper limbs should be evaluated on physical examination.

Electrodiagnostic studies can be helpful in localizing the site of neurologic injury (cervical roots vs. brachial plexus vs. peripheral nerve) and can be useful in objectifying the severity. They are not as helpful in determining whether an athlete is ready to return to participate because of the delay between nerve recovery and the normalization of electromyographic (EMG) findings. Once diagnosed, radiographs are usually indicated if symptoms are persistent or recurrent. They will most often be normal, but may show evidence of degenerative disc disease or intervertebral foraminal narrowing. A physical therapy program should be designed that addresses tackling technique, neck ROM, postural alignment, muscle shortening from spasm, strengthening of the involved muscles, and muscle retraining/proprioception so the neck and shoulder stabilizers “functionally” protect the neck and nerves from further potential injury. Adequate fitting of shoulder pads and lifters should be employed, and the use of a neck orthosis such as the “Cowboy Collar” should be considered. Neck orthoses are theorized to prevent excessive cervical extension, thereby decreasing the potential traction or compression that may result from contact. Recurrent bunners can be a serious problem. For those athletes who have recurrences or are left with neurologic deficits, an MRI may be needed. Since the long-term history of burners has not been studied, it is not known whether they have more long-lasting consequences.

It is very important to remember that “bilateral stingers” are extremely rare. If an athlete reports upper-limb symptoms in both arms, a spinal cord injury must be ruled out first and foremost.

**SHOULDER**

The rotator cuff includes the supraspinatus, infraspinatus, teres minor, and subscapularis (SITS) muscles. Its function is to maintain the humeral head centered on the glenoid (45a) during arm motion. The “shoulder” is made up of three true joints and one pseudojoint. These include the sternoclavicular joint, acromioclavicular joint, glenohumeral joint, and scapulothoracic pseudoarticulation. The glenoid and acromion are portions of the scapula. Dynamic muscular control of the scapula is vital to the function of the shoulder as a cohesive unit (46).

On physical examination, active and passive ROM should be evaluated, including isolated testing of the glenohumeral joint and scapulothoracic rhythm. Strength, sensation, and reflexes in the upper limb should be evaluated to completely evaluate the shoulder as well as rule out concomitant neurologic injury. Specialized testing for impingement, instability, and labral and rotator cuff pathology should be performed as indicated.

**Rotator Cuff Injury**

Rotator cuff tendinitis typically occurs either from scapular dysfunction causing subsequent impingement or rotator cuff dysfunction, allowing excessive movement of the humeral head and resultant impingement. Classically, Neer classified impingement into three stages (47). Stage 1 consisted of edema and hemorrhage of the tendon; stage 2, the development of fibrosis and tendinitis, usually occurring in patients older than age 25; and stage 3, the degeneration, bony changes, and tendon rupture, usually occurring after age 40. Jobe and Pink added a fourth stage, defining it as rotator cuff tears greater than 1 cm (48).

With chronic, impaired motion, degeneration or tearing of the tendons can result. Weakness of the scapula stabilizers, especially when there is an underlying nerve injury (i.e., long thoracic neuropathy), can lead to inadequate scapula rotation during shoulder elevation. Muscle incoordination of the same muscles can similarly lead to inadequate scapula rotation and further narrow the outlet through which the supraspinatus tendon passes. Additionally, excessive scapula motion can cause repetitive irritation of the undersurface of the acromion, leading to remodeling as demonstrated as spurs or acromial “hooking.” This bony degeneration can cause narrowing and repeated impingement with damage to the rotator cuff. It is easy to appreciate that weakened, deconditioned, and dysfunctional muscles can precipitate a viscous cycle (Fig. 53-4).

The supraspinatus tendon has a vulnerable area (watershed) where its vascular supply can be compromised by tendon lengthening, thus causing ischemia. Isolated bicipital tendinitis as the primary cause of shoulder pain is uncommon and exists along the continuum of rotator cuff tendinitis secondary to its related proximity. Impingement is a result of
biomechanical dysfunction of the shoulder and itself is not a pathologic diagnosis. This impingement causes pathology by repetitive compromise of the soft tissues of the shoulder girdle (Fig. 53-5).

Shoulder Dislocation/Instability

Thomas and Matsen classified shoulder instability into two groups (49): TUBS, which stands for traumatic, unidirectional, Bankart, surgery, and AMBRI, which stands for atraumatic, multidirectional, bilateral, rehabilitation, and inferior capsular shift. These entities are not mutually exclusive, but give a construct with which to think about shoulder pathology.

Dislocation treatment depends on the patient’s age, activity level, and the number of subluxations and dislocations experienced. Epidemiological studies of shoulder dislocations have shown an approximate 90% redislocation rate in those younger than age 20, moderate risk from age 20 to 40, and an approximate 10% rate for those older than age 40 (50). There appears to be no consensus of the period of immobilization after a dislocation and the resultant recurrence rate. Periods from a few days to 6 weeks, with and without a rehabilitation program, have been attempted. The redislocation risk may depend, in part, on the patient’s age, amount of trauma sustained, postdislocation activity level, and postdislocation rehabilitation program.

For the shoulder to maximally perform its function, it sacrifices bony stability for mobility. This highly mobile joint uses dynamic muscular control as its greatest stabilizer. Additional factors that contribute to stability are glenoid concavity, whose surface is increased by the labrum; negative intraarticular pressure; capsule and ligaments; and compression from the rotator cuff muscles. The static stabilizers include the anterior band of the inferior ligament complex for anterior translation, posterior band of the inferior glenohumeral ligament complex for posterior motion, and the superior glenohumeral ligament for inferior movement (51). The dynamic stabilizers include the rotator cuff muscles. These stabilizers contribute to varying degrees, depending on arm position.

On physical examination, one can use the apprehension and relocation test, sulcus sign, load and shift test (drawer), or “jerk” test. It is important to assess for generalized ligamentous laxity and the strength of the scapular and rotator cuff muscles, as well as ruling out neurologic causes for weakness and instability. In addition, many patients will also have posterior capsular tightness, which needs to be assessed and addressed in the rehabilitation of the athlete.

Rehabilitation of the Shoulder

Initially, pain control and inflammation reduction are required to allow progression of healing and the initiation of an active
rehabilitation program. This can be accomplished with a combination of relative rest, icing (20 minutes three to four times a day), electrical stimulation, and acetaminophen or an NSAI. This is the initiation of the therapy program. The following progression can be added as the patient tolerates. If these are not sufficient, more aggressive intervention, such as corticosteroid injections, can be considered. Passive modalities in isolation have never been proven to be of benefit.

After the pain has been managed, restoration of motion can be initiated. Specifically, internal rotation should be addressed because it is lacking in most patients. Poor internal ROM results from a tight posterior capsule or dysfunction of the rotator cuff, causing the humeral head to translate forward. With repetitive motions, this can cause a great deal of shearing on the anterior labrum and eventually result in either degeneration of the labrum or tearing.

The third phase of treatment is strengthening and should be performed in a pain-free range. The scapulothoracic stabilizers should first be addressed because the glenoid is the platform on which the arm moves. Strengthening can then progress to the rotator cuff muscles and then to the shoulder’s prime movers.

The fourth phase is proprioceptive training. This is important in order to retrain neurologic control of the strengthened muscles. It provides improved dynamic interaction and coupled execution of tasks for harmonious movement of the shoulder and arm.

The last phase of rehabilitation is to return to task- or sport-specific activities. This is an advanced form of proprioceptive training for the muscles to relearn prior activities. This is important and should be overseen to be sure that the task is performed correctly and to eliminate the possibility of reinjury or injury in another part of the kinetic chain from improper technique. The rehabilitation begins at a cognitive level, but must be practiced until it transitions to unconscious motor programming (52).

As with all musculoskeletal rehabilitation, the entire body must be taken into consideration. Abnormalities in the kinetic chain can also affect the shoulder. If there are restrictions or limitations in ROM or strength, the forces will be transmitted to other portions of the kinetic chain, resulting in an overload of those tissues and possible injury.

Nerve Injuries About the Shoulder
Several peripheral nerve injuries in the region of the shoulder are seen in the athletic population. Neuropathies of the axillary, suprascapular, and long thoracic nerves can be the result of the idiopathic condition Parsonage-Turner syndrome. All three can also be the result of trauma. The axillary nerve, which innervates the deltoid and teres minor, is the most commonly injured with anterior shoulder dislocations (53). The suprascapular nerve, which innervates the supraspinatus and infraspinatus, is rarely associated with glenohumeral dislocations, but can be injured from repetitive trauma that is probably the result of traction. Entrapment can also occur at the suprascapular notch, leading to involvement of both the supraspinatus and infraspinatus, or at the spinoglenoid notch, leading to involvement of only the infraspinatus. Compression can occur at either notch as the result of a ganglion cyst or a hypertrophied ligament. The largest case series of suprascapular neuropathies at the level of the spinoglenoid notch has been reported in volleyball players. Other sports with a high incidence of suprascapular neuropathies include tennis, baseball, and weight lifting. Long thoracic neuropathies have been reported in large numbers, but as individual case reports. The problem is that a good biomechanical etiology has not been well established. These nerve injuries not only lead to weakness of the serratus anterior and winging of the scapula, but also lead to inadequate scapula rotation during shoulder abduction. This commonly results in supraspinatus impingement syndrome.

Brachial Plexopathies
Brachial plexus injuries in sports are uncommon but can be devastating, with potentially lifelong consequences. They most commonly occur as the result of a “burner or stinger.” Classifying these injuries can be somewhat confusing, depending on how one defines the brachial plexus. The anatomist defines the plexus as beginning at the point where the cervical spinal nerves have joined to form the upper, middle, and lower trunks. Injuries proximal to the trunks would be considered cervical root or spinal nerve injuries. Clinicians more commonly define the cervical roots as part of the plexus. Based on the relationship of the dorsal root ganglion to the site of injury, they further define injury to the plexus as preganglionic or postganglionic. Because it is our contention that most burners are cervical root injuries, one can label these as preganglionic plexus as well. This type of classification would greatly increase the incidence of plexus injuries.

Probably the second most common cause of brachial plexus injuries are shoulder dislocations. These are usually anterior glenohumeral dislocations and can injure either a cord (usually the posterior cord) or an isolated nerve (usually the axillary nerve). These patients often have their shoulders immobilized for at least a brief period of time so that more subtle neurologic deficits may not be immediately recognized. Electrodagnostic studies are often necessary to objectify weakness and rule out a nerve injury.

Parsonage-Turner syndrome is an idiopathic cause of brachial plexitis. Strictly speaking, this is not a sports injury, but because of its idiopathic nature, the athlete often tries to attribute an athletic event or activity as the cause of the nerve dysfunction. The classic presentation of acute onset of pain lasting 1 to 2 weeks with no clear trauma should make one suspect this diagnosis. Weakness is delayed in its onset, and the pain usually will abate spontaneously. Symptoms will often follow a viral illness, and at the time of exam, most provocative maneuvers are negative. There is no current treatment, although antiviral drugs may prove to play a role in the future. Management usually is the palliative treatment of the pain followed by appropriate physical therapy when residual neurologic deficits exist. The most commonly involved nerves are the upper trunk.
of the brachial plexus, suprascapular, axillary, and long thoracic nerves. Long-term prognosis is usually good, with most patients having complete recovery within 2 years (54).

Isolated trauma to the plexus in the absence of a shoulder dislocation or burner is uncommon. Shoulder dislocation or burner usually occurs in conjunction with fractures of the humeral head/neck or the clavicle. A direct blow from a helmet or a significant impact distracting the head or shoulder can injure the plexus.

Thoracic outlet syndrome does not appear to have an increased incidence in athletes and is a relatively uncommon disorder, especially the neurogenic form. Athletes will usually report neurologic symptoms that correspond to the medial cord or lower trunk of the brachial plexus.

ELBOW

The elbow is comprised of the distal humerus and proximal ulna and radius. The humeroulnar joint allows flexion and extension. The humeroradial articulation occurs between the capitulum of the humerus and head of the radius. The proximal radioulnar articulation (superior radioulnar joint) is a pivot joint allowing rotation of the radius about the ulna for pronation and supination.

Tendinitis

The muscles (in descending order of frequency) most commonly involved in lateral epicondylitis are the extensor carpi radialis brevis and extensor digitorum communis; in medial epicondylitis are pronator teres, flexor carpi radialis, palmaris longus, flexor carpi ulnaris, and flexor digitorum superficialis. Posterior elbow tendinitis involves the triceps muscle. Tendinitis about the elbow appears to develop from overuse (demand greater than supply) or abnormal loading of the tendon. Nirschl and Pettrone noted that the histology of lateral epicondylitis (primarily affecting the extensor carpi radialis brevis) was “angiofibroblastic hyperplasia” and subsequently coined the term “angiofibroblastic tendinosis” (46). For the first time, it was theorized that this condition was a degenerative process and not inflammatory; in fact, no inflammatory cells were identified. In 1980 and 1981, they noted the same changes in medial and posterior epicondylitis, respectively. These authors feel that these conditions are primarily degenerative processes (55).

In lateral epicondylitis (aka, tennis elbow), pain is typically produced with palpation approximately 1 to 2 cm distal to the extensor origin. There is pain with resisted wrist extension and supination, and passive wrist flexion. Once diagnosed, the most important treatment aspect is to identify the offending behavior (e.g., poor tennis mechanics, repetitive job or recreational activity, etc.) and attempt to modify it. Restoration of motion and strength should not be forgotten as important components of the rehabilitation.

Medial epicondylitis is felt to occur from microtearing of the common flexor tendon’s origin of the medial epicondyle or subsequent failed healing response that alters the normal tendon biomechanics (56). Initially, this develops as an inflammatory response, but can then develop into a degenerative process as noted above. This is common in throwing athletes and can occur in the trail arm of golfers (the so-called golfer’s elbow). Chronic valgus stress to the medial elbow can lead to laxity of the ulnar collateral ligament and instability of the elbow.

ROM of the elbow, including pronation and supination, varus and valgus stress testing, and an upper-limb neurologic examination, should be incorporated with the patient’s history for evaluation. Flexibility of the wrist flexors and extensors as well as strength testing at the wrist can clarify reasons for patients’ pathology.

Neuropathies at the Elbow

The ulnar, median, and radial nerves are the three major nerves that traverse the elbow. The ulnar nerve lies posterior in the ulnar groove and is the most susceptible to injury. Additional physical exam maneuvers should be employed for further evaluation. With palpation of the ulnar nerve in the groove, one may be able to cause it to subluxate. Tinel’s sign produced by percussion over a nerve denotes membrane instability. The elbow hyperflexion test can place added tension to the nerve reproducing the patient’s symptoms. Care must be taken to avoid concomitant wrist flexion. With stress or stability test, traction of the nerve may occur causing provocation of symptoms.

Cubital tunnel syndrome is entrapment of the ulnar nerve in the cubital tunnel where the two heads of the flexor carpi ulnaris form a rooflike structure that overlies the ulnar nerve. It has been used more generally as a term indicating compression of the ulnar nerve in the region of the elbow. This may occur from “internal” compression by adjacent soft tissues or external compression on the nerve in the ulnar groove, or from repeated stretching by elbow overload. This can present with pain or paresthesias along the ulnar portion of the forearm with or without accompanying fourth- and fifth-digit symptoms.

The radial nerve innervates the triceps and anconeus in the arm and provides branches to innervate the brachioradialis and extensor carpi radialis longus, then transverses the elbow joint to the forearm. Once in the forearm, it splits into the superficial and deep branches. The deep branch is mainly motor, although noncutaneous sensory fibers also exist. It becomes the posterior interosseous nerve after its innervation of the extensor carpi radialis brevis and supinator. A lesion or entrapment of the posterior interosseous nerve must be considered in patients with presumed refractory lateral epicondylitis. Patients have lateral elbow aching localized to the region of the extensor origin or symptoms may be referred distally to the wrist. There can be reports of paresthesias in the lateral forearm or hand.

WRIST AND HAND INJURIES IN SPORTS

The wrist is comprised of two rows of four bones each. The proximal row contains the scaphoid, lunate, triquetrum, and
pisiform, and the distal row includes the trapezium, trapezoid, capitare, and hamate.

Palpation of each of the carpal bones and the digits can be performed easily because there is little soft tissue covering them. The mobility and stability between the bones also can be tested for possible ligament injuries. The joints of the digits can be tested with volar/dorsal and varus/valgus forces to test for ligament compromise.

**Common Wrist Fractures**

The most commonly fractured carpal bones are the scaphoid and triquetrum (57). Scaphoid fractures are often missed, and a high index of suspicion should be maintained with persistent radial wrist pain after trauma, even if initial x-rays are negative. The scaphoid and lunate are the bones most susceptible to avascular necrosis and should be monitored accordingly.

The wrist also includes the distal portions of the radius and ulna, as well as the distal radioulnar joint. A distal radius fracture is one of the most common injuries treated by orthopedists. Colles’ fractures are radial metaphyseal injuries usually of low energy and do not involve the articular surface.

The thumb can sustain extraarticular and intraarticular fractures and crush injuries. Axial loading of the thumb can commonly cause a Bennett’s or Rolando’s fracture, which occurs at the base of the metacarpal.

With suspected fractures, excessive motion or force may worsen the injury and should be avoided. A detailed knowledge of the wrist anatomy is important to guide the palpation and development of a differential diagnosis. With continued symptoms, a high index of suspicion should be maintained because fractures can be missed on initial x-rays. As always, the neurovascular status should not be overlooked.

**Ligament Injuries of the Wrist**

The most common carpal instability occurs between the scaphoid and lunate. Patients typically have a traumatic event with resultant dorsal- and radial-sided wrist pain and loss of motion with reduced grip strength. Physical examination should include palpation of the scapholunate junction and a scaphoid shift test, or Watson’s test (58), which will result in pain and increased motion. This can be identified by stress x-ray with a clenched fist ulnar deviated view if standard views fail to demonstrate flexion of the scaphoid and extension of the lunate. The fist places tension across the scapholunate ligaments, and increased joint space is seen. If the images are difficult to interpret, comparison views to the contralateral limb can be obtained. The second most common ligament injury is a lunotriquetral tear. Examination should include palpation of the joint and a lunotriquetral shear test (58).

**Ligament Injuries of the Thumb**

Gamekeeper’s thumb or skier’s thumb is a very common injury of the ulnar collateral ligament. When torn or attenuated, an athlete’s key pinch grip is severely impaired. When evaluating this injury, one should keep in mind that complete tears can result in entrapment of the adductor aponeurosis in the metacarpophalangeal (MCP) joint, or what is referred to as a Stener’s lesion. This will impair healing and cause sustained laxity. A reverse gamekeeper’s thumb refers to an injury of the radial collateral ligament and will not develop a Stener’s lesion. On examination, one can stabilize the radial portion of the metacarpal phalangeal while applying a force distally to stress the ulnar collateral ligament (Fig. 53-6). A grade I causes pain and no increased motion, whereas a grade III causes no pain from the absence of an intact ligament and continued motion while stressing. A grade II demonstrates increased opening with pain on stressing. When stressing the joint, care must be taken to avoid overstretcing the ligament and causing further injury. The thumb should be tested in extension and slight flexion to evaluate all fibers of the ligament (see Fig. 53-6).

**de Quervain’s Stenosing Tenosynovitis**

De Quervain’s stenosing tenosynovitis is inflammation of the first dorsal compartment (abductor pollicis longus and extensor pollicis brevis [APL and EPB]) and can occur with overuse in sports in which equipment is gripped and the wrist ulnar deviated, such as golfing, fly fishing, and playing racquet sports. Radial wrist pain is noted with resisted thumb extension or palpation over first dorsal compartment, which should elicit the patient’s symptoms. Many patients may have concomitant irritation of the superficial radial nerve, which partially overlies this area.

**The Triangular Fibrocartilage Complex**

The triangular fibrocartilage complex can be a cause of ulnar-sided wrist pain, along with tendinitis, fracture, or impaction. The fibrocartilage begins at the distal radius, blending with the hyaline cartilage of the lunate fossa. It narrows toward

![FIGURE 53-6. Skier’s thumb examination: A valgus force is applied to the interphalangeal joint of the thumb to stress the ulnar collateral ligament.](image-url)
the ulna, with deep fibers inserting into the fovea of the ulnar styloid. The cartilage thickens dorsally and palmarly, forming the radioulnar ligaments. There is a meniscal homologue from the distal ulna to the triquetrum, with an interval between the ulnotriquetral ligament known as the pre styloid recess (59). Understanding the anatomy guides palpation of the structures and development of a diagnosis. There exists little research on the natural history of these tears, which appear to do poorly with or without surgical intervention. It is unknown how many triangular fibrocartilage complex tears exist in asymptomatic persons. Shearing forces to entrap the tear or cause pain have been attempted with ulnar deviation and loading while pronating and supinating the hand. A positive test has been defined as a “pop” or “click” and not isolated ulnar wrist pain. The sensitivity and specificity of the test are unknown.

**Injuries to the Digits**

Coach’s finger or a jammed finger is most commonly a proximal interphalangeal dislocation reduced by the coach or athletic trainer. The dorsally directed dislocation is the most common, and volar plate injuries or fractures must be considered. If stable, it can be treated nonoperatively with splinting. Often this is an injury not mentioned to medical staff by the athlete and is self-reduced, but it can become stiff and painful if not treated appropriately.

Trigger finger is a rupture of the terminal extensor tendon of the proximal phalanx causing loss of active extension. Active extension of the distal joint must be carefully evaluated to avoid missing this injury. This is usually caused by forced flexion of the distal phalangeal joint and can be caused by being hit on the end of the finger with a ball, such as a baseball or basketball.

Jersey finger is the avulsion of the flexor digitorum profundus from the distal phalanx. This occurs most commonly in the fourth finger (60). Usually, an athlete is attempting to tackle a player when his finger is caught in the jersey, and forced extension occurs against the contracting long flexor. Active flexion of the digit should be evaluated. Retraction of the tendon into the hand usually occurs unless stopped by an attached bone fragment.

Trigger fingers are thickening of the proximal portion of the flexor tendon sheath, which may develop from chronic irritation of the palmar surface of the metacarpal phalangeal joint by stick or racquet. The digit may become “stuck” and fixed in flexion until passively maneuvered because the thickened sheath becomes hung up at the A-1 pulley. This “sticking” can usually be demonstrated by the patient on exam. If it is unclear, one can have the patient actively flex the digit and while palpating the volar MCP region, passively extend the digit. One usually feels a nodule, which “pops” during extension.

**Nerves of the Wrist and Hand**

The wrist and hand region includes the radial, median, and ulnar nerves. The superficial radial sensory nerve is superficial over the distal lateral radius and vulnerable to injury via contusion. It can also be caught between the brachioradialis and extensor carpi radialis brevis with repetitive pronation/supination. The median nerve can become contused, but more commonly is irritated by repetitive motion of either the wrist or flexor tendons. Athletes who grip equipment as part of their sport, including cycling, baseball, and racquet sports, should consider wearing gloves. The ulnar nerve can become contused or compressed at Guyon’s canal and is an additional reason for athletes to wear protective hand gear.

Strength and sensation should be tested. One should be mindful of testing different sensory modalities accurately, including pinprick, light touch (avoiding light pressure), two-point discrimination, and temperature.

**Differential Diagnosis for Low Back Pain in Athletes**

**Lumbar Disc Injury**

The intervertebral disc is a hydrodynamic elastic structure that has two components: the annulus fibrosus and nucleus pulposus. In early life, until the later thirties, blood vessels passing to the end plate are progressively obliterated, with most nutrition provided for via the process of imbibition, in which compression and distraction force nutrients into the disc. The annulus fibrosus is composed of layered sheets of collagen fiber, with the individual fibers enhancing overall mechanical efficiency. The nucleus pulposus is composed of a homogenous mucopolysaccharide matrix containing a network of fine protein-based fibrils. Biomechanically, intrinsic disc pressure functions to separate vertebral end plates and maintain tension within annular fibers. Flexion and extension forces are well tolerated by the annulus, whereas the addition of rotational forces causes excessive stress to the annulus leading to failure.

Lumbar disc injury is a relatively uncommon injury in the young competitive athlete (61). Micheli and Wood noted that 11% of adolescents as compared with 48% of adults with low back pain had evidence for disc injury (62). Although uncommon, there may be more significant changes to the structure of the disc in athletes as compared with nonathletes. Sward et al. noted disc degeneration in 75% of male gymnasts as compared with 31% in an age-matched group of nonathletes (63). Various sports have been evaluated in regard to the potential increased risk of disc herniation. Mundt et al. found only a weak association between bowling and herniation of the lumbar disc, with no significant association found in other sports, including baseball, softball, golf, swimming, diving, weightlifting, or racquet sports (64).

**Lumbar Facet Syndrome**

The lumbar zygapophyseal joints are true synovial joints and are subject to inflammation. It is of key importance to understand the anatomy of these joints, as it will clearly influence diagnostic and treatment decisions (65). The lumbar facet
joints maintain a sagittal orientation from L1 to L4, with a more coronal orientation encountered at L5 to S1. These joints are innervated by the posterior primary ramus, which supplies at least two zygapophyseal joints, with each joint receiving innervation from at least two spinal levels. The joints allow between 2 and 3 degrees of rotation at each segment within the lumbar spine and account for between 12% and 24% of compressive load, the remainder of which is absorbed by the intervertebral disc. The amount of compressive load increases as the intervertebral disc height decreases, such as is seen with degenerative disc disease. The facet syndrome may thus result from degeneration of the disc or the joint itself, positional overload such as may be seen with any repetitive overuse injury and secondary to trauma (66,67).

The athlete often describes a sudden flexion/extension maneuver often combined with rotation/side bending as the precipitating event. Lumbar facet syndrome commonly occurs in football, volleyball, gymnastics, figure skating, golf, and tennis, in which end-range sagittal and transverse plane motions are combined. No significant research has been performed to identify the true incidence and prevalence of facet syndrome, which is most likely secondary to the difficulty in obtaining reliable objective measures.

**Lumbar Spondylolysis/Spondylolisthesis**

Spondylolysis or fracture of the pars interarticularis can be related to an acute traumatic event or an overuse phenomenon causing stress fracture (68). Micheli and Wood reported spondylolysis as the final diagnosis of 47% of adolescents between the ages of 12 and 18 years presenting with reports of low back pain (62). Sports that require repetitive hyperextension will more commonly predispose to the development of a pars fracture. Rossi noted that 63% of divers, 36% of weightlifters, 33% of wrestlers, and 32% of gymnasts had evidence for spondylolysis as compared with 5% of the general population (Figs. 53-7 and 53-8) (69).

Spondylolysis is defined by a forward or backward subluxation of one vertebrae on another. There are several types of spondylolisthesis: the isthmic type has an anatomic defect in the pars interarticularis, dysplastic types have structurally inadequate posterior elements, degenerative types are the result of significant degenerative changes of the zygapophyseal joints and deficient supporting ligaments, traumatic types are the result of fracture of the posterior elements other than the pars, and pathologic types are the result of metabolic, malignant, or infectious disease (70). The Meyerding classification system separates slippages into 25% intervals, with a grade I defined by a 0% to 25% slip, grade II with a 26% to 50% slip, grade III with a 51% to 75% slip, and grade IV with a slip between 76% and 100% (71). Spondylolisthesis is an uncommon occurrence in the competitive athlete and is most likely secondary to superior dynamic muscular stabilization and skeletal maturity. Wilse et al. noted its occurrence in skeletally immature adolescent athletes between the ages of 9 and 14, and it was rarely seen in athletes above this age range (71).

**Sacroiliac Joint Dysfunction**

The sacroiliac joints are weight-bearing joints between the articular surfaces of the sacrum and ilium, which are located on the lateral surface of the sacrum. They are part synovial joint and part syndesmosis, with the synovial portion being the anterior and inferior one third of the joint (72). There is hyaline cartilage on the sacral side and fibrocartilage on the ilial side. There are no muscles that directly control movement of the sacroiliac joints, but many indirectly affect movement. Sacroiliac joint movement is mainly passive in response to the action of surrounding muscles. The psoas and piriformis muscles pass anterior to the sacroiliac joints, and imbalance of these muscles in particular may affect sacroiliac joint function. Imbalance in the length and strength of the piriformis strongly influences movement of the sacrum. Sacroiliac joint dysfunction occurs when there is an alteration of the structural or positional relationship between the sacrum and a normally positioned ilium (73). The sacroiliac joint plays a small but significant role in the cause of low back and buttock pain,

![FIGURE 53-7. A lateral lumbar roentgenogram of the lumbar spine demonstrating a lucency of the pars interarticularis, which represents a fracture.](image-url)
although the true incidence is unknown. It has been reported to occur in elite cross-country skiers, rowers, and gymnasts (74–76). It is a true synovial joint with extensive innervation from the lumbosacral region, accounting for the difficulty differentiating sacroiliac joint dysfunction from that of surrounding structures. Various medical conditions, such as rheumatologic disorders, infection, and neoplasms, may also affect the joint, and it is extremely important for the sports physician to rule out the possibility of sacral stress fracture. Johnson et al. demonstrated sacral stress fractures in five collegiate athletes presenting with pain over the region of the sacroiliac joint (77). It is therefore prudent to perform further diagnostic workup in those subjects diagnosed with sacroiliac dysfunction who have a history of or risk factors for stress fracture or remain refractory to conservative care.

**COMMON LOWER EXTREMITY INJURIES IN SPORTS**

**Hip**
The hip can be injured in a variety of sports, either from direct trauma or musculotendinous overload. In either case, the athlete is often greatly impaired in sport participation as a result of significant alterations in gait. It is important to quickly determine the etiology of the problem, allow for a period of rest, and maintain strength and flexibility as these injuries heal. Almost all respond to the previously outlined course of functional rehabilitation.

**Hip Pointer**
The “hip pointer” injury is a direct blow to the pelvic brim or hip region, which results in a contusion to the soft tissues and often the underlying bone. It is relatively common in sports such as football and hockey, where there are many collisions (both player to player as well as contact between the player and the field or arena). Common areas include the greater trochanter and iliac crest. The contact can result in hematoma formation, but often there is little visible swelling or ecchymosis. There is, however, a significant amount of pain and focal tenderness that is due to bony contusion and periosteal irritation. The athlete usually has difficulty with quick bursts of running and with any contact to the area.

The diagnosis is made by the above history or on field observation. On examination, it is important to note full ROM of the hip and the knee. If there is a great deal of pain with passive ROM, then x-rays should be obtained to rule out any significant bony pathology. For most injuries, imaging studies are not necessary. In cases of significant soft-tissue swelling, additional imaging such as CT or MRI may be indicated.

The treatment requires frequent and repeated icing, active ROM, and a period of rest until gait can be normalized. In a very painful hip, crutches with weight bearing as tolerated may be necessary for pain control and to unload the hip. NSAIDs can be helpful in the very early period to assist with pain control and inflammation but can generally be discontinued in less than a week. Once an athlete has no pain at rest or with running and jumping activities and minimal pain on palpation, he or she can be returned to play, usually with protective padding. Appropriate padding over the area is important to protect the area from recurrent injury. This can be coordinated with the trainer and generally consists of soft foam over the area or doughnut cutouts surrounding the area. Most of these injuries do not result in long-term sequelae.

**Hip Flexor Strain**
Hip flexor strains are commonly seen in sprinting as well as in other sports, such as soccer, baseball, and football. They occur as a result of an eccentric overload of the psoas muscle or as the athlete attempts to flex the fully extended hip, such as in hurdles. Most athletes are unable to continue to run, and for most sports, this means discontinuing participation. The examination is straightforward, with tenderness to palpation over the area and with resisted hip flexion and passive hip extension. ROM may also be painful; therefore, the majority of these injuries require plain x-rays of the hip (usually an anteroposterior and frog-leg lateral view) to exclude bony injury. This is particularly important in the adolescent or skeletally immature athlete, as injury to the apophyseal plate can commonly occur.
These various avulsions, whether in the adult or adolescent athlete, can also be treated nonoperatively.

Treatment of hip flexor strains and/or avulsions consists of protected weight bearing when there is a significantly antalgic gait, aggressive icing, and gentle active ROM as soon as possible. Strengthening exercises of the lower extremities should be avoided until the gait is nonantalgic and ROM is full and pain free. Then the athlete should be progressed through an aggressive strengthening consisting of both open and closed kinetic exercises. Eccentric and plyometric training should be added when the athlete is ready and can be invaluable in preventing recurrent injuries that are common.

**Greater Trochanteric Bursitis**
Greater trochanteric bursitis generally occurs secondary to repeated irritation of the bursa, or less commonly from direct trauma. Traumatic inflammation of the bursa is seen in collision sports, such as football and hockey, and at times in soccer and baseball after sliding hard into a base or hitting the ground after diving for a ball. Repeated irritation usually occurs secondary to other biomechanical abnormalities such as a tight iliotibial band (ITB) and/or weakness of the hip abductors. The athlete reports an aching pain along the lateral aspect of the hip that is worse with running and jumping and any contact on the area, including laying on the affected side at night. Physical examination will demonstrate relative weakness of the hip abductors on the affected side, a tight ITB on Ober testing, and significant pain with direct palpation over the bursa.

Treatment consists of icing, NSAIDs, and stretching of the ITB and strengthening of the hip musculature with attention to the hip abductors. In refractory cases, injection of a corticosteroid into the bursa can be very helpful in progressing the rehabilitation process. Ultrasound is not indicated in acute cases but can be helpful in chronic recurrent cases to assist with stretching the proximal ITB.

**Knee**
Knee injuries are common in almost all sports, particularly those that require running, jumping and pivoting, and cutting. In addition, contact and collision sports often place the knee joint at risk for injury. It is important to understand the anatomy and biomechanics of the knee joint and how the injury can affect the various structures of the knee, including the patellofemoral joint, the two menisci, and the four primary ligaments (ACL, posterior cruciate ligament [PCL], lateral collateral ligament [LCL], and medial collateral ligament [MCL]).

**Patellofemoral Syndrome**
Anterior knee pain, also referred to as patellofemoral syndrome, is one of the most commonly encountered problems of the knee seen in a sports medicine practice. The incidence of this syndrome is reported to be as high as one in four in the general population (78). It also occurs with an incidence of 19% after ACL reconstruction (79).

Many theories exist regarding the etiology of the patellofemoral syndrome. Goodfellow et al. (80) described a failure of energy absorption by the articular cartilage causing increased patellar subchondral bone pressure. Ficat and Hungerford (81) placed emphasis on lateral malalignment causing hyperpressure of the lateral patellofemoral compartment and hypopressure on the medial patellofemoral joint. Laurin et al. (82) radiographically confirmed that early signs of chondromalacia appeared first on the medial patellar facet, with osteoarthritis more advanced on the lateral patellar facet. Abnormal lateral tracking of the patella remains the most common theory at the present time. A number of causative factors are linked to patellofemoral syndrome, with weakness/tightness of muscle groups, bony abnormalities, and malalignment of the lower extremity being the main ones (81).

Vastus medialis obliquus (VMO) insufficiency is often cited as one of the abnormalities resulting in patellofemoral syndrome. The function of the VMO is to help maintain proper patella tracking during extension of the knee. It is the only dynamic medial stabilizer and, if weak, it allows lateral tracking. The fibers of the VMO originate from the tendon of the adductor magnus and vastus medialis. This becomes an important point during treatment. Mariani and Caruso (83) found a decrease in VMO activity compared with vastus lateralis activity in those with patellar subluxation. Voight and Wieder (84) showed an increased speed of firing of vastus lateralis in those with PFS. The exact importance of VMO firing remains somewhat controversial.

The Q (or quadriceps)-angle is the angle formed by the intersection of lines drawn from the bisection of a line from anterior superior iliac spine (ASIS) to the superior patella and a second line from the inferior patella to the tibial tubercle. It is increased by increased femoral anteverision, external tibial torsion, increased pronation, lateral displacement of the tibial tubercle, and any lateral rotation of the patella. The upper limit of normal is greater than 12 degrees in men and greater than 18 degrees in women (85). Tightness of various soft-tissue structures about the knee can result in a functional increase in the Q-angle, which results in abnormal kinematics which affect patellar tracking, resulting in patellofemoral pain.

ITB tightness is felt to result in abnormal patellar tracking as its distal-most fibers insert on the lateral patella, exerting a lateral pull during knee flexion. Hamstring tightness also contributes to patellofemoral syndrome as it increases the patellofemoral joint reaction force in stance. Rectus femoris tightness increases anterior compressive force and has a net vector lateral pull on the patella. Gastrocnemius tightness causes a decrease in ankle dorsiflexion, with a resultant compensatory pronation of the foot, via the subtalar joint, which results in an increase in the Q-angle and consequently, a lateral patellar deviation. Hyperpronation results in internal rotation of the leg and femur, also increasing the Q-angle, and results in malalignment of the patella.

The treatment for patellofemoral syndrome first lies in a thorough analysis of contributing factors. Biomechanical abnormalities should be diagnosed and addressed. Acutely, pain control is addressed with ice, NSAIDs, and occasionally, electrical stimulation. Patients should avoid activities causing...
pain, which include kneeling, excessive stair climbing, and prolonged sitting. Patients should be given appropriate shoe orthotics if hyperpronation is an issue. Proper stretching of gastrocnemius, hamstrings, and ITB is essential in treatment. McConnell taping can be done to improve the positioning and tracking of the patella and to facilitate an aggressive stretching and strengthening program. Patients are generally taught how to tape themselves until their symptoms resolve (86). VMO strengthening via straight leg raising with adduction, EMG biofeedback, and squats with abduction pillow (so that squat is performed with adduction to strengthen VMO) are carefully reviewed with the athlete to ensure proper technique and VMO activation. Athletes are also instructed on patella mobilization techniques (particularly medial and inferior glides) to improve proper tracking. For athletes with significant pes planus or hyperpronation, an in-shoe orthotic is generally prescribed to help normalize the overall lower extremity biomechanics. The results of treatment for patellofemoral syndrome are generally quite good. Ingersol and Knight (87), using EMG biofeedback for VMO strengthening, demonstrated radiographically improved patellar tracking. Doucette and Goble (88), treating patients with a program of VMO strengthening and ITB and hamstring stretching, demonstrated good results in 84% of patients. There have been many other treatments that have demonstrated efficacy as well; therefore, the specifics of treatment remain controversial. The importance of good motor control and tracking of the patella does appear to be well accepted.

**Medial and Lateral Collateral Ligament Sprains**

The collateral ligaments are important for maintaining mediolateral stability of the knee joint. They are susceptible to injury from direct trauma, producing varus (LCL) and valgus (MCL) loads.

The MCL is the most commonly injured knee ligament in sports. It occurs from a valgus force to the knee joint that stretches or tears the ligament. It can also be injured in association with other structures of the knee, including the ACL, the medial meniscus, and the joint capsule. It is important to rule out injury of any of these associated structures, as the rehabilitation will be different. Isolated, complete tears of the MCL can be successfully managed nonoperatively at all levels of sport participation and should be considered the state-of-the-art treatment for this injury (89). Once again, the previously described program is applied. The knee should be protected by a double upright hinged knee brace, which initially can be used in conjunction with crutches until a normal gait pattern is established. This same brace can then be used as a functional knee brace for returning to sport. It is important to promote active ROM of the knee. This can be done on an exercise bike, which will promote full ROM without causing mediolateral stress to the MCL. Strengthening should also progress along with the protection, ROM phases moving as quickly as possible to closed kinetic chain strengthening, which are more functional. Once again, plyometric and power-strengthening issues should be addressed once there is no pain, full ROM, and symmetric strength. Serial examinations should be used to readdress the stability of the MCL on valgus stress testing. This should demonstrate decreasing pain and improved stability. Return-to-sport criteria require the athlete to be pain free and have full ROM, normal strength, no effusion, no pain, and minimal to no laxity on valgus stress testing. Once again, with a complete MCL tear, knee bracing would be recommended for 6 to 9 months (usually for the remainder of the sport season).

Although isolated LCL sprains are less common, they too can be managed nonoperatively; however, great care must be taken to rule out injury to other structures, such as the posterior lateral corner, PCL, and ACL. If there is any question regarding the integrity of these structures, additional imaging and referral to an orthopedic surgeon specializing in sports medicine are recommended.

**Anterior and Posterior Cruciate Ligament Injury**

The cruciate ligaments cross within the knee joint, providing predominantly anterior and posterior stability, along with lateral rotatory stability to the knee joint. ACL injuries most commonly occur following a hyperextension injury or from a significant valgus force of blow to the knee. O’Donohue described the “unhappy triad” following a valgus injury to the MCL, ACL, and medial meniscus, although lateral meniscal tears are also quite common (90,91). In general, ACL injuries do not heal, although there are rare reports of spontaneous healing, especially injuries at the attachment to the femur (92,93).

In ACL injuries, patients often describe hearing or feeling a “pop” with an unstable sensation of the knee. If chronic, giveaway episodes of the knee are common with rotatory activity, such as pivoting. An effusion (hemarthrosis) usually develops rapidly and is generally quite obvious on physical examination as a result of the significant vascularity of the ACL. On examination, a positive anterior drawer and Lachman test should be noted. Other structures about the knee should also be carefully examined. If the findings on examination are unclear, then additional imaging (usually MRI) of the knee is indicated.

PCL injuries generally occur with a posterior force to a bent knee, such as after being tackled or receiving a blow with the knee in 90 degrees of flexion as can occur with a dashboard injury during a motor vehicle accident. There usually is little swelling, and often the athlete may attempt to play. On examination, there is a positive posterior drawer sign, which is the most sensitive examination test for a PCL injury (94). Once again, care must be taken to avoid missing injuries to the lateral structures of the knee.

Nonoperative treatment with a stepwise approach as previously described can be tried for both ACL and PCL injuries; however, for a young athlete who wishes to continue with high-demand sports (i.e., jumping/twisting/pivoting-type sports), then ACL reconstruction is recommended. For those who attempt a nonoperative treatment course, the results can be favorable, particularly if there is a strong focus on the later stages of strengthening and proprioceptive training, along
with some type of functional bracing. A 15-year follow-up study by Ageberg and colleagues revealed that good functional performance and knee muscle strength can be achieved and maintained over time in patients with ACL tears undergoing aggressive and early functional rehabilitation and not reconstructive surgery (95). Functional bracing is actually considered both for surgically repaired and nonsurgically repaired ACLs. A 15- to 20-degree extension stop on these braces should be strongly considered, as the authors believe that this accentuates the proprioceptive aspect of the brace.

Isolated PCL injuries have been found to do quite well. A functional stepwise approach works very well with a bias toward greater quadriceps strength and closed kinetic training. We have not found bracing to be necessary or helpful.

**ACL Injury Prevention**

ACL injury is common in sports as has been previously mentioned. As female participation in athletics has increased since the inception of Title IX, there has been a dramatic rise in the rate of these injuries (96). Furthermore, we are learning that the long-term morbidity is not favorable. A literature review on the topic by Lohmander and colleagues suggests that 10 to 20 years after an ACL tear, osteoarthritis will develop in the knee, regardless of whether the ACL was surgically repaired or not (97). Therefore, prevention of ACL injuries is extremely important not only for the immediate consequences of an ACL tear on the athlete, but also because of the long-term structural damage that can occur in the knee at an early age for most of these athletes as they progress into their 30s and 40s.

Studies have shed light and understanding into the biomechanics of ACL injuries in young female athletes. These studies have analyzed the mechanics that these athletes use when the knee is placed in positions known to cause ACL tears, such as landing from a jump, sharp cutting, or decelerating. These positions often place the knee in flexion and internal tibiofemoral rotation. Hewett and his colleagues have described their theory that dynamic knee instability is caused by three factors: ligament dominance, quadriceps dominance, and leg dominance (96).

Ligament dominance refers to the fact that females tend to depend on their ACL for joint control, rather than on their muscles and on their neuromuscular firing and recruitment, when compared to their male counterparts. Quadriceps dominance refers to young female athletes having significantly weaker hamstrings than quadriceps. Building on the known observation that at angles greater than 45 degrees, the quads act as antagonists to the ACL, deep knee flexion and strengthening of the knee flexors (i.e., the hamstrings) have been proposed as a protective mechanism for the joint (96).

These researchers have therefore devised a dynamic neuromuscular analysis and training program for the prevention of ACL injuries, in particular in the female athletic population. It is a program that incorporates biomechanics, plyometrics, core stability, and strength training of the hamstrings to develop adequate neuromuscular adaptation for the knee and to enhance kinetic chain control and balance. The goal of this program is to analyze and correct the biomechanical that are made by these young female athletes when the knee is in ACL-stressing positions. By correcting such errors early in the life of these athletes, they will be protected from injury later in their athletic endeavors (96,98). The incorporation of such programs into communities could have a tremendous impact on reduction of ACL injury rates in the female athletic population.

**Meniscal Tears**

The menisci serve as shock absorbers and force dissipaters to protect the knee joint. They can be injured in isolation or in conjunction with other structures, such as the collateral ligaments, ACL, etc. These injuries can be from either a direct blow to the knee or a twisting type injury, as the menisci are susceptible to compression/rotation forces. Degenerative meniscal tears in the older athlete may become symptomatic following a prolonged period of knee flexion.

Patients will generally report swelling or “tightness” within the knee from the associated synovitis that occurs, and bleeding in the younger athlete if there is injury to the vascular portion of the meniscus. There can be mechanical symptoms such as catching or locking with an associated or intermittent clicking. Symptoms are generally increased with knee flexion and are often localized to the joint line.

On examination, there is usually a small to moderate effusion, with pain on flexion and on palpation of the joint line over the side of the injured meniscus. As noted above, McMurray test, while very specific, has a low sensitivity and therefore should not be used as the determinant of meniscal injury in isolation (99).

Many meniscal tears can be treated conservatively with a combination of relative rest, vigorous icing, and NSAIDs initially, followed by a course of physical therapy with restoration of motion and an aggressive strengthening program stressing closed kinetic-type exercises. If the effusion persists in spite of compression and icing, aspiration and injection with a corticosteroid can be very helpful in decreasing the reactive inflammatory response, which inhibits the normal muscle firing of the knee. In those patients with a locked knee or severe mechanical symptoms and those unresponsive to a 2- to 3-month course of physical therapy, arthroscopic treatment with either resection of the unstable portion or repair (if possible) has been very effective.

**Ankle**

**Achilles Tendinitis**

Achilles tendinitis has been described in any sport requiring running or rapid changes in direction, such as in racquet sports such as tennis or badminton (100–102). It typically begins as an inflammatory reaction around the Achilles tendon/paratenon. The paratenon surrounds the tendon but is not a true synovial sheath. If not treated, microtears, mucoid degeneration, longitudinal fissuring, and scarring develop. Running is the most commonly associated activity, associated
with 53% of those cases studied by Kannus et al. (103). The etiology is multifactorial, with overuse being the most common cause (104). It is also associated with a rapid increase in mileage, increased interval training, and running downhill or sloping hills. Numerous biomechanical problems are correlated with Achilles tendinitis, and a thorough biomechanical analysis is necessary in resistant cases. Treatment consists of decreasing inflammation, followed by aggressive stretching of the gastrocnemius/soleus complex. Strengthening should be initiated in concert with stretching to improve the ability of the tendon to withstand repetitive changes in length and load. Eccentric strengthening is the most important aspect of strengthening and is based on three parameters: length, load, and speed (105). Stretching helps to increase the length of the muscle tendon unit and reduce strain with joint movement (106). Increasing load to myotendinous unit helps to increase its tensile strength. Finally, increasing speed of contraction helps to increase force of contraction.

Ankle Sprains
Inversion ankle sprains are the most common traumatic injuries encountered in sports that require running or jumping, such as football, basketball, tennis, soccer, and gymnastics. Deltoid ligament and syndesmotic ankle sprains occur much less commonly and are usually associated with more significant trauma to the ankle. They have been well described in football and soccer (107–109).

As noted above, the most common injuries to the ankle involve the lateral ligament complex, consisting of the anterior talofibular, calcaneofibular, and posterior talofibular ligaments. A grade I ankle sprain involves a mild sprain of the anterior talofibular with a negative ankle drawer and talar tilt test. A grade II sprain involves disruption of the anterior talofibular with sprain of the calcaneofibular, with a positive ankle drawer and a negative talar tilt test, whereas a grade III ankle sprain involves disruption of the lateral ligament complex with both positive ankle drawer and talar tilt tests. The careful clinician must completely evaluate the individual with ankle sprain to rule out concomitant fracture of the fibula or fifth metatarsal, which may occur at the time of inversion injury. The Ottawa ankle rules have been used to determine when radiologic assessment is necessary in the face of inversion ankle injury. Some positive indicators for testing include difficulty bearing weight and tenderness about the medial or lateral malleolus or fifth metatarsal.

Deltoid ligament injuries constitute only 5% of ankle injuries and result from eversion injury to the ankle (110). These injuries may occur concomitantly with inversion ankle sprains; therefore, the medial aspect of the ankle must be inspected with any significant inversion sprain. The majority of isolated injuries to the deltoid ligament are mild, with more significant injuries often associated with injury to the anterior and posterior tibiofibular ligaments, along with diastasis of the syndesmosis. Syndesmotic ankle sprains are more common than significant deltoid ligament sprains, which may relate to the overall tensile strength of the deltoid ligament complex.

They may be missed if a comprehensive evaluation is not performed, including palpation in or about the region of the anterior tibiofibular ligament or manual loading of the ankle mortise. Evaluation of the individual with deltoid ligament or syndesmotic sprain must also include a careful assessment of the fibula to rule out distal or proximal (Maisonneuve’s) fracture. Syndesmotic ankle sprains can be a significant source of disability in the competitive athlete with a significantly prolonged recovery time as compared with isolated inversion ankle sprains (109).

Treatment of ankle sprains involves the principles of pain control through the use of ice and anti-inflammatory medication, relative rest, and early mobilization. In concert with early mobilization, the individual with an ankle sprain must be instructed in balance and proprioceptive exercises with progression toward strengthening of the musculature of the foot and ankle. Return to sport may depend on satisfactory completion of a rehabilitation program, compliance with a home exercise program, and the use of taping or bracing. More significant ankle inversion (grade III) or deltoid ligament sprains may require surgical intervention if chronic instability results in residual functional deficits.

Foot
Metatarsophalangeal Sprain (“Turf Toe”)
Injury to the metatarsophalangeal joint generally occurs with hyperextension injuries of the great toe in sports such as football and soccer. It has been termed “turf toe” because of the unforgiving and stiff surface of synthetic surfaces. In addition to hyperextension, there are often added forces, such as tacklers or other players falling on the injured athlete. This is associated with localized pain at the metatarsophalangeal joint and pain with weight bearing, but especially during push off in running. The examination can demonstrate some localized tenderness with some localized swelling and decreased or painful ROM. In recurrent injuries, loss of motion and degenerative changes about the joint may occur. The treatment consists of the PRICE regimen, with taping techniques to limit motion at the joint. A long rigid shoe orthotic may assist in decreasing force across the joint to facilitate return to play. This can be a very disabling injury, requiring several weeks of rest before return can be accomplished. Occasionally, corticosteroid injection is necessary to control pain and inflammation of the joint and to facilitate a progressive rehabilitation program.

Plantar Fasciitis
The plantar fascia runs across the medial aspect of the foot from the medial calcaneus to the metatarsals to maintain the medial arch of the foot. A sudden loading of the feet can injure it, but more often it is inflamed and irritated insidiously by repeated overload. It can occur in both a pes planus foot and a pes cavus foot. The exact cause remains obscure, but it does appear that both muscular weakness and soft-tissue tightness play an important role. The symptoms are generally that of insidious onset of heel and plantar foot pain that is worse when
first rising in the morning or after a period of inactivity such as sitting. It can also increase after the end of a competition or practice or at the end of the day.

On examination, there is usually focal tenderness at the origin of the plantar fascia at the medial calcaneus and along the plantar arch. Pain can be elicited by hyperdorsiflexion of the great toe with palpation along the plantar fascia. There is usually associated tightness of the gastrocsoleus complex and weakness of the soleus and possibly the tibialis posterior and tibialis anterior as well. Plain x-rays are often obtained and may reveal a tractionsal calcaneal spur that should not be interpreted as the cause of the symptoms but rather as an additional sign of overload of the plantar fascia. In addition, these spurs are often noted on the asymptomatic foot. Therefore, plain x-rays are not necessary for most patients who present with plantar fascia pain.

Treatment includes aggressive stretching and at times a night splint to produce a passive prolonged stretch of the tissues. Strengthening of the gastrocsoleus, tibial, and foot intrinsic muscles helps to restore dynamic stability of the arch and foot and ankle joints. Patients with significant biomechanical abnormalities of the foot may benefit from a custom orthotic. The role and timing of corticosteroid injection are not clear. There is anecdotal evidence that they can be very helpful, but at what point and how often they should be injected is not clear (111–114). Surgical release of the plantar fascia is rarely necessary but can be helpful in recalcitrant cases (115,116).

**Midfoot Sprain**

It is important for the sports medicine physician to be aware of midfoot sprains as they can be easily overlooked. The injury occurs with an awkward landing of the forefoot, usually in inversion. This can appear as a simple ankle sprain, and therefore the injury can be initially missed. On careful examination, the tenderness to palpation is localized to the dorsal-medial midfoot. Occasionally, there is swelling and ecchymosis in the area. With a varus stress to the foot, pain can be elicited as well. Most of these athletes are unable to even walk without an antalgic gait pattern. It is important to get plain x-rays of the foot and note any widening of the first and second metatarsal rays. Widening of greater than 5 mm is unstable and requires surgical fixation. If there is no significant widening on x-rays, 4 to 6 weeks of immobilization in plantar flexion and supination will allow for healing. The athlete should then progress through a functional rehabilitation program. As the injury heals, it is important to maintain aerobic conditioning of the athlete. For mild sprains, icing, wrapping, and crutch ambulation are used for maintaining strength and ROM as the tissue heals.

**CONCLUSION**

The physiatrist is the ideal specialist to provide and coordinate care for injured athletes at all levels of competition. The team approach provides the best method of maximizing an athlete’s recovery from an injury. This chapter presents an overview on the basic principles of functional rehabilitation and some common injuries encountered when treating athletes.

**REFERENCES**


Several studies have reported that people with disabilities are more likely to be sedentary (1,2), have health problems (3), and experience substantially more barriers to physical activity participation compared to the general population (4–6). This increases the likelihood that, as they age, they will have greater difficulty maintaining their ability to work, participating in recreational activities, performing self-care activities (7,8), and engaging in various activities in their community (9). The subsequent physical limitations that many people with disabilities experience may not be entirely associated with the primary effects of the disability, and some or much of the limitations may be related to the effects of disuse and deconditioning (10,11).

Improving health and function in people with disabilities has the potential to facilitate greater levels of participation in all aspects of society, including work, leisure, and recreation; reduce or mitigate secondary conditions (12); and achieve a higher satisfaction with life (13). An important first step in this process is for clinicians and health professionals to find effective ways to identify and remove the many barriers to participation that people with disabilities encounter when attempting to become physically active or increase their current levels of physical activity.

The focus of this chapter is to guide health care professionals in their efforts to increase physical activity in patients or clients with disabilities. The first section discusses how community-based physical activity programs can be used as a mechanism for transitioning patients/clients from rehabilitation to self-maintenance of health and function across the life span. Within this section is a discussion of an approach to promote physical activity in people with disabilities, which includes tailoring physical activity for a patient/client and a model for systematically prescribing physical activity for individuals with physical disabilities. The second section focuses on specific types of recreation, fitness, and sports activities that individuals with disabilities can engage in and identifies appropriate community-based resources.

**TRANSITIONING FROM REHABILITATION TO COMMUNITY-BASED PHYSICAL ACTIVITY**

While most individuals who are recovering from a serious injury or accident will receive a certain amount of rehabilitation, the gains in health and function are often lost after the person returns home if he/she does not continue some type of home- or community-based physical activity program (14,15). Most rehabilitation is limited to the subacute period and is designed to restore or improve the most critical functional skills needed to assist the person in performing BADL (basic activities of daily living) and/or IADL (instrumental activities of daily living) (15). Once the person returns home, much of the responsibility for continuing to perform the recommended rehabilitative exercises is left to the individual and/or his or her caregiver. Figure 54-1 provides a conceptual illustration of what often occurs after an injury, accident, or onset of a new health condition that leads to a progressive decline in health and function. Typically, an individual receives a few to several days of rehabilitation in a hospital setting or outpatient facility, though this varies according to the injury. Individuals with more severe disabilities (i.e., spinal cord injury) receive longer hospital stays. Significant progress is often made as a result of intensive rehabilitation and the natural phase of recovery. But after the person returns home to what is often a considerable alteration in health, function, and lifestyle, there is, theoretically, a gradual decline in health and function resulting from physical inactivity. A primary goal of every rehabilitation program should be to increase physical activity postrehab to avoid a state of severe deconditioning that could impose serious limitations in performing BADL and IADL. In some ways, the key to a successful rehabilitation program is identifying effective strategies that can support individuals in participating in some form of regular physical activity after they are no longer receiving therapy. People with disabilities who reported a higher level of physical activity also indicated a higher level of community reintegration compared to participants who described their physical activity as low or inactive (16,17).

**Tailoring Physical Activity to Meet the Unique Needs of the Patient/Client**

Tailoring physical activity programs that recognize the unique circumstances surrounding each individual is an important aspect of prescribing physical activity programs that have a greater likelihood of successful initiation and adherence (18). Research on physical activity indicates that generic programs not targeted to the needs of any particular end user are far less likely to result in long-term maintenance of health-promoting behaviors (19,20). By assessing a combination of factors,
including the person’s motivational level (readiness to change), physical activity profile, health and mobility limitations, and barriers to participation, a program can be developed that meets each person’s specific needs, interests, and circumstances. Thus, physical activity recommendations for people with disabilities may be more effective if they are individually and culturally appealing and are implemented in a setting of the individual’s own choosing. This includes establishing realistic, achievable goals to meet the person’s needs, while also finding solutions for the barriers to participation that frustrate even the most enthusiastic participant. The program also has to be dynamic (i.e., interesting, enjoyable) and have the opportunity to change frequently to accommodate changes in the life of the participant (i.e., boredom, getting a new job, experiencing some pain performing certain types of exercises, etc.). So many programs fail because the individual or his or her environment changes and the program is no longer interesting or challenging, or the environment changes and the person no longer has access to the same resources needed to participate in the program or activity.

**PEP Intervention Model**

An example of an online resource that can assist rehabilitation professionals in providing more tailored physical activity recommendations to their patients/clients as they transition from in-patient rehabilitation to community-based activities is the Personalized Exercise Program (PEP). The PEP intervention model (shown in Fig. 54-2) begins with problem identification accomplished through a detailed assessment of the individual’s needs, interests, activity level, health status, functional ability, and readiness to change. This assessment of individual health, lifestyle behaviors, and function, in conjunction with an evaluation of the individual’s family support, community resources, and environmental barriers, allows a personalized physical activity program to be tailored for the individual. Jointly, individual-, family-, and community-level strengths and resources are harnessed to help the individual overcome barriers to physical activity.

While implementation of the resulting PEP intervention in the context of long-term support is expected to result in increased physical activity and improved health outcomes, more research needs to be conducted to determine the effectiveness of this model (21,22). Establishment and maintenance of social support networks are an important component of maintaining physical activity behaviors in people with disabilities. The individualized assessment, in conjunction with an evaluation of community resources (e.g., community swim programs, accessible fitness centers), allows for a personalized physical activity program that is based on customized information. Both individual and community-level strengths and resources are utilized to help the patient/client overcome barriers to participating in a sport or other type of physical activity at home or in the community. Establishment and maintenance of social support networks that include building community friendships with other people, in addition to parental and/or caregiver support, are important components for maintaining participation over an extended period until health benefits are achieved (21,22). These networks may include personal contacts with family members or caregivers or may involve friendships formed through exercise classes or nutritional cooking classes offered as part of a community-based program.

The dynamic design of the PEP intervention model and its person-centered approach allow the rehabilitation professional to revise and modify the program at any time until the recommendations are calibrated or recalibrated (as in the case with an illness or new secondary health condition) to the user’s needs and interest level. In Figure 54-2, the first step of the PEP intervention model is illustrated in the left column, comprehensive needs assessment, which includes the following components: (a) physical activity profile and activity preferences, (b) barriers to participation, (c) health status and mobility limitations, and (d) motivational level (readiness to change). Each of these components is described below.

**Components of the PEP Intervention Model**

**Physical Activity Profile**

Understanding the patients/clients’ physical activity participation history in sports, recreation, and/or fitness before (when possible) and after the onset of a disability is critical to
designing an effective program. An assessment of a patient/client’s specific disability and functional level, along with personal activity preferences, can help determine potential activities. For example, if a person with a spinal cord injury was an enthusiastic softball or basketball player before his or her disability, finding links to participating in wheelchair softball or basketball, whenever possible, may be more appealing and increase the individual’s participation. As discussed in the second section, any recreation, fitness, and sport activity can be adapted for a person with a disability that should motivate rehabilitation professionals to learn more about how this can be accomplished for certain clientele.

Barriers to Physical Activity

Personal and environmental barriers can impose substantial limitations in maintaining a physically active lifestyle. Among people with disabilities, some of the more common barriers are pain, lack of transportation, insufficient financial resources to pay for a health club membership, lack of awareness of community-based programs, and not knowing how to perform certain types of exercises or recreational activities. Usually, more specification is required in developing exercise programs for disabled populations because certain impairments and activity limitations can limit an individual’s access to a program or activity. As certain barriers are identified and removed, participation in various types of physical activities will likely increase.

One of the major concerns associated with increasing physical activity participation among people with disabilities is the lack of access to many community sports, recreation, and fitness programs (23). People with disabilities encounter enormous barriers in the built and natural environment (6). Indoor and outdoor structures have a major effect on participation in physical activity among people with disabilities (24,25). Structures such as gyms, fitness centers, outdoor trails, parks, and swimming pools often have poor signage, lack detail on how to use the equipment or participate in a program, or provide poor access routes to and from the facility or program.

Indoor environments of many fitness and recreation facilities also need to become more accessible for people with disabilities. One of the major barriers is inaccessible exercise equipment (26). Most manufacturers do not consider in their design specifications how to make their equipment accessible for people with physical, cognitive, and sensory disabilities. Typically, commercial cardiovascular exercise equipment requires propulsion by using the musculature of both lower extremities (i.e., treadmills, stationary bikes,
elliptical cross-trainers, and steppers), thereby restricting use among people with lower-extremity disabilities (e.g., paralysis, limb loss). While a few fitness facilities may be able to purchase a commercial quality arm cycle or wheelchair ergometer, the vast majority of fitness centers either cannot afford this equipment or do not find it cost-effective to purchase one for a small percentage of their clientele. Another problem with inaccessible equipment is that offering clients with disabilities the opportunity to use one piece of adaptive exercise equipment while the rest of the membership has access to all of the equipment clearly limits the amount of enjoyment and benefit that can be obtained from a more diversified program.

People with visual or cognitive disabilities also have difficulty using various types of exercise equipment (27). Display panels are often difficult to read or understand, getting on and off the equipment presents some risk of falling, and machines are often hard to propel or lift (e.g., weights are too heavy to lift) for people with low strength levels.

Programmatic issues are also barriers for people with disabilities (28). Fitness and recreation classes such as aerobic dance, yoga, and tai chi are often taught by instructors who have minimal knowledge of how to adapt their program for someone with a disability. Spoken directions regarding the types of exercises or movements the class is being asked to perform tend to be terse or vague. Rehabilitation specialists must have a good understanding of how to overcome certain barriers in order to ensure that the patient/client can successfully participate in various types of community-based programs.

Health Status and Mobility Limitations
A rehabilitation professional must also understand the health and mobility limitations that prevent a patient/client from engaging in certain types of physical activities. Various types of disabilities include impairments such as loss of balance, vision, hearing, pain, fatigue, decreased cognition, paralysis, and many others. Each of these impairments can have an effect on the physical activity recommendations. Similarly, mobility limitations such as difficulty walking, climbing steps, transferring from a wheelchair, etc., must all be considered when designing a physical activity program that will match the individual’s functional level.

Motivational Level
Finally, for successful participation in any form of physical activity, the rehabilitation professional must find innovative ways to encourage participants to obtain regular physical activity. People with disabilities are often demotivated because of a number of barriers preventing them from becoming physically active (6). The stages of change model (29,30) includes five stages: precontemplation—no intention to change a behavior in the foreseeable future; contemplation—awareness that he needs more physical activity but has not yet made a commitment to take action; preparation—intending to take action in the next month but has unsuccessfully taken action in the past year; action—interested in modifying behavior to increase physical activity; and maintenance—individual is engaging in physical activity. The stage of change that a patient/client is experiencing is often dependent on personal and environmental barriers. Therefore, assistance in identifying and addressing these barriers, helping the individual to appreciate the benefits of regular physical activity, as well as greater access to various forms of physical activity can help improve physical activity participation.

Community-Based Physical Activity Options
The need to continue the recovery process after an injury or disability must include access to various types of physical activities in the person’s community (31,32). In order to avoid boredom or lack of interest in a specific physical activity program, it is optimal to offer as many activity options as possible. Certain individuals may prefer group exercise (i.e., an aqua-aerobics class) in a community-based setting, while others may choose to exercise alone in their home or engage in some form of outdoor activity. Regardless of preference, a variety of options should be available. In the second section, such community-based physical activity options and resources are reviewed.

Types of Physical Activities
The pyramid in Figure 54-3 illustrates different types of physical activities that can be engaged in over the course of the day. With rising obesity rates and low levels of energy-requiring activities at work and home, it is critical that all forms of physical activity be increased over a 24-hour period. Even simple movements (i.e., moving between commercials) can result in substantial accumulations of energy expenditure over the course of the day.

At the base of the pyramid is residual activity. This type of activity is generally performed at a very low intensity threshold and involves frequent unstructured movements, such as getting up and down from a chair, walking into different areas of the home, fidgeting, moving arms, etc. While residual activity does not offer any formal mechanism for structuring physical activity, the more an individual moves during the day, the higher the energy expenditure. Any activity above sitting, lying down, or sleeping requires greater energy expenditure and therefore higher calorie output.

![Physical activity pyramid](image-url)
The next level of the pyramid is task-oriented physical activity. These types of activities relate to certain indoor and outdoor household tasks such as cleaning, doing laundry, shopping, driving a car, doing yardwork, etc. Several of these activities can be strenuous in nature and provide effective ways to increase energy expenditure in individuals who generally have low levels of physical activity or are somewhat deconditioned and/or uninterested in engaging in structured activities.

The third level of the pyramid is leisure physical activity, which includes structured and planned physical activities such as leisure walking, bike riding, hand cycling, outdoor yard games, bowling, team and individual sports, fishing, etc. This form of physical activity can range from very low to moderate intensity activity. At the top of the pyramid is fitness-related activity, which is the most strenuous type of activity, with the highest rate of return in terms of improving health and function.

Rehabilitation professionals can assist their patients/clients to incorporate each level of physical activity within their daily regimen so that daily energy expenditure is increased. For example, residual activity can be performed at regular intervals during the day (i.e., get up or push wheelchair to another room between commercials, spend 5 minutes at the end of every hour moving arms or legs); to increase task-oriented physical activity, more chores can be completed inside and outside the home (i.e., dusting, cleaning windows, mopping or sweeping floor, cleaning appliances, grocery shopping, etc.); leisure physical activity can include new hobbies, such as tennis or bowling, leisurely walks or rolls for a few minutes several times a day, joining a chair exercise class, etc. Fitness-related activities can generally be performed for less than 1 hour, compared to the other forms of physical activity that are demonstrated at lower intensity thresholds. The next section provides many different recommendations for increasing two types of physical activities, leisure/sports, and fitness-related activity.

**RESOURCES IN PHYSICAL ACTIVITY AND DISABILITY FOR REHABILITATION PROFESSIONALS**

One of the most important elements of this chapter is to provide rehabilitation and health professionals with a one-stop resource for identifying physical activity materials that can be used to facilitate/promote/prescribe participation among people with various types of physical, cognitive, and sensory disabilities. While textbooks provide a framework for guiding professionals in planning and developing rehabilitation programs, when it comes to physical activity, the Internet is extremely important for identifying key resources that can be passed on to the patient/client including specific information related to his or her disability and engaging in various types of physical activities, resources on how to exercise or use an exercise facility, and, most importantly, video clips demonstrating various types of exercises or physical activities that are easier to visualize than explain. The National Center on Physical and Disability has been funded since 1999 to promote physical activity among people with disabilities. This resource center will allow clinicians to identify materials, programs, etc., for their patients/clients.

**The National Center on Physical Activity and Disability**

As has been discussed in the first section, participation in fitness, recreation, and sport is more likely to succeed if programming is tailored to the participant’s abilities and interest levels, while simultaneously addressing the barriers that these individuals may encounter for participating in such activities. The National Center on Physical Activity and Disability (NCPAD, www.ncpad.org) addresses these problems by making sports and physical activity programs more accessible to people with disabilities through advanced Internet technologies that can assist rehabilitation professionals to develop appropriate, accessible physical activity programming (Fig. 54-4).

NCPAD serves as a central repository of information on physical activity and disability, actively collecting information from research, best professional practices, information on public and private recreation and fitness facilities serving people with disabilities, and businesses that provide equipment and services supporting physical activity participation by people with disabilities. In addition, NCPAD has actively promoted the importance of physical activity in attaining and maintaining optimal health for people with disabilities. This is being accomplished through a variety of promotional resources and outreach activities in partnership with advocacy organizations, service providers, and individual consumers.

Rehabilitation professionals can use the NCPAD web-based resources when designing treatment programs for patients/clients that include acute care guidelines, in addition to web-based physical activity assessment tools, and information on community-based resources and activity programs that promote long-term physical activity maintenance.

NCPAD’s information is centralized on its Web site (www.ncpad.org), which provides a range of resources on physical activity and disability: information on physical activity and disability, networking opportunities, searchable databases, assessment tools, and research. NCPAD Information Center can be reached at 1-800-900-8086.

**FIGURE 54-4.** National Center on Physical Activity and Disability.
Specialists are available at 800-900-8086 or ncpad@uic.edu to answer questions, including but not limited to appropriate exercise for individuals with a specific disability, available adaptive equipment, the location of accessible fitness programs and sport team opportunities, and more.

**Information on Physical Activity and Disability**

Fact sheets include physical activity (i.e., lifetime sports, competitive sports, exercise, fitness, and leisure), disability/health conditions, nutrition, and wellness. Fact sheets on specific disabilities, for example, encompass an overview of the disability; aerobic, strength, and flexibility components for developing an appropriate fitness program; precautions and safety guidelines for exercising with a specific disability; and key organization contacts and references by local community. Sports and recreation fact sheets provide general information about the benefits of the activity, rules and regulations, and key contact organizations. Many fact sheets have embedded video clips so that the specific sport or exercise technique can be demonstrated to the user (http://www.ncpad.org/videos/), and some are available as videos/DVDs/quick series booklets that can be purchased at NCPAD’s web shop (http://www.ncpad.org/shop/). NCPAD’s free monthly newsletter (http://www.ncpad.org/newsletter/newsletter.php?letter=current) also provides updates on new NCPAD resources, fitness trends for people with disabilities, secondary condition prevention, as well as information on events and conferences, grants, and employment information.

**Searchable Databases**

NCPAD’s online searchable databases can assist professionals in locating appropriate physical activity programs and organizations, adaptive equipment and devices, and qualified personal trainers. These resources can help rehabilitation professionals involve their patients/clients in community-based physical activities and sports and/or obtain physical activity support and training from fitness professionals.

A web-based directory of accessible physical activity programs (including sports, recreational, and exercise/fitness activities) for people with disabilities and chronic health conditions (http://www.ncpad.org/programs/) and organizations providing health promotion and community-based resources for people with disabilities (http://www.ncpad.org/organizations/) may be searched by geographic area or topic. A suppliers database (http://www.ncpad.org/suppliers/) contains information on equipment and assistive devices for sports and recreation for people with disabilities, and a personal trainers database (http://www.ncpad.org/trainers/index.php), also searchable by geographic area, supplies information about personal trainers who have experience working with people with disabilities and chronic health conditions.

**Networking Opportunities**

Web-based forums in the areas of networking (exercise professionals, exercise buddies), adaptive equipment, community resources (services, funding, programs), research (grants, articles), adherence ideas/suggestions, and best practices on accessibility (http://www.ncpad.org/) can help rehabilitation professionals exchange information, above and beyond that available through the NCPAD Web site and the Information Specialists. The “Your Writes” section (http://www.ncpad.org/yourwrites/) also provides articles and information written by people with disabilities, family members, and caregivers on health promotion opportunities and resources for people with disabilities.

**Assessment Tools and Research**

NCPAD and its related projects include assessment instruments and resources on evaluating physical activity levels and accessibility for people with disabilities to participate in fitness and recreation programs and facilities. Additionally, resources are available to assist professionals who are developing evidence-based programs, applying for grants, or designing a research study.

Evaluation tools include the PADS (Physical Activity Disability Survey) and B-PADS (Barriers to Physical Activity and Disability Survey) instruments. The PADS is designed to assess low-level physical activity among persons with physical disabilities and chronic health conditions (http://www.ncpad.org/) and the B-PADS, barriers to physical activity encountered by persons with disabilities, specifically personal (e.g., lack of motivation, fear of leaving home, perception of exercise as difficult or boring) and environmental barriers (e.g., lack of transportation, costs, and availability of fitness and recreation facilities). Assessment tools can also help to regularly measure steps logged per day, as well as body mass index and weight.

NCPAD’s AIMFREE (Accessibility Instruments Measuring Fitness and Recreation Environments) manuals are a validated series of questionnaire measures that can be used by persons with mobility limitations and professionals (i.e., fitness and recreation center staff, rehabilitation professionals, owners/managers of fitness centers) to assess the accessibility of recreation and fitness facilities, including fitness centers and swimming pools. The instruments are available for purchase at NCPAD’s web shop (http://www.ncpad.org/shop/).

For rehabilitation professionals interested in research, the Web site references section provides listings of books, journals, newsletters, videos, pamphlets, reports, theses, and proceedings (http://www.ncpad.org/refs/books/). Abstracts of current research articles on physical activity and disability are available at http://www.ncpad.org/research/ and presentations on physical activity and disability at http://www.ncpad.org/presentations/. The monthly NCPAD newsletter also contains up-to-date information on grants and events/conferences (http://www.ncpad.org/newsletter/newsletter.php?letter=current) and includes information on new or interesting recreation, leisure, and sports programs offered to people with disabilities throughout the country.

**Fitness Programming for People with Disabilities**

Physical fitness is an important area of emphasis in improving health among people with disabilities. The three primary components of fitness include cardiorespiratory endurance, muscular strength and endurance, and flexibility.

Cardiorespiratory endurance or aerobic capacity refers to the ability of the heart (cardio) and lungs (respiratory) to provide sufficient blood flow to working muscles for sustained
physical activity. Muscular endurance is defined as the muscle’s ability to exert force for a sustained period of time, and muscular strength refers to the ability to generate force one time. Both contribute to improved balance, mobility, and stability, as well as improvements in BADL and IADL. Flexibility is the movement capability (i.e., range of motion) of muscle groups around a joint and can help to reduce injury, improve posture, and perform ADL and IADL (33,34).

Below are a series of guidelines for developing fitness center programming for people with disabilities with respect to cardiorespiratory endurance, strength, and flexibility (Table 54-1). The NCPAD (www.ncpad.org) resources can be used for finding a fitness center in a specific geographic area, as well as locating adapted equipment. Note that these considerations do not generally apply to leisure and competitive sports and activities, which are less likely to be closely monitored for safety and specific conditioning effects.

Before beginning an exercise program, a patient/client’s physician should be informed so that adequate precautions can be taken, and possible side effects of medications considered (Table 54-2). The American College of Sports Medicine (ACSM) (35) recommends that individuals with certain risk factors for exercise have a graded exercise test to determine how the heart responds to stress and whether there is adequate blood flow to the heart during increasing levels of activity. See the NCPAD factsheet on “Exercise Guidelines for People with Disabilities” (http://www.ncpad.org/exercise/fact_sheet.php?sheet=15&view=all) for more information.

In order to locate an exercise professional who has adequate education and experience, seek out recommendations from health professionals or national disability organizations. Though there are no national standards or educational requirements necessary to be an exercise professional, the ACSM, the American Council on Exercise (ACE), and the National Strength and Conditioning Association (NSCA) all issue certifications recognized nationally by exercise professionals. Moreover, some organizations assist with finding a certified professional in an area, including NCPAD’s personal trainer database (http://www.ncpad.org/trainers/index.php) and the ACSM ProFinder.

### TABLE 54.2 Parameters for Choosing an Exercise Professional and Fitness Center

- Education and experience
- Personal preferences
- Training frequency
- Location
- Cost
- Social skills

Other important considerations include the personality and gender of the professional, training frequency, the availability of the professional, whether training is desired at home or in a gym, and cost. For more tips, see the NCPAD factsheet “How to Choose a Personal Trainer” (http://www.ncpad.org/fitt/fact_sheet.php?sheet=384).

In choosing a fitness center, similar considerations apply. Depending on the patient/client’s disability and functional level, accessibility of both the center and the equipment is critical. Exercise equipment using universal equipment design is preferable, as it is accessible for persons with and without disabilities. Moreover, other factors such as availability of staff on site, location, transportation options, and cost should also be considered. See transformation in a “Before-After Fitness Center Makeover” (http://www.ncpad.org/get/fitnessCenter/index.html), and view a virtual accessible community fitness center at http://www.ncpad.org/get/VirtualTour/Welcome.html. The NCPAD’s programs database (http://www.ncpad.org/programs/) can help to locate fitness centers by geographic location and other search parameters.

Various indicators of exertion should be monitored during fitness programming so that the individual stays within an optimal and safe cardiovascular workout zone (Table 54-3). Rating of perceived exertion (RPE) should be monitored by

### TABLE 54.3 Cardiorespiratory Fitness Suggestions

- An exercise prescription for individuals with neurologic conditions should include the same four elements that are used with the general population: frequency, intensity, duration, and type. Check the ACSM Web site (www.acsm.org) for resources on developing exercise prescription guidelines.
- Any activity that increases energy expenditure above rest safely and effectively can be used to improve cardiorespiratory fitness.
- Patients taking β-blockers and those with autonomic dysfunction can have a blunted heart rate response. In these instances, consider using the RPE scale.
- Monitor RPE, heart rate, and blood pressure during early stages of activity and new programs.
- Teach proper breathing techniques (i.e., deeper vs. shallow breaths and “think tall” to maintain good posture) to avoid dyspnea (breathlessness).
- Vary workout to maintain interest.

### TABLE 54.1 Considerations Prior to Involvement in Fitness Training

- Inform physician of intentions to engage in physical activity if you are uncertain the patient/client is safe to participate in higher intensity levels of exercise.
- Determine side effects of medication on exercise.
- Consult a trained exercise professional to design an exercise prescription or identify resources for exercise programming for people with disabilities through the ACSM Web site (www.acsm.org).
- A graded exercise test may be warranted for some patients/clients if there is any concern about safely performing moderate- to vigorous-intensity exercise.
using the Borg RPE scale, which measures intensity of exercise on a scale from 6 to 20, or the conversation rule, in which one is able to converse while exercising. Heart rate must be gauged by finding the pulse and comparing this to the target heart rate.

In some cases, blood pressure should also be checked by an exercise professional during exercise or sport. For some individuals with physical disabilities, hypertension (i.e., stroke) or hypotension (i.e., spinal cord injury) is a common health condition. Once the individual adjusts to the program without wide fluctuations in blood pressure, it can be measured before and after activity.

Though a wide range of activities and sports are available to obtain a cardiovascular workout, including walking, wheelchair rolling, cycling, and swimming, as well as recreational and competitive sports (see http://www.ncpad.org/lifetime/ for more suggestions), specific types of cardiovascular exercise equipment can assist wheelchair users and individuals with lower-extremity impairments in obtaining a cardiovascular workout (Table 54-4).

An ergometer is an upper and/or lower body–driven exercise device that is powered from a stationary position and allows for a cardiovascular workout. The common stationary exercise bike is one type, though other variations exist for individuals with physical disabilities. Most ergometers use some form of hand or foot pedal for the basic motion, and some models allow for wheelchair access. See the NCPAD “Ergometers and Exercise Cycles” factsheet (http://www.ncpad.org/exercise/fact_sheet.php?sheet=5&view=all) for more information.

Ergometers come in upper-, lower-, and dual-extremity models. For example, the Saratoga Cycle can be used from a table top for an upper-body workout and/or placed on the floor for lower-body exercise through pedaling (Fig. 54-5). Hand grips or foot straps can assist with full, partial, or nonexistent hand or foot grip. Motorized movement trainers, such as the MotoMed Viva, allow for passive, active, and active-assist exercise for a lower-extremity workout for individuals with limited function. Wheelchair ergometers allow for upper- and lower-body workouts through a wheelchair being mounted on a support with the driving wheels mounted on an elongated central shaft.

Other types of ergometers and exercise cycles that can be powered by upper, lower, and dual extremities include the Nu

**TABLE 54.4** Sample Cardiovascular Exercise Equipment for People with Disabilities Who Are Unable to Use Their Legs to Exercise

- Upper arm ergometer
- Wheelchair ergometer
- Nu Step recumbent stepper
- BioStep Semi-Recumbent Elliptical
- Schwinn Airdyne (i.e., arm device only)
- VitaGlide

Step recumbent stepper and the Schwinn Airdyne exercise bike (Fig. 54-6). The Nu Step recumbent stepper includes a reclining backrest and a swivel seat, which simultaneously exercises arms and legs in a smooth, full range motion. Design features allow for either the lower and/or upper body to power the stepper, depending on the individual’s disability and strength capacity. The Schwinn Airdyne exercise bike is also designed with a progressive resistance pedal system, which can be used jointly or separately by the upper or lower body.

A variety of guidelines should be followed to help ensure the safety and effectiveness of a physical activity program or sport (Table 54-5). Most importantly, the patient/client must

**FIGURE 54-5.** A Saratoga Cycle® can be positioned on a table top for an upper-body workout or placed on the floor for lower-body exercise through pedaling. Courtesy of the Rehabilitation Institute of Chicago and the National Center on Physical Activity and Disability (www.ncpad.org).

**FIGURE 54-6.** Nu Step recumbent stepper, with a reclining back rest and a swivel seat, allows for an upper- and/or lower-body workout. Courtesy of the Rehabilitation Institute of Chicago and the National Center on Physical Activity and Disability (www.ncpad.org).
be aware of how his or her body is responding to the activity, so that appropriate adjustments can be made. Ideally, an exercise program should be varied, which can help to avoid overuse injuries. Balance should also be assessed to determine readiness for standing activities.

For wheelchair users, adequate precautions must be taken to avoid overuse injury, as well as additional injuries, due to the activity. Stretching is necessary to prepare individuals with spasticity for the workout, and using workout gloves and push rim padding can help to protect hands.

Strength training prescriptions for people with disabilities are based largely on disability severity, functional muscle mass, and health status (Table 54-6). Some patients/clients can train at very high intensity levels, while others will only be able to perform at minimal levels of resistance (i.e., lifting a body part against gravity). The training load (number of sets and repetitions, frequency, rest interval between sets) also varies by disability type, health status, and functional muscle mass. With respect to disability type, individuals who do not have a progressive disorder (i.e., spinal cord injury, cerebral palsy) can work at higher intensity levels than persons with progressive disorders (i.e., multiple sclerosis, postpolio).

An individual’s health status also affects how quickly a person is able to progress during the conditioning stage. Many individuals with physical disabilities who have been inactive for much of their lives need only a small amount of resistance exercise during the initial stage of the program to obtain a training effect. For example, a person with stroke and hypertension should not perform high intensity exercise. Individuals who are seizure prone or fatigue easily require a reduction in training volume.

The amount of functional muscle mass also affects training volume. Persons with paralysis, hemiplegia, impaired motor control, or limited joint mobility have less functional muscle mass and will therefore require a lower training volume. For individuals who cannot lift the minimal weight on certain resistance machines, resistance bands or cuff weights or even their own body weight can be used as the resistance.

Modes of resistance exercise consist of three general categories: free weights, portable equipment (i.e., elastic bands, tubing), and machines. While free weights may be preferable because they resemble functional daily activity (36), they require good trunk stability and may be difficult to perform for individuals with severe limitations in motor control and coordination. For individuals with very low strength levels, gravity-resistance exercise may be all that the person is capable of doing, and active-assistive exercise may be required for certain muscle groups that are too weak (paresis) or tight (spasticity) to be moved independently.

Precautions for assisting persons with disabilities with strength training include avoiding fatigue and exacerbations, assisting with asymmetrical weakness and hand dysfunction, and checking blood pressure routinely. Fatigue and delayed-onset muscle soreness in individuals with physical disabilities should be avoided, though this is common in new resistance training programs. Exercise should be stopped temporarily if soreness in certain muscle groups prevents the person from performing routine daily activities, and a physician should be contacted if fatigue persists. Similarly, individuals with asymmetrical weakness tend to hike their body disproportionately in resistance training and must be encouraged to use proper form.

For individuals with hand dysfunction who have difficulty grasping barbells or handles on different strength machines, a variety of specially-designed gloves allow the person’s hand to

---

**TABLE 54.5**  
Safety Guidelines and Reducing Injury

- Stop exercise if pain, discomfort, nausea, dizziness, light-headedness, chest pain, irregular heartbeat, shortness of breath, or clammy hands are experienced.
- Adequate fluid intake is necessary.
- Appropriate clothing can help to avoid overheating.
- Realistic short-term and long-term goals must be established.
- The exercise program must meet the patient/client’s goals.
- Overuse injuries can be avoided by varying exercise routine (e.g., cross-training) and using proper equipment.
- Pain and fatigue levels must be monitored closely.
- Balance must be assessed prior to engaging in standing activities (e.g., weight routine, aerobic dance class).
- Adequate stretching is necessary prior to the workout.

**TABLE 54.6**  
Strength Training Guidelines

- Strength training prescription and training load are determined by disability severity, functional muscle mass, and health status.
- Resistance goals must be clarified to determine strength training schedule.
- Proper breathing techniques and a complete range of motion must be followed.
- Blood pressure should be monitored during strength training.
- Adaptations may be necessary for individuals with hand dysfunction and asymmetrical weakness.
- Focus on stability, coordination, ROM, and timing.
- Wide benches, low seats, and trunk and pelvic strapping will help support and protect the person from injury during the exercise routine.
- Avoid “hiking” the body on the weak side.
- Avoid person holding breath while lifting weight (i.e., Valsalva maneuver).
- Use straps whenever necessary to keep body part in contact or alignment while lifting weight.
maintain contact with the resistance equipment. Gloves will also protect the hand from injury while performing resistance training routines. Individuals who do not have good grip strength can use wrist cuffs or leather mitts with velcro and buckles to secure their hands to dumbbells or weight equipment (Fig. 54-7). Blood pressure should also be monitored frequently during the early stages of the program, and as for some disabilities, hypertension or hypotension can be a common problem.

Finally, safe and effective strength training must include proper breathing techniques and extend through a complete range of motion. Breath should not be held during strength training, but the patient/client must exhale or breathe out while pushing the weight up or out and inhale or breathe in while letting the weight down or in. Proper posture facilitates effective breathing techniques.

Flexibility exercises help to improve range of motion, balance, and coordination, as well as the ability to carry out regular activities of daily living (Table 54-7). They can be performed every day and should be conducted before and after every cardiovascular and strength workout. A stretching program should last at least 10 minutes and should not be painful. Every muscle group used in a workout should be thoroughly stretched in a slow and fluid manner. They should be held and progress slowly, and more time should be spent on tight muscle groups. Types of flexibility training that can be easily adapted for people with disabilities include yoga and pilates.

Since many individuals with physical disabilities will have some degree of spasticity (tightness), flexibility training should always be combined with resistance training. An exercise professional must identify the “spastic” muscle groups and develop

FIGURE 54-7. Activity mitts or gloves can assist a person with hand dysfunction to grip equipment handles. Courtesy of the Rehabilitation Institute of Chicago and the National Center on Physical Activity and Disability (www.ncpad.org).
a long-range plan to increase range of motion. If the joint has been in a “fixed” position for many years, or if the spasticity is severe, it may not be possible to fully extend the joint. An exercise professional should consult with a physical therapist, physician, or appropriate medical professional to determine how to stretch a spastic muscle without causing injury. As with resistance training, use active-assistive stretching for certain muscle groups that are too weak (paresis) or tight (spasticity) to be moved independently.

### Involvement of People with Disabilities in Leisure and Recreational Physical Activities

As has been discussed with respect to the physical activity pyramid, rehabilitation professionals can motivate and assist patients/clients to incorporate a range of activities within their daily schedule, including residual, task-oriented, leisure, and fitness-related physical activities (see Fig. 54-3). In this section, types of leisure and recreational activities adapted for people with disabilities are reviewed, along with resources that the professional can use in connecting the patient/client with community programs and resources. Table 54-8 identifies activities falling into leisure and competitive categories along the physical activity continuum. While team sports are more likely to be competitive in nature, both individual and team sports can be competitive or enjoyed for recreational purposes.

### Leisure Physical Activity

Many types of leisure physical activities are now available to people with disabilities (Table 54-8). The importance of leisure physical activity is not necessarily to improve fitness, although with many activities, fitness benefits can be attained, especially in deconditioned individuals: the important element is that the person is participating in an enjoyable activity that raises energy expenditure above resting levels and reduces sedentary behavior. The other nice feature of leisure physical activity is that many of the activities can be done for longer than 30 minutes and sometimes up to a few hours. Examples of beneficial leisure activities include gardening, dog walking/wheelchair rolling, bowling, dancing, and recreational swimming, kayaking, and sailing.

### Gardening

Gardening is considered an excellent leisure activity for increasing energy expenditure and for using various muscle groups that are not typically used in general activities. Bending, reaching, twisting, digging, etc., are all aspects of gardening that provide health benefits. Gardening can help improve balance, strength, and flexibility, depending on the types, intensity, and duration of activity (Fig. 54-8). Thus, one can adjust one’s gardening for a heavier strength- and endurance-building workout or focus on less intensive activities using fine/gross motor skills, eye/hand coordination, as well as flexibility and balance.

A garden can be adapted for access in a variety of ways, starting with appropriate grades and paving, then careful selection and placement of planters, vertical gardening techniques such as hanging baskets, and, where possible, larger raised beds. These are all used to position soil and plants safely and comfortably within reach. A barrier-free or enabling garden can be as simple as an easily-reached window box hung from a balcony railing at waist height or an entire home landscape designed to be accessible and maintained by a wheelchair user.

Gardening tools and equipment also enable the individual to reduce effort, maximize abilities, and encourage independence while working in the garden. For example, long-handled tools decrease the need to extend or bend the body to work

---

**TABLE 54.7** Flexibility Guidelines for People with Disabilities

- Flexibility training should be incorporated before and after each cardiovascular and strength workout.
- Stretches should be held and progress slowly.
- Each major muscle group should be thoroughly stretched.
- Special precautions should be taken to avoid spasticity.

**TABLE 54.8** Sample Physical Activities for People with Disabilities

<table>
<thead>
<tr>
<th>Leisure</th>
<th>Competitive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bocce</td>
<td>Basketball</td>
</tr>
<tr>
<td>Canoeing</td>
<td>Cycling</td>
</tr>
<tr>
<td>Fishing</td>
<td>Dancing</td>
</tr>
<tr>
<td>Gardening</td>
<td>Football (American)</td>
</tr>
<tr>
<td>Horseback riding</td>
<td>Ice skating</td>
</tr>
<tr>
<td>Karate</td>
<td>Kayaking</td>
</tr>
<tr>
<td>Scuba diving</td>
<td>Snow skiing</td>
</tr>
<tr>
<td>Swimming</td>
<td>Table tennis</td>
</tr>
<tr>
<td>Tai chi</td>
<td>Tennis</td>
</tr>
<tr>
<td>Water skiing</td>
<td>Volleyball</td>
</tr>
<tr>
<td>Wheelchair basketball</td>
<td>Wheelchair rugby</td>
</tr>
</tbody>
</table>

---

**TABLE 54.8** Sample Physical Activities for People with Disabilities

<table>
<thead>
<tr>
<th>Leisure</th>
<th>Competitive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bocce</td>
<td>Basketball</td>
</tr>
<tr>
<td>Canoeing</td>
<td>Cycling</td>
</tr>
<tr>
<td>Fishing</td>
<td>Dancing</td>
</tr>
<tr>
<td>Gardening</td>
<td>Football (American)</td>
</tr>
<tr>
<td>Horseback riding</td>
<td>Ice skating</td>
</tr>
<tr>
<td>Karate</td>
<td>Kayaking</td>
</tr>
<tr>
<td>Scuba diving</td>
<td>Snow skiing</td>
</tr>
<tr>
<td>Swimming</td>
<td>Table tennis</td>
</tr>
<tr>
<td>Tai chi</td>
<td>Tennis</td>
</tr>
<tr>
<td>Water skiing</td>
<td>Volleyball</td>
</tr>
<tr>
<td>Wheelchair basketball</td>
<td>Wheelchair rugby</td>
</tr>
</tbody>
</table>
areas beyond the gardener’s reach. Comfort and gripping ability can be improved by modifying handles with soft padding. Knee pads or kneelers can be useful for tasks close to the ground. Such tools help protect muscles and joints from fatigue and injury in order to find the best match for the garden tasks being done.


• Internet key words: Search terms, such as enabling garden, healing garden, accessible garden, horticultural therapy, and barrier-free gardening, etc., to identify links to many different resources related to gardening for people with disabilities.

• American Horticultural Therapy Association: http://www.ahta.org can help identify a nearby horticultural therapist who is adept at working with people with disabilities in gardens. The association also provides referrals to local health care agencies and public gardens with horticultural therapy programs, educational opportunities, and publications.

• Chicago Botanic Garden’s Horticultural Therapy Services program: http://www.chicagobotanic.org/therapy provides demonstrations on enabling gardening, an extensive adaptive tool collection, publications, and educational programming.

Bowling
Bowling is a popular recreational and competitive activity that can be enjoyed by individuals with disabilities through minimal equipment adaptations. Traditionally a sport in which players score points by rolling a bowling ball to knock down pins, both indoor and outdoor variations on the game can be employed. Benefits include improved flexibility and strength, as well as social engagement.

Adapted equipment is available for individuals of all ability levels to participate and compete in bowling. These include ball ramps, assisting those unable to throw a bowling ball; ball pushers or bowling sticks, which provide the participant with increased control over the force of the throw and the angle at which the ball travels down the lane; and aids for persons with visual impairments (Fig. 54-9).

Organized bowling opportunities for people with disabilities include groups associated with Special Olympics International (SOI), the United States Deaf Bowling Federation (USADBF), the National Disability Sports Alliance (NDSA), the American Wheelchair Bowling Association (AWBA), and the American Blind Bowling Association (ABBA).

Resources: For more information, see the NCPAD fact-sheet on bowling at http://www.ncpad.org/lifetime/fact_sheet.php?sheet=16&view=all.

• NDSA: http://www.ndsaonline.org, T: 401–792–7130
• AWBA: T: 434–454–2269, bowlawba@aol.com
• USADBF: http://www.usdeafsports.org/, T: 605–367–5760

Dancing
Dance is a recreational or competitive activity for individuals with varying abilities. It involves creativity and expression while promoting movement, flexibility, and endurance. Dance
can decrease stress and increase muscle tone, physical endurance, and self-confidence. It may also increase brain activity and stimulate memory through the use of choreography and help prevent or lessen the effects of secondary conditions such as depression and social isolation.

The basic rules of wheelchair dance sport include the pairing of a male and a female partner, one of which must be a wheelchair user with at least a minimal disability that makes walking impossible (Fig. 54-10). Participants with disabilities in the lower parts of their bodies such as amputation, paralysis, cerebral palsy, and leg shortening (at least 7 cm) are eligible for competition. Participants are expected to have normal upper-body functions. There are two functional classes based on ability to maneuver a wheelchair, trunk rotation, and arm function. Competitive dance includes the waltz, tango, slow foxtrot, quickstep, samba, cha-cha-cha, rumba, and paso doble.

Social or recreational dance opportunities for people with disabilities can be located through accessible dance troupes and organizations. See resources below, and seek out similar opportunities locally. Classes may include ballet, hip hop, capoeira, modern dance, and jazz and can be adapted according to ability level.

Resources: For more information, see the NCPAD factsheet on dance at http://www.ncpad.org/fun/fact_sheet.php?sheet=35&view=all.


Swimming
Swimming is an excellent recreational, competitive, and therapeutic activity that can be enjoyed indoors or outdoors. In addition to swimming's therapeutic role, other aquatic activities are popular for recreational purposes. The benefits of swimming and aquatic exercise include improved cardiovascular function, reduced stress on joints, and water resistance improving muscle strength. Additionally, an individual's ability to swim and feel comfortable in the water create opportunities to become involved in other water sports including kayaking, canoeing, sailing, and water skiing.

For competitive swimming, a sport classification ensures that swimmers compete against others with similar functional abilities (Fig. 54-11). Swimmers with disabilities are encouraged to train under the supervision of a qualified coach who understands stroke mechanics and can recommend special stroke adaptations. USA Swimming teams (http://www.usa-swimming.org) are an excellent place to find a coach. Swimmers with disabilities can be easily integrated into existing clubs to obtain the necessary training and experience. With a few exceptions, swimmers are expected to perform strokes according to the rules of nondisabled swimming, and rulebooks are available from USA Swimming. The most elite swimmers compete in the Paralympic Games every four years.

Stair systems and pool lifts can be used for accessing pools, and all-terrain wheelchairs help to reach natural bodies
of water. Though not permitted in sanctioned swim events, different prosthetic devices can reduce resistance in the water during the recovery portion of a swimming stroke. Similarly, flotation devices help stabilize individuals in the water.


Organizations especially pertinent for persons with disabilities interested in swimming include

- USA Swimming: http://www.usa-swimming.org, T: 719-866-4578

Kayaking

Kayaking is a recreational activity in which individuals can experience the outdoors and obtain an excellent cardiovascular workout (Fig. 54-12). Kayaking can be practiced in many water-based locations and is easily adapted to various ability levels. The American Canoe Association (ACA) promotes the benefits of canoeing and kayaking while offering courses on the basics of paddling, safety and rescue courses, and instructor certification.

Adaptations and modifications include paddles that can be adapted for individuals with amputations or hemiplegia due to a stroke or cerebral palsy. Grip adaptations can assist paddlers with visual impairments in learning the proper grip and hand placement on the paddle. Proper fit and comfort within the kayak are essential because water increases the breakdown of skin and can lead to an increased risk of pressure sores. Closed-cell foam and adaptive seating systems may be used for increased contact with the boat, as well as to create a backrest or seat cushion.

Types of kayaks include tandem, open-decked, whitewater boats, inflatables, and sea kayaks.


- ACA: http://www.americancanoe.org/, T: 703-451-0141
- Paddlesports Industry Association: http://www.paddlesportsindustry.org/, T: 800-789-2202

Sailing

Sailing for people with disabilities can be practiced in both leisure and competitive contexts. It made its debut as a demonstration event in the 1996 Atlanta, GA, Paralympics, and became a full medal sport at the 2000 Sydney Paralympic Games. Using the International Association for Disabled Sailing (IFDS) Functional Classification System, Paralympic sailors with disabilities are classified according to their ability and can compete on a relatively equal basis.

With some adaptations, many standard boats are suitable for competition by people with disabilities. Types of adaptations include accessible and stable seats, transfer benches, and steering assists (Fig. 54-13). Seats and transfer benches can consist of low-tech furniture, such as lawn chairs or coolers, and specifically-designed equipment, such as a translating seat. Steering systems can also be adapted for the user. Bars and handles are secured throughout the boat for sailor stability and safety.
Local community organizations provide recreational sailing opportunities for people with disabilities. See the NCPAD’s program database to search for a program in your area: http://www.ncpad.org/programs/.

- IFDS: http://www.sailing.org/disabled
- U.S. Sailing: http://www.ussailing.org/swsn/links.asp, T: 401-683-0800

**Recreational Sports**

Key recreational physical activities that can be adapted for people with disabilities include sports such as water and snow skiing, tennis, golf, and cycling and often more competitive team sports such as basketball, softball, volleyball, and quad rugby. While team sports are more likely to be competitive in nature, both individual and team sports can be competitive or enjoyed for recreational purposes.

In addition to elite Paralympic sports, the U.S. Paralympics also runs a series of nonelite programs, including those based in the community, an academy program for youth aged 12–18, and a military division for veterans wounded through combat. The goal of such programming is to involve people with disabilities from the community to develop healthy habits and a healthy self-concept through engagement in sport and competition. For more information, see the Paralympics web link: http://www.usolympicteam.com/paralympics/.

**Water Skiing**

Water skiing is a recreational and competitive warm-weather activity that can be easily adapted for individuals with disabilities. Benefits include the opportunity to develop aerobic and strength capacity, to learn new skills, and to compete.

Adaptive equipment allows skiers with disabilities to enjoy a variety of water skiing activities (Fig. 54-14). Skiers with arm and leg amputations generally use the same equipment as skiers without disabilities, but with adaptations. A ski harness assists skiers who have difficulty holding onto the towrope. A ski boom, used by beginner skiers, is a 15-foot bar that attaches onto the boat. Outriggers can be attached to the sit ski to help skiers who have difficulty with balance. Skiers with disabilities may also use a board (ski) that is similar to a surfboard with a seat (cage) attached to the top. The cage, typically made of padded metal tubes, is deep so that the skier can sit securely.

Competitive water skiing includes three events—slalom, trick skiing, and jumping—and six broad category divisions for the disability water skiing events. The sit skier division includes individuals with spinal cord injuries (paraplegia, quadriplegia), bilateral leg amputations, or other conditions that affect the lower limbs. The vision impairment division includes people with complete and partial vision impairment. The leg division is for skiers who have an above or below knee amputation, in which skiers can ski with or without their prosthesis (Fig. 54-15). The arm division includes skiers with either an arm amputation or an impairment of one arm, and the arm and leg division for skiers with an impairment of both upper and lower limbs.

**Resources:**

- USA Water Ski: http://www.usawaterski.org/, T: 863-324-4341, a national governing body for water skiing, which works through the Water Skiers with Disabilities Association to serve individuals with disabilities.
Snow Skiing

Skiing is a competitive and recreational winter activity in which individuals use skis to maneuver downhill or cross-country. Benefits include improved cardiovascular conditioning, muscle core strength, balance, coordination, and flexibility. Skiing can be readily adapted for persons with disabilities and offers opportunity for competition or recreating with friends.

Adapted ski programs have grown due to the development of adaptive equipment for individuals with disabilities. Various organizations offer skiing events to encourage individuals of all abilities to participate. Information on adaptive equipment, teaching techniques, workshops, and certification clinics can be obtained from the Professional Ski Instructors of America (http://www.psia.org). Competitive alpine events include slalom, giant slalom, super giant slalom, and downhill. Cross-country events include men’s 10K and 30K and women’s 5K, 10K, and 4 × 10K relay races.

Skiing adaptations for people with disabilities include adapted skis and prosthetics, as well as adequate communication and instruction for the individual with the disability. The monoski consists of a bucket-style seat attached to a single ski, with outriggers to maintain stability. A biski has a bucket-style seat attached to two skis and outriggers attached to the sled, three track skiing requires two skis and two hand held outriggers, and four track skiing requires two skis and two hand held outriggers (Fig. 54-16). Skiing for persons with visual impairments involves good communication between the visually impaired skier and a sighted guide providing the ski instructions. The U.S. Deaf Ski & Snowboard Association (http://www.usa.org/), affiliated with USADSF (http://www.usdeafsports.org/), sponsors various championships, selects U.S. Deaf Ski and Snowboard Teams, and can provide information on skiing with a hearing impairment.


• United States Ski and Snowboard Association: http://www.usa.org/, T: 435-649-9090
• Ski for Light, Inc: http://www.sfl.org, T: 612-827-3232
• SKIFORALL Foundation: http://www.skiforall.org/, T: 425-462-0978
• United States Deaf Ski and Snowboard Association: http://www.usdssa.org, T: 435-752-2702
• USADSF: http://www.usdeafsports.org/, T: 605-367-5760
CHAPTER 54 | PHYSICAL ACTIVITY FOR PEOPLE WITH DISABILITIES

Tennis
Tennis is a competitive recreational activity that has become increasingly popular among individuals with disabilities. Benefits include improved aerobic and anaerobic fitness, gross and fine motor control, agility, balance, flexibility, hand-eye coordination, as well as bone strength and density. The sport is a cross-disability activity and can be played indoors or outdoors, in pairs or teams.

The two versions of tennis within disability sport include ambulatory tennis and wheelchair tennis. Ambulatory tennis is governed by two disability sport organizations, the United States Deaf Sports Federation and Special Olympics. Wheelchair tennis was developed in 1976 and is currently governed by the International Tennis Federation (ITF, http://www.itftennis.com/) and Wheelchair Tennis Committee, with advisement by the International Wheelchair Tennis Association (IWTA).

A wheelchair tennis match is played between two or four athletes with loss of motor function in one or both lower extremities (Fig. 54-17). Separate divisions exist for individuals by disability (quadriplegia or athletes with loss of motor function in the lower extremities and functional disability or amputation in the upper extremities), gender, skill level, and age. This allows for persons with disabilities to compete against able-bodied players.

To score a point, the ball must be hit in the opponent's court without the opponent returning the shot. The winner of the match is the athlete or pair to win three “sets” of six “games” each (best-of-three set match). Wheelchair tennis follows the same rules as able-bodied tennis as endorsed by the ITF, with several key exceptions, the most significant of these being the “two-bounce rule,” where the wheelchair tennis player is allowed two bounces before the ball must be returned across court. For more information about wheelchair tennis or an extensive list of rules, please refer to the ITF Wheelchair Tennis Handbook (http://www.itftennis.com/wheelchair/rules/wheelchairtennishandbook.asp).

Wheelchair tennis requires only a few pieces of equipment, a tennis racquet, ball, and a wheelchair, and is played on the same courts and surfaces as able-bodied tennis. No modifications are necessary for the racket and the ball. Players use a sports wheelchair that is lighter than everyday chairs to allow for ample flexibility, and the wheelchair is considered part of the player so that general rules of contact apply. To keep the player stable on the chair, a positioning strap across the waist and/or thighs is used. Grip devices are designed specifically for players who do not have the grip strength to hold a racquet. Arm and leg prosthetics are available for individuals who are amputees.

• ITF: http://www.itftennis.com/abouttheitf/, T: 44-020-8878-6464
• United States Tennis Association: http://www.usta.com/, T: 800-990-USTA

Golf
Golf is one of the most rapidly growing recreational activities in America, as it offers an opportunity for individual skill development, companionship, and socialization and can be played by persons of all abilities. Particularly for those who walk the course, golf provides for cardiovascular fitness, strength, and flexibility, as well as improved mental health.

Assistive devices for golfers with disabilities include specially-designed golf clubs, mobility devices, gripping aids, and practice facility equipment, including automated ball teeing devices, ball retrieval aids, etc. (Fig. 54-18). More information on golfing assistive devices can be found on the National Center on Accessibility (NCA) Web site at www.ncaonline.org. The U.S. Golf Association (USGA) maintains a “Modification of the Rules of Golf for Golfers with Disabilities” (http://www.usga.org/playing/rules/golfers_with_disabilities.html) for individuals wishing to play golf by these adjustments. Specific modifications are available for golfers with visual impairments, amputations, as well as those using canes, crutches, or wheelchairs.

Resources: The NCPAD golf factsheet (http://www.ncpad.org/lifetime/fact_sheet.php?sheet=245&view=all#1) provides more information about the game, golf professionals, organizations, and equipment suppliers.
• NCA: http://www.ncaonline.org/, T: 812-856-4422
• National Alliance for Accessible Golf: http://www.access-golf.org, T: 812-856-4422

FIGURE 54-17. A wheelchair tennis player. Wheelchair tennis follows the two-bounce rule, where the wheelchair tennis player must return the ball before it hits the ground a third time. Courtesy of the Rehabilitation Institute of Chicago and the National Center on Physical Activity and Disability (www.ncpad.org).
Cycling
Cycling is a recreational and competitive activity that allows individuals to obtain a cardiovascular workout while strengthening corresponding muscle groups. Individuals can cycle with upper or lower extremities, in indoor or outdoor locations. It is an official sport of DSUSA, SOI, the USABA, the USADSF, and the NDSA. Competitive hand cycling is governed by the United States Handcycling Federation (USHF).

Each of these disabled sports organizations work cooperatively with USA Cycling to develop cycling opportunities throughout the United States. Events range from 400-m junior races offered by NDSA to 40-km road races offered by DSUSA and USABA. A wide variety of adapted cycles exist, including handbikes, arm-driven cycles, dual-recumbent cycles, and tandem cycles (Fig. 54-19). Ergometers and stationary cycles reviewed earlier in this chapter are also suitable for indoor activity for people with disabilities.

Resources: For additional information, read the NCPAD factsheet on cycling at http://www.ncpad.org/competitive/fact_sheet.php?sheet=21&view=all.

- NDSA: http://www.ndsaoonline.org, T: 401-792-7130
- USADSF: http://www.usdeafsports.org/, 605-367-5760
- USA Cycling: http://www.usacycling.org, T: 719-866-4581
- USHF: http://www.ushf.org/, T: 303-459-4159

Basketball
Basketball is a team sport that can be played by individuals with a variety of disabilities. Though numerous variations exist, the two major versions of basketball popular within disability sport are ambulatory and wheelchair basketball. Benefits of playing basketball include the development of aerobic capacity and strength, as well as the opportunity to participate in different roles and positions on a team.

In wheelchair basketball, the wheelchair is considered part of the player, and therefore, general rules of contact apply (Fig. 54-20). The National Wheelchair Basketball Association (NWBA) has an extensive list of rules, and rule modifications exist from SOI, DAAA, and the NDSA. Basketball variations include bankshot basketball, which relies entirely on shooting skill and no running, dribbling, jumping, or body contact, and twin basketball, which is set up for athletes with cervical-level spinal cord injuries, with players surrounding a free throw circle.


- NWBA: http://www.nwba.org, T: 719-266-4082
- NDSA: http://www.ndsaoonline.org, T: 401-792-7130
- USADSF: http://www.usdeafsports.org/, 605-367-5760

Softball
The game of softball can be easily adapted for individuals with disabilities and played within competitive or recreational contexts. Benefits include participation in a team sport and social activity involving skill and strategy, as well as improved...
Variations on the game include wheelchair softball and beep baseball, and different softballs can be used for variations on the game. Batting and fielding equipment are available for individuals with upper-extremity amputations. Competition is offered by the USADSF (http://www.usdeafsports.org/), the DAAA (http://www.daaa.org), SOI (http://www.specialolympics.org), the National Wheelchair Softball Association (NWSA) (http://www.wheelchairsoftball.org/), and the National Beep Baseball Association (http://www.nbba.org).

Wheelchair softball requires that all players participate from a wheelchair, and all teams must have a team member with quadriplegia in play (Fig. 54-21). When on offense, the individual must be in the batting lineup, and on defense, the individual with quadriplegia must play the field. A team that chooses not to play a team member with quadriplegia must incur an automatic out every tenth batter. Each team must abide by the classification system that requires that no team can earn more than 22 points at any time during a game. If a player leaves his or her chair to gain a fielding advantage, all base runners advance two bases. For throwing errors in which the ball leaves the playing area, each base runner advances one base. Rule modifications (http://www.wheelchairsoftball.org/rules.htm) and general information are available from the NWSA (http://www.wheelchairsoftball.org/).

Beep baseball is an adaption on the game for individuals with visual impairment and blindness. The game consists of 4-ft-high padded cylinders with speakers and a sighted pitcher and catcher. Beeping tones are emitted as the pitcher places the ball, and as the ball is hit, the base operator activates a buzzing sound for the batter to identify and run to the base prior to the defense fielding the ball. For more information, contact the National Beep Baseball Association (http://www.nbba.org).

Resources: Access the NCPAD factsheet on softball (http://www.ncpad.org/competitive/fact_sheet.php?sheet=32&view=all) for more information on the game.

- DAAA: http://www.daaa.org, T: 972-317-8299

Volleyball

The game of volleyball can be played by persons of different ability levels in recreational or competitive settings. It provides a cardiovascular workout and helps to build muscle mass through rapid muscular movements. While standing and sitting volleyball are the two main forms of this activity, adapted versions of volleyball include wallyball and water volleyball.

Adaptations on the game allow for greater participation by persons with disabilities. Players with amputations can participate with or without prostheses in standing volleyball (Fig. 54-22). “Wallyball” is another version of volleyball played on a racquetball court. With the proper equipment, water volleyball can be played in many indoor or outdoor swimming pools. Sitting volleyball is played with six players per team on a smaller court with a lowered net. This version of volleyball enables double-leg amputees and individuals with spinal cord injuries, polio, and various other lower-extremity disabilities to participate. In certain recreational volleyball settings, a beach ball or balloon can be substituted for a standard volleyball.
United States Deaf Volleyball Association (USDVA) and the American Deaf Volleyball Association (ADVA) guide the game of volleyball for recreational and competitive players who are hearing impaired. Competitive standing and other variations of volleyball are played according to Federation Internationale de Volleyball (FIVB) standard international rules, with variations.


- World Organization Volleyball for Disabled (WOVD): http://www.wovd.info/

Wheelchair Rugby
Developed in 1977, wheelchair rugby is a competitive game in which two teams of four players in wheelchairs attempt to maneuver the game ball over their opponent’s goal line while in possession of the ball. Benefits include improved cardiovascular endurance, upper-body strength, mobility skills, and use of different muscle groups, as well as the opportunity to participate in an aggressive team sport combining strategy, power, and quickness. The game is played primarily by individuals who have quadriplegia, as well as those with cerebral palsy or postpolio, with seven classification divisions that range from 0.5 to 3.5 points based on functional ability.

The game moves at a fast pace with aggressive players. Offensive strategies are designed to exploit defensive weaknesses (Fig. 54-23). Defensive strategies involve blocking, forcing turnovers, and intercepting or batting away the ball from the offensive team. Only the four wheels of the chair (and the antitip device, if applicable) may touch the floor during contact with the ball. A goal is scored when a player carrying a ball crosses over the opponents’ goal line with two wheels. After a
goal is scored, a throw-in is used from any point on the end line to put the ball back into play. For more detailed information about quad rugby rules, consult the United States Quad Rugby Association Web site (http://www.quadrugby.com/rules/rules2006.htm).

Most rugby chairs have metal guards on the front, sides, and/or back of the chair to prevent opponents from hocking the chair during play. In addition, the wheels are attached at an angle (camber) for greater stability. Trunk, waist, leg, and foot strappings are allowed, depending on individual needs. Gloves and the use of taping at the forearms can protect against skin abrasions.


- United States Quad Rugby Association: http://www.quadrugby.com/toc.htm; T: 561-964-1712
- WSUSA: http://www.wsusa.org/; T: 732-422-4546

CONCLUSION

The transition from rehabilitation to community-based physical activity is a critical one. There is a small window of opportunity for rehabilitation and health professionals to establish effective community-based physical activity recommendations that are more likely to be adhered to if the program meets the person’s interest level, is readily accessible, and has an element of social engagement. The focus of this chapter was to provide professionals with the necessary resources to recommend tailored physical activity programs for individuals with disabilities. There are many types of physical activities that can keep people with disabilities engaged in this healthful behavior, including a variety of community-based fitness, leisure, and sport opportunities.

Most individuals recovering from an injury or health condition will appreciate the benefits of rehabilitation during the subacute period, but once the person returns home and stops performing certain rehabilitative exercises and increases his or her sedentary behavior, these gains are often lost or diminished and the risk of further health complications increases. Considering that generic physical activity programming not targeted to the needs of the particular end user is far less likely to result in long-term maintenance of health-promoting behaviors, a tailored approach is necessary for maintaining good adherence. The PEP model combines key factors including a person’s motivational level (readiness to change), physical activity profile, health and mobility limitations, and barriers to participation, so that the rehabilitation and health professional can prescribe a program that meets each person’s specific needs, interests, and circumstances.

Rehabilitation and health professionals play an essential role in connecting the patient/client to such community-based services. Armed with knowledge of the patient/client’s disability, functional limitations, and health status, as well as resources for encouraging various types of physical activities with appropriate adaptations and/or modifications, they can help transition the individual into a healthier lifestyle within a community-based setting. And as such, patients/clients maintain a physically active lifestyle. They will not only experience reduced secondary conditions and improved functional and health status, but they will also experience increased social integration leading to a higher quality of life.

ACKNOWLEDGMENT

This work was supported by the National Institute on Disability and Rehabilitation Research; Rehabilitation Engineering Center on Recreational Technologies and Exercise Physiology for People with Disabilities, Grant #H133E070029; and the Centers for Disease Control and Prevention, Grant #U59/CCU522742.

REFERENCES


“The body says what words cannot.”
Martha Graham, Dancer and Choreographer, 1894–1991

INTRODUCTION

Performing arts medicine is a branch of physical medicine and rehabilitation devoted to the care of musicians and dancers. A key concept of performing arts medicine is the ability to prevent, recognize, treat, and rehabilitate musculoskeletal injuries as related to each student, amateur and professional performing artist. This chapter discusses in further detail common injuries found in musicians and dancers, including injury management, rehabilitation, and barriers to care. The injuries primarily focus on the upper limbs for musicians and the lower limbs for dancers. Practitioners of performing arts medicine should have a solid understanding of musculoskeletal medicine and an appreciation for the physical effort and specific movement required to perform in each of the various arts.

DANCE MEDICINE

Most current medical literature focuses on ballet (1–4). Research must therefore be extrapolated to treat other types of dance. The majority of this chapter focuses on the injuries associated with ballet, understanding that the performing arts practitioner treats all types of dance including modern, jazz, tap, folk, ethnic, ballroom, and hip hop, all of which can lead to medical problems (5,6). Each style places demands on the body and may vary with respect to footwear, dance surfaces, training, and alignment.

Dance Training

Ballet training typically begins for female dancers by 5 or 6 years of age. Young dancers often begin training daily for 2 to 6 hours at a time with annual or semiannual performances by age 11 or 12. In order to pursue ballet professionally, preteen and teenage students with high-level ability and skills attend intensive summer programs, and those with the greatest ability and potential move to preprofessional ballet schools for year-round training. These exceptional students may participate in work-study programs with school in the morning, blended with class and rehearsals in the afternoon and evening. Some dancers obtain their general education degree early to join professional companies, while others continue to dance through college and join companies later. A large component of a dancer’s early career involves auditions. These tryouts can be critical to obtaining roles in performances and in the development of a professional career. Performing arts physicians need to acknowledge the high importance of auditions and should inquire about upcoming dates. This knowledge may aid in the decision to tolerate a higher risk of reinjury than under normal circumstances, as long as documentation of discussion and decision is thorough.

Terminology and Positions

Knowledge about the technical requirements for dance is important. Performing arts medicine practitioners excel at understanding the biomechanical demands placed on dancers and the maladaptations many dancers use to compensate for inadequate anatomy, conditioning, or biomechanics. Dancers should be evaluated for onset of symptoms and triggers, in addition to demonstrating the movements or positions that reproduce the symptoms (7) (Fig. 55-1). Common dance terms are

- *Plié*—heel remains on the floor with the knee in flexion (Fig. 55-2)
- *Grand plié*—heels lift off the floor with deep knee flexion
- *Pointe*—dancing on the tips of the toes, ankle in maximal plantar flexion (Fig. 55-3)
- *Demipointe*—weight bearing on the metatarsal joints, ankle plantar flexion, and metatarsal joint extension (Fig. 55-4)
- *Barre work*—dance classes are divided into three stages: barre, center, and movement from the corners across the floor. Barre work involves holding on to a bar during warm-up exercises and conditioning.
- *Center floor work*—The second stage of a dance class performed in the middle of the studio involves balancing, jumps, and short sequences of movement.

Causes of Injury and Risk Factors

Extrinsic

There are many common extrinsic and intrinsic causes of injury and risk factors in dance. One extrinsic factor is the type of dance floor. The presence of adequate shock absorption and the degree of surface friction can lead to injuries of the knee, foot, and ankle if there is too much surface resistance. Likewise, a floor that is too firm leads to fatigue and injuries such as tendonitis of the leg and foot (8).
Footwear can also influence injuries. Most ballet dancers wear either slippers, composed of only fabric or leather and no structural support, or pointe shoes, which are made of a solid toe box, metal or wood shank, and fabric. Theoretically, pointe shoes act as an additional stabilizer of the foot as demonstrated by cadaveric studies (9). However, these shoes are not designed to provide adequate stability or shock absorption and have changed very little since the 1600s (4,8). Furthermore, principal dancers may wear out one to three pairs of pointe shoes per performance because if the shoe becomes too soft it can often lead to injury, such as tendonitis and stress fractures (9).

Intrinsic
Intrinsic causes of injury or risk factors include, but are not limited to: malalignment of the lower limbs, muscle imbalance, and inappropriate training (7,8,10,11). Proper alignment in ballet is based on turnout (the maximal amount of external rotation at the hips), which will ideally enable a dancer to stand with the feet placed at 180 degrees in first position. Although dancers may appear to be able to achieve 160 to 190 degrees of turnout at the feet, they may not have the corresponding degree of external rotation at the hip (12,13). When the lower extremity is properly aligned in first position (14,15), the knee flexes directly over the foot. Many students will not be able to attain this ideal position and will compensate by forcing their turnout. "Rolling in" forces the feet into a position with a valgus heel and forefoot pronation (Fig. 55-5), leading to subsequent collapse of the medial arch and thereby increasing the torque on the ankle and knee (Fig. 55-6) (7,9,11). Young dancers, both female and male, will sometimes increase their lumbar lordosis in order to tip the pelvis anteriorly and thus obtain increased hip rotation. In addition to potentially leading to low back pain, other complaints may arise from the increased torque placed on the lower limb (7,11).

Muscle imbalance is often due to inadequate strength or altered flexibility. Ballet emphasizes hip flexion, external rotation, and abduction, which can lead to weakening of the antagonist muscles (8). Deficits in strength and endurance at the hip may lead to injuries of the foot, ankle, and knee (21–23). Focusing the treatment on the injured joint without addressing any underlying hip muscle imbalances may frustrate both dancers and physicians alike as pain and dysfunction are likely to recur. By nature, dancing en pointe involves loading the ankle joint in maximal plantar flexion with accompanying distal pressure placed on the first and second toes. Significant ankle flexibility and intrinsic foot strength are required to maintain the position (9). Elite and professional dancers usually display appropriate alignment as such alignment problems typically extinguish careers prior to this level. Generally speaking, at the professional level, limb injuries more often result from overuse rather than from issues of alignment. In young dancers, immature skeletal formation may contribute to musculoskeletal injury. Similarly, growth spurts may further lead to muscle imbalances, where a gain of 1 to 3 in. in growth might overwhelm the muscular strength required to move and change positions. As a result, young dancers can develop muscle tightness relative to bone length and/or inadequate strength, particularly in the hip rotators (16).

The effects of inappropriate training are often related to factors such as excessive duration and intensity (7,10). As in other aesthetic sports that place a large degree of importance on the appearance of the performer such as diving and figure skating, dancers often have delayed menarche due to caloric restriction coupled with large amounts of exercise (17). As dancers are often subjectively judged on their appearance, eating disorders are common in preprofessional dance schools with some reports estimating up to 30% of the students being effected (18–20). Such behavior can often place dancers at increased risk of injury.
risk of developing the “female athlete triad”: disordered eating, amenorrhea, and osteoporosis (7,10).

Incidence of Dance Injuries

In 2005, a UK national survey reported that 80% of 1,056 professional dancers sustained at least one injury per year (24). In a similar study of dancers at a sports medicine center, 90% of injuries occurred in the lower extremity and two-thirds were classified as overuse injuries (25). Of all the disciplines in dance, classical ballet necessitates the longest training, places the most physical demands on the musculoskeletal system, and studies indicate that 67% of all dance injuries occur from the practice of classical ballet (4).

Common Dance Injuries

Spine Injuries

Dancers routinely suffer spine injuries similar to many other athletes, including but not limited to fractures of the pars interarticularis and sacroiliac dysfunction. Delay in the treatment of pars fractures may progress to fibrous unions that will not calcify and potentially result in spondylolisthesis. If underlying biomechanical issues are corrected within an appropriate time frame, dancers may continue dancing with spondylolisthesis. However, dancers should restrict activity with relative rest in the event of a new pars fracture and may require either an anti-lordotic modified Boston brace or a nonspecific conventional lumbar corset (26).

Sacroiliac pain occurs frequently and is often difficult to resolve (7,27). Jumping and leg extension on the affected side may incite pain in dancers experiencing sacroiliac dysfunction. Special physical examination maneuvers such as Gaenslen’s test and the Gillet’s test are useful in diagnosis. Treatment of sacroiliac dysfunction may include mobilization of the joint and pelvic stabilization exercises.

Repetitive microtrauma from extreme extension or hyperextension may result in damage to posterior elements, such as the pars interarticularis, facets, pedicles, and spinous processes (28). Poor technique and control may cause painful injuries to the periosteum during hyperextension. Although this injury is painful, it requires no special treatment other than correction of technique unless it proceeds to pars fracture.

Hip Injuries

“Snapping hip” accounts for 50% of hip problems in dancers, occurring either medially or laterally with or without pain (4). When it occurs laterally, the iliotibial band or tensor fascia lata snaps anteriorly and posteriorly over the greater trochanter when landing from a jump in poor turnout and with excessive anterior pelvic tilt (7) or as the thigh moves between the anterior and posterior positions. “Snapping hip” may also occur medially and is sometimes referred to as “iliopsoas syndrome” (29) (Fig. 55-7). This phenomenon occurs when a hypertrophied iliopsoas crosses the femoral neck when bringing the hip into extension during a semicircular motion known as rond de jambe (7) (Fig. 55-8). A dancer with this injury will experience intense anterior groin pain during this arcing movement, particularly as the iliopsoas stretches over the femoral head from medial to lateral; it is also typically more painful while in the air than with the foot on the floor (30). “Snapping hip” rarely causes significant disability and responds well to
faced stretching, hip strengthening (especially the external rotators, adductors, and internal rotators), and antilordotic postural control (11,30). Other hip injuries such as labral tears, osteoarthritis, stress fractures of the femoral neck, hip avulsion injuries, and femoral neuropaxia have been described in the medical literature as well (7).

Knee Injuries
Knee problems are often attributed to insufficient hip strength in external rotation or forcing greater foot turnout than physiologically possible (as noted previously). Hip weakness or alignment problems are the usual causes of knee pain. The typical sports-related knee injuries also occur in dancers, including anterior and medial collateral ligament sprains, patellar subluxation, and patellofemoral syndrome (31,32). Tears and injuries often result from slips, twisting falls, and improper landings from jumps. Diagnosis and rehabilitation of such injuries follow standard sports medicine guidelines.

Acute Ankle Injuries
Ankle sprains are the most common acute dance-related injury (7). These sprains commonly occur when a dancer loses balance while landing from a jump while the foot is in plantar flexion. They can also occur from rolling over the lateral aspect of the foot while on demi-pointe (9). Interestingly, in full pointe, the ankle is relatively stable as the posterior lip of the tibia locks on the calcaneous and thus the subtalar joint locks in varus. It should be noted, however, that slight dorsiflexion relaxes the chain and may predispose the ankle to injury (10). As noted with other sports, the anterior talofibular ligament is the weakest ankle ligament and most susceptible to injury (9,10). The Ottawa ankle rules recommend ankle radiographs if a patient is unable to walk three steps or if there is tenderness over either malleoli (10). Specific rehabilitation of acute ankle injuries includes pool therapy, plantar flexion exercises, proprioceptive retraining, and peroneal strengthening with an emphasis on resuming full mobility of the talar, subtalar, and transverse tarsal joints (7,9).

Overuse Ankle Injuries

Anterior Ankle Impingement
In anterior ankle impingement, the anterior distal tibia and the talus pinch the bony or soft tissue during ankle dorsiflexion, particularly during grand plie (7) (Fig. 55-9). The medical history is often notable for chronic anterior or anterolateral ankle pain on jump landings or a limited demi-plié (9). Other mechanisms and the source of pathology of ankle impingements are found in Table 55-1.

Physical examination might include an effusion, audible click, palpable tenderness, and limited dorsiflexion when compared to the contralateral side. Pain is also present with passive dorsiflexion when the knee is bent (9). This is primarily
a clinical diagnosis. Treatment includes restricting pointe and center floor work, encouraging exercises at the barre, non-steroidal anti-inflammatories (NSAIDs), local injections, modalities, and in refractory cases, arthroscopic surgery (7). Return to a full plié position may take 3 to 4 months (9).

**Posterior Ankle Impingement**

Posterior ankle impingement is caused by the posterior tibia and calcaneus compressing the posterior talus and surrounding structures. This impingement may frequently be caused by a symptomatic os trigonum (7,9). The os trigonum is an accessory ossicle posterior to the talus which is often asymptomatic (11) (Fig. 55-10). When a dancer positions herself in extreme plantar flexion, the os trigonum is pinched between the tibia and calcaneus. Patients will present with posterior ankle pain, worsened with plantar flexion (11), and the pain and inflammation may limit the dancer’s ability to perform pointe, as well as the movement from flat foot to demi-pointe. Os trigonum syndrome will also cause pain with both active and passive plantar flexion (7). This pain is usually posterolateral, behind the peroneal tendons and often with associated ankle stiffness (9). Diagnosis is again primarily clinical; however, a standard lateral view ankle x-ray can demonstrate the presence of an os trigonum (7). Magnetic resonance imaging (MRI) may also show bony edema of the distal tibia and inflammatory changes surrounding soft tissue (7). Treatment of posterior impingement syndrome includes dance training modifications including limiting pointe-work, NSAIDs, physical therapy modalities, and exercises, with the refractory case leading to surgical excision (7). Changing pointe shoe wear to a half or three-quarter shank to more easily facilitate pointe-work may also provide some relief (9).

**Achilles Tendinosis**

Achilles tendinosis is classically seen in pointe overuse, excessive pronation and “forced turnout” during plié, or the wearing of too tight pointe shoe ribbons (10,11). Pain is localized to the zone of avascularity of the distal tendon, approximately 2 to 8 cm proximal to the insertion, and is often aggravated during landing from jumps (7). Physical exam is often notable for palpable tendon nodules, swelling, and crepitus (10). MRI and ultrasound are
useful tools for demonstrating tendon degeneration. Treatment includes gastrocnemius stretching, concentric to eccentric muscle strengthening, with approximately 25% of cases proceeding to surgical intervention to excise and debride the focal area of tendon degeneration (10). Stretch boxes strategically placed in the studio and theater have been shown to provide some degree of prevention for some dance companies (9).

Acute Foot Injuries

Subluxation

Cuboid subluxation may result in acute or chronic lateral midfoot pain (7). It is often associated with lateral ankle sprain injury (10) or is sustained during landing from a jump onto demi-pointe with the lower limb in external rotation (4). The medial border of the cuboid subluxes inferiorly, with dorsal displacement of the fourth metatarsal base, and plantar displacement of the metatarsal head (9,10). Physical exam is notable for tenderness over the cuboid bone, decreased mobility in passive supination/pronation, and a step-off of the fourth metatarsal head (10). A dancer may walk on the lateral edge of the injured foot as extreme supination is the most rigid position. Early diagnosis and treatment are key to preventing progression toward chronic injury. The cuboid should be immediately realigned using manipulation techniques (9).

Subtalar subluxation is caused by ankle hyperflexion, external rotation, and slight inversion, often occurs when performing grand plié en pointe, or landing from a jump on demi-pointe (33). Dancers may experience a sudden sharp talonavicular pain and a sensation of forward displacement of the foot, interfering with ambulation (7,33). Physical exam shows tenderness to palpation of the talonavicular ligament, anterior taloﬁbular ligament, anterior tibiotalar ligament, posterior joint, and hypomobility of the subtalar joint (33). Imaging has not shown to be useful due to anatomic variation (4). Treatment includes manual reduction, taping, and relative rest for 6 weeks. Dancing may be started in a pool after the initial 2 weeks. Proprioceptive training may begin after 3 to 4 weeks followed by continued close observation (7,33).

Metatarsal Fractures

Metatarsal fractures generally occur at the fifth metatarsal bone. These fractures are usually caused by missed landings and rolling over the lateral border of the foot on demi-pointe (9). Dancers present with tenderness, swelling, and ecchymosis over the fifth metatarsal bone, and lateral foot pain particularly with weight bearing (9,10). The most common proximal fifth metatarsal fracture is an avulsion fracture, which is commonly caused by an inversion injury (10). A “Jones fracture” involves the metaphyseal-diaphyseal junction, and often leads to nonunion due to the poor blood supply. “Dance’s fractures” are oblique or spiral fractures of the middistal fifth metatarsal associated with twisting and inversion (10). Surgical intervention is needed for high-level dancers to avoid prolonged immobilization for Jones fractures and diaphyseal fractures (9,10). Avulsion fractures and Dancer’s fractures can typically be treated conservatively and symptomatically with immobilization with a removable cast boot (9,10). Once fractures have healed adequately, rehabilitation goals return to strengthening the intrinsic foot musculature, retraining proprioception, initiating aquatic therapy, and gently returning to dance.

Lisfranc’s Joint Dysfunction

Midfoot sprains most commonly involve injury to Lisfranc’s joint, where the second metatarsals articulate with the intermediate cuneiform. The Lisfranc’s joint is crucial to midfoot stability and the longitudinal and midfoot arches of the foot (9,10). This type of injury is typically caused by a loss of balance on pointe, or spinning and landing jumps which subsequently result in an inability to roll through demi-pointe or full pointe (9,10). Physical examination is significant for midfoot swelling, ecchymosis, and tenderness to palpation and pain with both passive pronation and supination (9,10). Imaging may reveal avulsion fragments or diastasis between the first and second metatarsal bases (9). Treatment depends on the severity of the injury, from non-weight-bearing cast immobilization to operative fixation if diastasis is present (10). However, a simple strain without evidence of instability should be treated nonsurgically (9).

Second metatarsal stress fractures at Lisfranc’s joint are almost exclusive to women practicing pointe-work (30). Ballerinas with long second toes are particularly vulnerable to Lisfranc’s fracture (34), as the load through the foot is carried mostly on the first two toes with partial weight bearing of the third toe. Studies implicate that the fully plantar ﬂexed or hyperﬂexed forefoot places the highest risk of injury due to the anatomy of the Lisfranc’s joint (30), and increased load on a long second toe can thus lead to fracture at the base of the second metatarsal head or middle cuneiform bone (35–37).

Overuse Foot Injuries

Metatarsal Stress Fractures

Most stress fractures in dancers occur in the metatarsal bones (9). Multiple risk factors have been identiﬁed in placing dancers at risk for stress fracture including, but not limited to, long duration of practice hours per day and the female athlete triad (9). Early diagnosis improves outcomes (7), and dancers typically present with poorly localized midfoot pain initially after dance, but progressing to pain during activity and at night. Bone scans and MRI will reveal bone marrow edema and microfractures. Diaphyseal stress fractures are often seen with repetitive adduction, movements that involve cutting and pivoting. These fractures have poor healing potential due to their poor blood supply in the metaphyseal-diaphyseal region (9,10). Treatment of stress fractures involves immobilization or use of a removable cast boot with custom molded footplate for 6 to 12 weeks. The contralateral shoe may need to be modified, as a cast boot or footplate may cause a leg length discrepancy leading to sacroiliac dysfunction, and adjusting the height of the opposite shoe may help prevent imbalance and pain. Weight bearing is progressed as tolerated and aquatic therapy or Pilates
starting as soon as possible is recommended for approximately 6 weeks and until the fracture has healed. A padding system or toe cap for the second metatarsal and properly fitted pointe shoes may subsequently support the Lisfranc region, enabling a more balanced weight distribution (7).

**Flexor Hallucis Longus Tendon Dysfunction**

Flexor hallucis longus (FHL) tendon dysfunction is also known as “dancer’s tendonitis” (38). This condition often occurs with symptomatic os trigonum. The FHL is used as an accessory push-off muscle for plié, transitioning from flat foot to demi-pointe, and during jumps (30). These repetitive movements from full plantar flexion to dorsiflexion cause the FHL to become compressed within the fibro-osseous tunnel posteromedial to the talus with subsequent inflammation (9). Dancers present with pain when jumping and dancing on pointe and may have crepitus and triggering, or locking of the big toe in a flexed position much like a trigger finger of the flexor tendons of the hand (10). Physical examination may further reveal tenderness to palpation along the course of the tendon with painful passive extension of the big toe (7). Conservative treatment includes limiting pointe-work and jumps, correcting malalignments related to forced turnout, NSAIDs, physical therapy, and aquatic barre exercises (4,7,9).

**Sesamoid Disorders**

Sesamoid injuries occur in dancers due to the increased load on the sesamoid bones when rolling through demi-pointe, forcing turnout, out-toeing gait, and during abrupt landings (7). Sesamoids are located both in the tendon of the flexor hallucis brevis (FHB) and on the plantar surface of the proximal first metatarsal-phalangeal joint (MTP) and increase the mechanical advantage of the FHB. Dancers typically present with pain on the plantar aspect of the first MTP joint (10). Physical examination demonstrates tenderness to palpation of the sesamoid, pain with resisted MTP flexion, and restricted passive dorsiflexion (10). Injecting a small amount of local anesthetic can confirm the diagnosis (9). MRI can help differentiate bipartite sesamoids with smooth edges versus acute sesamoid fractures (9). Treatment is to correct alignment problems including avoiding excessive turnout, restricting demi-pointe, using sesamoid pads, wearing a stiff-soled shoe outside of class, physical therapy for mobilization of the joint, restoration of range of motion, balance/proproprioeptive retraining and intrinsic foot strengthening, and occasionally corticosteroid injections (9,10).

**Hallux Rigidus**

Hallux rigidus presents as arthritic changes of the first MTP joint. Dancers experience pain at the first MTP joint with dorsiflexion, particularly on demi-pointe, which causes the dancer to roll onto the lateral metatarsals. Physical examination demonstrates tenderness to palpation of the first MTP, decreased passive dorsiflexion, and imaging may reveal osteophytes, joint space narrowing, or subchondral sclerosis (9,10). Early treatment is usually conservative, including avoidance of demi-pointe and strengthening of the intrinsic foot muscles (9). While surgery may be an option, it may not lead to full restoration of movement, and some of these dancers may need to retire (9).

**Hallux Valgus**

Hallux valgus is a painful bunion on the medial border first MTP and is commonly exacerbated with pointe-work. Dancing in pointe shoes does not cause bunion deformities (9); however, the condition can be exacerbated by poor technique, inadequate shoe inserts, flat first MTP surface, and long second toe (4). Treatment includes orthotics for regular walking shoes, NSAIDs, modalities, and foot strengthening of FHL, adductor hallucis and flexor digitorum longus with the use of exercise bands to simulate releve, pushing the foot from demi-pointe to pointe, progressing to weight-bearing exercises (7), with surgery reserved for retiring professionals as intervention may lead to loss of range of motion (9).

**Dance-Specific Rehabilitation**

Prevention remains the key in dance rehabilitation. Dance screening clinics may provide screening for adequacy of training, physical conditioning, malalignment, poor technique, prior injuries, and other medical problems (7). Ideally, dancers should be assessed in functional foot movements of plantar flexion, dorsiflexion, pronation, supination as well as plié, demi-pointe, parallel position, and at the individual affected joint (10). Much of dance rehabilitation is similar to sports medicine management including pain management, restoration of range of motion, muscle strengthening, and balance and proprioceptive retraining. Studies also indicate that supplementary physical conditioning or cross-training can both improve fitness and decrease injuries in dancers (39). Modification of dance activities may include strengthening initially in neutral position where the feet are in parallel position and not turned out, and then progressing from barre exercises to the center floor (10). Pilates-based exercises are often very effective for promoting balance and coordination (7) and enable injured dancers to remain relatively fit while recovering from injury. While acknowledging that improper technique may increase the likelihood of injury recurrence, rehabilitation should emphasize biomechanical reeducation, possibly with a transitional dance class supervised by both a specialized physical therapist and dance coach (7) or supervision with both a dance-medicine trained practitioner and dance instructor (30). Additional intervention may include psychological care for stress management, anxiety and coping with limitations, identifying proper shoes, nutritional assessment to screen for eating disorders, and advocating for general health improvement including smoking cessation (7).

The benefit of dance rehabilitation is cited in several studies (40,41). One study demonstrated savings in one ballet company of $1.2 million over a 5-year period as well as a decrease in annual incidence of injury with the provision of in-house medical and therapy services (40). A second
MUSCULOSKELETAL PROBLEMS IN THE MUSICIAN

Musicians of all ages and playing levels are susceptible to musculoskeletal pain and injury due to repetitive and unique biomechanical movement required to play an instrument. Many studies have reported the incidence of such injuries. The International Conference of Symphony Orchestra Musicians published a survey in 1988 that revealed 76% of 2,212 musicians had experienced an injury that affected their performance (42). A study of professional orchestra musicians in Puerto Rico published in 2007 revealed increased risk in younger musicians (ages 22 to 29) and older musicians (ages 50 to 61). Risk was also higher for females, increased playing time and technical difficulty (43).

Most of the injuries observed in musicians are similar to musculoskeletal injuries seen in the nonmusicians. It is imperative, however, to understand the biomechanics involved in developing these injuries in order to treat this population. In many cases, an isolated musculoskeletal injury can be career ending or life altering. In the last two decades, there have been numerous articles published on this subject. Terms describing musician-related injuries include overuse injuries, cumulative trauma disorders, muscle-tendon pain syndromes, playing-related musculoskeletal disorders, and strains. The incidence of specific injuries depends on the instrument played (Table 55-2). The remainder of this chapter reviews common musculoskeletal disorders in the musician organized primarily by body part. The emphasis is on risk factors and treatment unique to musicians.

For those not familiar with instruments and terminology, below is a brief overview (44):

Woodwind instrument—produces sound when the player blows against the edge of or opening in the instrument, causing air to vibrate within a resonator. These include the flute, piccolo, clarinet, bassoon, saxophone, and oboe.

Brass instrument—one is produced by vibration of the lips, as the player blows into a tubular resonator. These include the trombone, trumpet, tuba, and the French horn.

String instrument—produces sound by means of vibrating strings. These include the violin, viola, cello, bass, electric guitar, bass guitar, acoustic guitar, harp, and banjo. The violin is the smallest of all the strings and the most common orchestral instrument.

Percussion instrument—any object which produces a sound by any action which causes it to vibrate and produce sound. These include the snare and bass drums, cymbals, and in some references, the piano.

Keyboards: such as the piano, organ, and accordion are sometimes classified in the percussion group and sometimes in the strings. For the purpose of this chapter, the term keyboards or the specific instrument is used.

Embouchure—position of the lips and tension of the face to produce a good tone on a brass or woodwind instrument.

Disorders of the Head and Face

The temporomandibular joints (TMJs) connect the temporal bone to the mandible. TMJ disorder occurs when the surrounding muscles become hyperactive causing malocclusion or when the meniscus of the joint becomes displaced causing poor articulation (45). TMJ occurs in violin and viola players due to pressure on the mandible, clenching mastication muscles, and from vibrations transmitted from the instrument. It may occur in trumpet, trombone, and tuba players due to displacement of the mandible. Treatment includes modifying the shoulder rest, use of occlusal splints, and physical therapy. The goal of physical therapy is to increase joint flexibility and decrease muscle tightness. Modalities include ultrasound, TENS, massage, biofeedback, and myofascial spray and stretch (46).

Embouchure overuse syndrome is the most common performance injury suffered by brass players. Symptoms include lip pain, swelling, embouchure weakness, and lack of endurance. Chin bunching indicates embouchure fatigue or evidence of a problem in technique (Fig. 55-11). Rest is essential to recover from these symptoms. Once they improve, working with an appropriately trained music instructor to optimize technique and avoid excess strain can prevent further injury or even rupture (47). Louis Armstrong suffered a rupture of the orbicularis oris from repeated straining to play the trumpet. Surgical intervention is required if rupture occurs. A useful website for patients and clinicians is www.embouchures.com.

Disorders of the Neck and Shoulder

Playing an instrument involves dynamic and static positions that can cause discomfort in the neck. Violin and viola players support the instrument between their left shoulder and left side of the neck and mandible. This prolonged static position can lead to myofascial pain and eventually, degenerative changes in the cervical spine. Cervical radiculopathy is more common on the left in violinists due to prolonged neck rotation and flexion (47). Support devices such as shoulder rests and chin rests can be used to optimize position. A properly fitted violin allows the player to turn his/her head and tilt it slightly to stabilize the violin. The hand is not required to support the instrument distally and the neck should not strain to hold it (Fig. 55-12). Neck pain can occur as the artist clamps the chin down on the chin rest to hold a poorly balanced instrument. The chin and shoulder rests are often placed too low, causing

retrospective cohort study conducted with the Alvin Ailey Dance Company revealed that comprehensive management inclusive of prevention, intervention, and case management, over a 5-year period, significantly decreased the number of worker’s compensation cases from 81% to 17% and decreased the number of lost workdays by 60% (41).
the musician to raise her shoulder or thrust the head forward to compensate (47).

Keyboard players may develop neck pain from looking down at their hands, especially on the right side. Cervical radiculopathy is more common on the right side in pianists (48). Orchestra partners must keep their head turned to one side to read the music on the stand. The flute player is also at risk for cervical degenerative changes due to static position of the tilted and rotated head. There is a mouthpiece available that has a 30 degree obtuse angle just past the lip plate. This allows for a near neutral position. Shoulder injury is included in this section because many of the positions that contribute to neck pain also contribute to shoulder pain. Myofascial pain in the shoulder is common on the same side of the neck from positions described above. Rotator cuff impingement is likely

<table>
<thead>
<tr>
<th>Woodwinds</th>
<th>Clarinet</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flute</td>
<td>Throracic outlet syndrome (L or R)</td>
</tr>
<tr>
<td></td>
<td>Back and neck pain</td>
</tr>
<tr>
<td></td>
<td>DeQuervain’s tenosynovitis (L or R)</td>
</tr>
<tr>
<td></td>
<td>Focal dystonia fourth and fifth digits (L)</td>
</tr>
<tr>
<td></td>
<td>Ulnar neuropathy (L)</td>
</tr>
<tr>
<td>Oboe</td>
<td>Lateral epicondylitis (R or L)</td>
</tr>
<tr>
<td></td>
<td>Ulnar neuropathy (R)</td>
</tr>
<tr>
<td></td>
<td>Back or neck pain</td>
</tr>
<tr>
<td></td>
<td>DeQuervain’s tenosynovitis</td>
</tr>
<tr>
<td>Clarinet</td>
<td>Thoracic outlet syndrome</td>
</tr>
<tr>
<td></td>
<td>Lateral epicondylitis (R or L)</td>
</tr>
<tr>
<td>Saxophone</td>
<td>Lateral epicondylitis (R)</td>
</tr>
<tr>
<td></td>
<td>Temporomandibular joint disorders</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Brass</th>
<th>French horn</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trombone</td>
<td>Focal dystonia of lip</td>
</tr>
<tr>
<td></td>
<td>Lateral epicondylitis (R)</td>
</tr>
<tr>
<td></td>
<td>Orbicularis oris strain</td>
</tr>
<tr>
<td>Tuba</td>
<td>Orbicularis oris strain</td>
</tr>
<tr>
<td></td>
<td>Temporomandibular joint disorders</td>
</tr>
<tr>
<td></td>
<td>Orbicularis oris strain</td>
</tr>
<tr>
<td></td>
<td>Lateral epicondylitis (R)</td>
</tr>
<tr>
<td></td>
<td>Temporomandibular joint disorders</td>
</tr>
<tr>
<td></td>
<td>Maxillofacial and lip trauma</td>
</tr>
<tr>
<td></td>
<td>Pharyngeal dilatation</td>
</tr>
<tr>
<td></td>
<td>Violin/viola</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Strings</th>
<th>Thoracic outlet syndrome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guitar</td>
<td>Carpal tunnel syndrome</td>
</tr>
<tr>
<td></td>
<td>Dorsal interosseous strain (L)</td>
</tr>
<tr>
<td></td>
<td>Back and neck pain</td>
</tr>
<tr>
<td></td>
<td>Focal dystonia first, second, third digits (R)</td>
</tr>
<tr>
<td>Harp</td>
<td>Neck pain</td>
</tr>
<tr>
<td></td>
<td>Medial epicondylitis (L)</td>
</tr>
<tr>
<td></td>
<td>Flexor and extensor tenosynovitis of thumbs</td>
</tr>
<tr>
<td>Cello/string bass</td>
<td>Neck pain</td>
</tr>
<tr>
<td></td>
<td>Ulnar neuropathy (L)</td>
</tr>
<tr>
<td></td>
<td>Rotator cuff tendonitis (R)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Percussion</th>
<th>French horn</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drums</td>
<td>Medial and lateral epicondylitis</td>
</tr>
<tr>
<td></td>
<td>DeQuervain’s tenosynovitis</td>
</tr>
<tr>
<td></td>
<td>Carpal tunnel syndrome</td>
</tr>
<tr>
<td></td>
<td>Medial and lateral epicondylitis</td>
</tr>
<tr>
<td></td>
<td>Carpal tunnel syndrome</td>
</tr>
<tr>
<td></td>
<td>DeQuervain’s tenosynovitis</td>
</tr>
<tr>
<td></td>
<td>Neck and back pain</td>
</tr>
<tr>
<td></td>
<td>Focal dystonia of thumb, finger, hand, foot</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Keyboards</th>
<th>Thoracic outlet syndrome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Piano/organ/accordion</td>
<td>Medial and lateral epicondylitis (R and L)</td>
</tr>
<tr>
<td></td>
<td>Carpal tunnel syndrome</td>
</tr>
<tr>
<td></td>
<td>DeQuervain’s tenosynovitis</td>
</tr>
<tr>
<td></td>
<td>Neck and back pain</td>
</tr>
<tr>
<td></td>
<td>Focal dystonia of thumb, finger, hand, foot</td>
</tr>
</tbody>
</table>

|           | Rotator cuff tendonitis (R) |

in musicians that must keep their arms in a raised position with elbow pointing outward or forward. Instruments that require this position include the violin, viola, cello, and bassoon. In flute players, the right shoulder is more commonly affected than the left because it is abducted and externally rotated (Fig. 55-13A and B).

The physician who treats musculoskeletal injuries is sure to see numerous patients with neck and shoulder pain. Current literature does not suggest that these problems are more common in musicians than in the general population, but when a musician presents with neck and shoulder discomfort, the instrument is usually the cause as it requires repetitive motion and unnatural positions. A study published in 2004 (49) in Sweden compared the risk of musculoskeletal injury in actors and orchestra musicians. The neck and shoulders were the most frequently injured anatomical areas and the prevalence of 25% was similar in both groups. The orchestra musicians, however, had a threefold to fivefold increased risk for pain affecting their performance capacity (49). Upper body strengthening and stretching to compensate for these postural asymmetries is helpful in these patients.


FIGURE 55-12.  A violin that fits properly requires no hand support and does not strain the neck.

FIGURE 55-13.  A: Shoulder abduction and external rotation required to play the flute may lead to impingement and shoulder fatigue. B: The flautist may drop the shoulder to avoid discomfort which strains the neck.
Playing an instrument in a slumped position contributes to tight anterior muscles such as the pectoralis muscles and scalenes.

Thoracic Outlet Syndrome
Thoracic outlet syndrome (TOS) is a condition which may result in compression of the brachial plexus or subclavian vessels. This may cause paresthesias of the digits, vascular changes, and/or weakness of the hand. Musicians can be particularly susceptible owing to their need to maintain shoulder abduction or extension for prolonged periods of time (Fig. 55-14) (50). Musicians commonly have tight pectoralis and scalene muscles, which can cause compression of the neurovascular bundle. Irregular breathing patterns and breath-holding may be a risk factor in woodwind and brass instrumentalists (51). The presence of a cervical rib and the first rib are known contributors to TOS. It is imperative to stretch the tight anterior muscles and to strengthen the posterior scapular stabilizing muscles to treat this condition.

TOS has been reported to be more common on the left side in violin, viola, and guitar players, and bilaterally in flute and piano players. In one study, there was a 12% prevalence of TOS among musicians, most prevalent in pianists, string, and wind players (48). In an earlier study, the same author found that physical therapy to correct postural errors and range of motion was successful in 76% of 17 musicians with TOS (52).

Disorders of the Arm and Hand
Entrapment of the ulnar nerve may occur between the two heads of the flexor carpi ulnaris in the cubital tunnel or in the bony sulcus between the medial epicondyle and the ulnar olecranon process. As in ulnar neuropathy of non-musicians, excessive elbow flexion is the most significant risk factor. It is more likely to occur in the left arm of viola, violin, and guitar players, due to continuous elbow flexion. Treatment consists of position awareness and limiting elbow flexion. This is not usually feasible while playing, but needs to be emphasized during other activities and sleep.

Medial epicondylitis occurs with repetitive wrist flexion, finger flexion, or pronation of the forearm. Medial epicondylitis occurs in the harpist on the left side. Lateral epicondylitis is due to repeated wrist extension, finger extension, or supination. It is more likely seen in clarinet players bilaterally and in oboe and trombone players on the right side. Percussionists and keyboard players are at risk for bilateral medial and lateral epicondylitis. Proper bench height is important to maximize a neutral wrist and elbow position. Forearms should be parallel to the floor when touching the keyboards (Fig. 55-15A–C).

Repetitive wrist flexion with ulnar deviation leads to DeQuervain tenosynovitis. This is an inflammation of the extensor pollicis brevis and adductor pollicis longus muscles on the radial side of the wrist. The clarinet, oboe, flute, keyboard, percussion, violin, and viola are the primary culprits. Use of a spica splint to immobilize the thumb as much as possible is helpful. A localized steroid injection provides relief in 81% of patients at 6 weeks (53). Ergonomics can be improved by repositioning the proximal arm so the left hand does not have to twist so radically to perform fingerling work (Fig. 55-16).

Carpal tunnel syndrome (CTS), entrapment of the median nerve at the wrist, is the most common entrapment neuropathy in musicians (54). As in non-musicians, the primary risk factor is repetitive wrist flexion. Because CTS is such a prevalent disorder among the general population, causality from instrument playing is less clear than in other musculoskeletal disorders. Dawson studied 98 musicians with CTS and reported that only 18.4% had no risk factors for CTS other than playing music (55). Among right-handed musicians, the left hand in violin, viola, and guitar players is more likely to develop CTS. Treatment includes rest with splinting, icing, physical therapy and especially, technique modification (Fig 55-16). Fortunately, for musicians with CTS, over 90% have complete recovery with surgical and nonsurgical treatment (55).

The actual manipulation of keys, valves, and strings is performed mostly by the small intrinsic muscles of the hands (56). The intrinsics include the thenar, hypothenar, interossei, and lumbrical muscles. Involvement of these small muscles is less studied and referenced in the literature. Dawson, however, addressed this problem in a retrospective chart review of 1,354 instrumentalists with hand and upper extremity injuries, 51 of whom had strains of hand intrinsic muscles. Risk factors included playing keyboards, string instruments, female sex, and higher-level musicians. It is hypothesized that females are more affected due to smaller, and therefore, weaker hand muscles (56). Treatment includes relative rest, modification of technique, analgesics, and occupational therapy, which can address ergonomics of playing.

Disorders Related to Joint Laxity
Many authors have described joint laxity as a risk factor for injury among musicians. Brandfonbrener found that 34% of 128 musicians presenting with arm and hand pain had hyperlaxity, which appeared to play a role in the etiology of pain (57). Many joints can have hyperextensibility. Having more than 10 degrees beyond 180 degrees of proximal interphalangeal...
joint extension and/or 10 degrees or more beyond 90 degrees of metacarpophalangeal joint extension seems to put musicians at most risk (58). More women than men have joint laxity. Smaller hands are also a risk factor for hand injury in musicians, thus the clinician should be especially vigilant when dealing with hand pain in women musicians (58). Awareness, education, supportive splints or taping, and joint protection are the cornerstones of treatment (Fig. 55-17).

**Disorders Related to Focal Dystonia**

Focal dystonia is characterized by involuntary muscle contractions that selectively interfere with the execution of specific motor tasks such as writing or playing a musical instrument (59). Similar to writer’s cramp, a focal dystonia of the hand, dystonia may affect the fingers of instrumentalists and more rarely, the embouchure in wind instrumentalists (60). The dystonia causes the finger or lip to contract when it should not, or otherwise move involuntarily (Fig. 55-18). Lederman found a pattern of involvement of the left hand in string players and the right hand in keyboard and wind players (61). Keyboard players appear to be most commonly affected of all instrumentalists, usually involving the fourth and fifth digits of the right hand which plays melody more than the left hand (62). Dystonia of the fourth and fifth digits may be more common in patients who have developed ulnar neuropathy (62).

A study of patients with focal hand dystonia suggested that these patients have alterations in the cortical somatosensory representation of the digits. The representation of the fingers spans outside the area normally occupied in persons without focal dystonia. This may affect sensorimotor feedback and cause involuntary co-contraction of other muscles (63). The onset is insidious and usually painless (64). It should be
considered when a musician begins to experience difficulties with speed and precision of playing and when he or she has to work harder to play. It is helpful for the examiner to observe the musician playing and watch for subtle signs of slow, hesitant, or abnormal movements of fingers or lips (65).

Treatment of focal dystonia is challenging. An evidence-based review by the American Academy of Neurology in 2008 concluded that botulinum toxin is potentially effective in the treatment of focal upper limb dystonia (66). Priori et al. found that immobilization of the forearm and hand for 4 to 5 weeks resulted in improved function (59). The hypothesis is that immobilization leads to plastic changes at the motor cortical level, reducing excitability. Unfortunately, treatments are not always successful and in more than half of the musicians with dystonia, it is career ending (67).

**Upper Back Pain**

Myofascial pain commonly occurs in the trapezius, levator scapulae, and periscapular muscles. As in the general population, it is caused by maintaining a static position, poor posture, and muscle fatigue. The tendency toward counterclockwise rotation of the guitarist’s spine facilitates right shoulder protraction and placement of the strumming hand. Widening of the strap can help decrease tension on the left trapezius and periscapular muscles (68). The physician may offer trigger point injections or make a referral to physical therapy for myofascial release and neuromuscular therapy, but restoring postural symmetry, strength, and flexibility is again the crux of treatment.
FIGURE 55-18. Focal dystonia in the left fourth and fifth digits. The affected fingers flex off the keys out of the musician’s control.

Low Back Pain
An important risk factor for muscular strain in the low back is maintenance of a position in which the weight of the spine is not well balanced (69). This is a likely scenario in many musicians. A discogenic etiology of pain is more likely with lumbar flexion. Musicians who play leaning forward, especially while seated, are likely more at risk for disc bulging and degenerative disk disease. Lifting and carrying heavy instruments is also a risk factor for muscular strain and discogenic pain. The clinician should recommend wheeled carts to transport heavier instruments, for instance, the harp, drums, and bass (42).

The practitioner and music instructor should focus on the posture of the musician while playing. Paull and Harrison, authors of “The Athletic Musician,” recommend a chair designed specifically for musicians with the front end sloped and the rear horizontal (70). This puts about one-third of the body weight through the feet, taking the weight off the spine (70). Similarly, a wedge cushion can be added to a standard chair. If using a standard chair, it is best to have the legs parallel to the ground and the back straight (Fig. 55-19A and B). It is helpful for guitarists to have a footstool so that he/she can rest the guitar on the elevated thigh. A string player with low back pain may improve symptoms by playing standing up, though the technique is different. A strap can be added to support the guitar while standing.

General Guidelines for the Treatment of the Injured Musician
It is widely accepted that athletes should undergo proper strengthening, conditioning, and stretching to maximize their performance. The musician is using smaller muscles to perform, but ignoring the basics of maximizing function and performance can have devastating results. The physician, therapist, and music instructor should reinforce the need to address muscular weakness, tightness, and asymmetries. Emphasizing a healthy lifestyle, incorporating diet, regular exercise, abstinence from tobacco, and avoiding excessive use of alcohol are also essential. In comparison to athletes, musicians are more likely to be deconditioned, and the clinician should inquire about exercise habits. The impact of physical problems may be exacerbated by the stress of performing, thus referral to a psychologist may be indicated.

Novice musicians should gradually increase their practice time and rest at appropriate intervals to avoid muscle fatigue. There are several suggested guidelines such as 5 minutes of rest for every 20 minutes of playing, 10 minutes of rest for 50 minutes of playing, and 10 to 15 minutes of rest for 30 minutes of playing (51). During rest, the musician should stretch and change positions. An abrupt increase in practice time is the most important risk factor for playing-related musculoskeletal disorders (71).
Communication with the music instructor and musician is essential to optimize performance. Together, they can assure proper instrument size, adjustment pads, benches, straps, and harnesses. Most of the problems discussed can be prevented with education regarding posture, technique, and instrument specific conditioning.

**Barriers to Treatment of the Performing Artist**

Barriers to care often include lack of medical insurance or financial coverage. Worker's compensation is usually only found in large dance companies or in the university setting for dancers. Similarly, worker's compensation and health insurance are often unobtainable for musicians. The median hourly wage for musicians and singers in 2007 was $19.92, and it is rare for them to have guaranteed work beyond 3 to 6 months (72). Physical therapy, occupational therapy, physician visits, medications, and durable medical equipment may be unaffordable.

It is necessary to prescribe rest at times, which in many cases, results in lost income. There is often pressure from parents or managers of dancers and musicians to continue through the discomfort. More seasoned musicians may choose to perform injured during their busy seasons and audition periods when they are in the most need of treatment and rest.

**CONCLUSION**

Musicians and dancers are a wonderfully unique group of people. Knowledge of the body mechanics required to play an instrument and dance, and the accompanying lifestyle is vital to optimize the outcome of musculoskeletal problems. A reference for current concepts in performing arts medicine is the Performing Arts Medicine Association (PAMA) at www.artsmed.org.

**ACKNOWLEDGMENTS**

We would like to acknowledge the following individuals for their help with this chapter: Konrad Weatherhogg, PT for assistance with rehabilitation techniques; Stephanie Levi, PT, for sharing her library of arts medicine articles; the music department at Charlotte Christian School—Don Humphries, Kelly Goley, and Jane Mendlik for being our musician models, and Laura Goodyear and Terry Efird, Charlotte Christian School for facilitating our photography session.

**REFERENCES**

24. Dobson R. Eight in 10 dancers have an injury each year, survey shows. BMJ. 2005;331:594.
SPECIAL POPULATIONS

The rehabilitation of children with physical impairments in some ways resembles and in many ways differs from that established for adults. It is a challenging combination of normal child care and the best of rehabilitation intervention strategies. Successful rehabilitation interventions require the understanding that a child is not merely a miniature adult, and that specific physiologic parameters exist that either complicate or allow unique intervention opportunities. This chapter reviews the scope of disabling disorders that occur in childhood, the specific differences between children and adults that relate to their special needs, and the basic principles of management of disabled children. The specific management of various childhood disorders will be found both within this chapter and within relevant sections of other chapters.

The various disabling disorders that occur in childhood may be characterized as congenital if they are acquired before birth and are not due to known external environmental factors during the birth or postbirth period. Otherwise, conditions are considered to be acquired. Congenital problems may be further specified by cause as either genetic or influenced by some extrinsic factor, even though the effect was expressed in the prenatal period (e.g., fetal alcohol syndrome). Acquired disabilities usually are the result of trauma, infection, or other causes.

Developmental disabilities in children exert a major impact not only on the child’s ability to function in the family and in society, but they also result in 1.5 more doctor visits and 3.5 more hospital days per year than for nondisabled children of similar age. In addition, these children typically lose twice the number of school days annually, and there is a 2.5-fold increase in the likelihood of repeating a grade in school when compared with the general population of children. This impact is much greater in children with multiple disabilities or with either cerebral palsy (CP), seizures, delays in growth and development, or emotional or behavioral problems (1).

**Physiologic Performance**

Children change with age and size in a number of physiologic parameters. Normal heart rate, respiratory rate, heat transfer behavior, and various chemical assessments all change as a function of age. For example, serum alkaline phosphatase levels may be elevated compared to adult reference values in an adolescent not because of the presence of occult heterotopic ossification but rather because of normal accelerated bone growth.

The question of enhanced neural plasticity in youth remains open. Conflicting data appear in the literature to support or reject this concept, but the clinical management implications are generally well accepted: the more treatment that is administered earlier after disease onset and at a younger age, the better the outcome seems to be.

A number of neurologically mediated reflex behaviors are age and development dependent. For example, the asymmetric tonic neck reflex (ATNR) (Fig. 56-1) is a normal behavior when elicited at 2 to 6 months of age, but may be distinctly abnormal when it is persistent and dominant many months later.

**Primitive Reflex Patterns**

Because they are more commonly observed in children with physical disabilities, the major primitive reflex patterns that may either interfere with or facilitate skilled motor actions should be well understood. The times of appearance and disappearance of these reflexes in the developmental sequence are summarized in Table 56-4.

The most basic proprioceptive patterns are the flexion and extension synergies of the upper limbs and lower limbs. In the upper extremity, the full flexor pattern is more commonly seen in CNS lesions, while in the lower limbs, extensor patterns are usually seen. Upper extremity flexor patterns include shoulder adduction, flexion, and internal rotation with elbow flexion, wrist pronation and flexion, and finger and thumb flexion (Fig. 56-2). The thumb is frequently tightly adducted and flexed into the palm. The extension posture of the lower limb includes hip adduction, extension, and internal rotation, along with knee extension, internal tibial rotation, and equinovarus foot posturing (Fig. 56-3). In both these postures, the fingers

---

**The Pediatric Patient: Differences to Consider**

Knowledge of the patterns of growth and development is key to understanding, anticipating, and managing the difficulties that disabled children experience. In the child’s early years, head circumference, weight, and height are important parameters to monitor. Standard tables of growth and development may be used to record and compare disabled children with the healthy population (Tables 56-1 to 56-3).
TABLE 56.3 Examples of Gross Motor Milestones

<table>
<thead>
<tr>
<th>Age</th>
<th>Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 mo</td>
<td>Head in midline</td>
</tr>
<tr>
<td>3 mo</td>
<td>Prone prop on extended elbows</td>
</tr>
<tr>
<td>4 mo</td>
<td>Rolls prone to supine</td>
</tr>
<tr>
<td>5 mo</td>
<td>Rolls supine to prone</td>
</tr>
<tr>
<td>Infants</td>
<td>“swimming”</td>
</tr>
<tr>
<td>Pivot circles in prone</td>
<td></td>
</tr>
<tr>
<td>6 mo</td>
<td>Sits with straight back</td>
</tr>
<tr>
<td>9 mo</td>
<td>Crawls on hands and knees (“creeps”)</td>
</tr>
<tr>
<td>Transitions into sit from four-point</td>
<td></td>
</tr>
<tr>
<td>10 mo</td>
<td>Pulls to stand through half-kneel cruises</td>
</tr>
<tr>
<td>12 mo</td>
<td>Independent ambulation</td>
</tr>
<tr>
<td>15 mo</td>
<td>Comes to stand independently</td>
</tr>
<tr>
<td>18–19 mo</td>
<td>Climbs into adult-size chair</td>
</tr>
<tr>
<td>21–24 mo</td>
<td>Up and down stairs with hands on rail</td>
</tr>
<tr>
<td>30 mo</td>
<td>Jumps clearing ground and lands on feet together</td>
</tr>
<tr>
<td>3 y</td>
<td>Pedals tricycle</td>
</tr>
<tr>
<td>Climbs up stairs alternating feet</td>
<td></td>
</tr>
<tr>
<td>True run</td>
<td></td>
</tr>
<tr>
<td>4 y</td>
<td>Hops, gallops (not true skipping)</td>
</tr>
<tr>
<td>Walks down stairs alternating feet</td>
<td></td>
</tr>
<tr>
<td>5 y</td>
<td>True skipping</td>
</tr>
</tbody>
</table>
tonic neck reflex (Fig. 56-4) describes midline effects of flexing and extending the head on the body. Flexion of the neck facilitates flexion in the upper limbs and extension of the lower limbs. Extension produces the opposite pattern.

The vestibular system mediates static postures and dynamic postural reactions. These are important movement patterns that facilitate the development of mobility skills. The most commonly seen static vestibular pattern is the tonic labyrinthine reflex. This pattern, facilitated by the supine position of the head, demonstrates lower extremity symmetrical extension with upper extremity shoulder abduction and external rotation. In the prone position, shoulder adduction and internal rotation are accompanied by lower extremity flexor posturing.

Cutaneously mediated reflex patterns include palmar and plantar grasp, elicited by tactile pressure over the respective sites. In young infants, the stepping response is an example of a kinesthetic reflex and is seen as a result of loading one limb (e.g., by vertical suspension) and stimulating the dorsum of the opposite foot, which is presumably a cutaneous influence as well. The positive supporting reaction occurs pathologically in older children when a loading of the suspended child’s plantar surface results in a symmetric extension pattern of the lower limbs.

**Child Development**

In any assessment of children, one cannot overemphasize the importance of measuring performance against age-expected norms. An understanding of all areas of normal development is essential if an abnormality is to be recognized (2,3).

---

**TABLE 56.4** Normal Acquisition and Regression of Primitive Reflex Behaviors

<table>
<thead>
<tr>
<th>Reflex</th>
<th>Age of Onset</th>
<th>Age Reflex Disappears</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moro</td>
<td>Birth</td>
<td>6 mo</td>
</tr>
<tr>
<td>Palmar grasp</td>
<td>Birth</td>
<td>6 mo</td>
</tr>
<tr>
<td>Plantar grasp</td>
<td>Birth</td>
<td>9–10 mo</td>
</tr>
<tr>
<td>Adductor spread of patellar reflex</td>
<td>Birth</td>
<td>7 mo</td>
</tr>
<tr>
<td>Tonic neck</td>
<td>2 mo</td>
<td>5 mo</td>
</tr>
<tr>
<td>Landau</td>
<td>3 mo</td>
<td>24 mo</td>
</tr>
<tr>
<td>Parachute response</td>
<td>8–9 mo</td>
<td>Persists</td>
</tr>
</tbody>
</table>

Psychosocial Development

Frequently, children may be infantilized by family and caretakers as a response to guilt or pity for the child. Conversely, some may be assumed to be more mature than is really the case, particularly if much time has been spent in the company of adults in hospitals and other health care settings.

Children who grow up with a physical limitation usually do not have a sense of loss of ability. It is usually around adolescence, when social sensitivity and maturity cause adaptation problems to surface. In the early years, it is important to help the child and family identify the child’s strengths and abilities so that, despite the disabilities, a sense of confidence and self-worth can be built (4).

COMMON CONCERNS IN TREATING DISABLED CHILDREN

Although the disabling disorders of childhood are widely varied, they share a number of common medical issues and potential as well as active problems. These will be discussed below as well as under specific clinical conditions. Routine well child care should not be overlooked, with careful attention paid to monitoring of growth parameters, nutrition, and especially immunizations, since the more seriously ill child may slip through the cracks of the well child system and fall behind in immunizations. Due to their particular susceptibility to infection, pneumococcal and influenza vaccinations should be given routinely.

Additional areas of medical concern will be reviewed under appropriate sections of specific clinical conditions as they differ somewhat in their presentation and treatment approach. These include spasticity and the various orthopedic conditions of contractures, scoliosis and hip dysplasias (5), bowel and bladder incontinence, decubitus prevention, and gastroesophageal reflux.

Functional issues are common to many diseases and disabilities. Specific management strategies vary with the particular disease, but the determination of goals and objectives is largely independent of the specific disease.
The desire and ability to communicate is the highest priority in managing a child with a disability. This allows the child to interact with his environment in spite of what may be profound physical impairment. Communication need not be vocal in nature to be effective. Gestures and signs may be used if motor control allows. As simple a movement as control of eye gaze may allow use of assistive devices as primitive as object boards or as sophisticated as computer-assisted voice output devices. What is important is not the technology but the identification of the potential for communication and the development of a system that allows a child to live up to his or her potential.

Second to communication in priority is the child’s ability to control his or her motion in three-dimensional space. Frequently, children are provided with passive transportation instead of the technical ability to achieve independent mobility. From a developmental perspective, newborns start experiencing self-directed movement in space in their cribs, and a child who is crawling and creeping is acquiring substantial knowledge about the world by navigating within it. It is important not only to provide convenient transportation for parents but also to give the child control over his or her own mobility. As in communication, the devices used are dictated by whatever movement control the child demonstrates. Early prone mobility may be facilitated by wheeled prone boards, while once the child can be positioned in sitting, this can be replaced by caster carts and later by appropriately designed wheelchairs. For the child who is unable to propel a manual wheelchair, there are powered vehicles. Limited research has supported clinical observations regarding the age at which a child can control a powered wheelchair (6). By 18 months, the child may be able to acquire the eye-hand coordination necessary to control a wheelchair with a joystick. More sophisticated steering controls such as wafer, head or mouth switches require more advanced developmental levels. For the child with ambulation potential, lower extremity orthotics provide optimum positioning of pivots for safe gait, while upper extremity assistive devices may provide necessary stability.

The more severely impaired children frequently need adapted seating and positioning systems to achieve a number of developmental goals. General goals include normalization of tone, symmetric positioning, and improved trunk alignment. This often facilitates use of the head or upper limbs for communication, self-care, and educational activities. Side-lying devices may facilitate function and maintain trunk flexibility for floor and bed-level activities. Adapted seats are useful for spasticity management and the facilitation of upright activities. Car seating that is safe from both a postural and a crash safety perspective gives secure travel capacity to the child and family. Standing devices allow vertical alignment, weight bearing, and experiences in the upright posture although the effectiveness in reducing osteoporosis has not been shown.

**Self-Care Activities**

Appropriate goal setting is most important. Infantilization will often limit the achievement of the child’s potential. However, unrealistic goal setting should also be avoided. Few methods for the quantified assessment of functional abilities in children exist.

**Education**

Whereas vocational rehabilitation is an important consideration for the adult with a disability, special education is important to the child with a disability. The laws and services for children with disabilities vary with state and local school systems, but some common features are present. According to federal law (PL94-142 and PL99-457), the goals of special education are to provide “free and appropriate public education” in the “least restrictive environment” for a child. Part H of PL 99-457 mandates that participating states also provide early intervention services for children with developmental disabilities from birth to 2 years of age. Services that may variably be incorporated into this type of educational program include special education, physical therapy, occupational therapy, speech and language therapy, adaptive physical education, psychological and social work services, and nursing services. In each case, these services become school system responsibilities in that they are necessary for the child to participate in an individualized education program, a specific educational plan with goals, objectives, definition of services, and time frames.

The role of health care professionals in the special education system is an advisory and participatory one. Programs developed within a school setting should be consistent with those established out of school, in both the home and other therapy settings. School programs should not let therapeutic goals obscure their educational objectives. Programs that effectively integrate medical rehabilitation needs with those of education are the most effective (7).

**Play**

Children with disabilities frequently need special assistance in achieving the ability to play successfully. Recreational therapy, music therapy, art therapy, play therapy, and other interventions may be helpful in allowing a child to find mechanisms to express himself or herself and to experiment with future skills and roles.

Adapted toys and games are useful. Battery adapters that allow external switch control of any electric or electronic device may give the child with a severe disability the option to play with age-appropriate toys while possessing only limited physical skill.

Children also may need special assistance in experiencing group play. Participation in nursery school programs, play groups, and other endeavors will allow the child to experience play with other children both with and without disabilities. Parental counseling and resource identification may be necessary to facilitate these activities.

**Social Skills**

Children with disabilities are often found to be deficient in adaptive social skills because of a variety of factors, including
limited normal childhood experiences and relatively more intensive involvement with the health care community. Efforts may need to be taken to assist a child and family to identify specific behavioral issues and find methods to overcome them. The concept of a child having acquired learned helplessness is useful as a perspective for parents and health care professionals in dealing with these issues.

As these children reach adulthood, they frequently remain in the home setting with their parents long after most young adults have elected to live independently. They also demonstrate a much higher rate of unemployment than their nondisabled counterparts (8).

Parenting Skills
Just as normal child parenting is a challenging experience, so too is parenting a child with a disability. It is further complicated by the challenges of health and social impairments experienced by the child and family. Many parents need assistance and guidance in coping with what are really normal parenting issues in their care for a child with a disability.

Common problems include discipline maintenance and difficulty in setting appropriate levels of expectation of responsible behavior. It may be helpful to counsel the parents on how to distinguish the special limits and expectations that are appropriate for the disability from normal parenting issues, such as control and authority challenging. Introducing families to parents of other children with disabilities is also a very positive strategy.

Sexuality
Managing the emerging sexuality of a child with a disability requires knowledge, an openness and willingness to discuss, and anticipatory strategies. Many of the early needs of these children are simply for accurate and age-appropriate information about sex and reproduction in general, as well as the child’s specific abilities or limitations based on the disability. Frequently, knowledge about the child’s sexual and reproductive potential is also needed by the parents.

Education and counseling for the adolescent child are often necessary. Children may express their underlying sexual concerns through other behaviors, including social withdrawal and depression. A high index of suspicion of the need for sexual education and counseling should be maintained by the involved health care professional.

Independent Living
A long-term perspective on the child’s potential to live independently should be adopted from an early age. Realistic goal setting is the essential first step for any long-term rehabilitation program and for attaining independence. It is frequently possible to distinguish at an early age the child who will need some type of supported living situation in the long term. Helping the family and school system to identify these expectations early on will facilitate appropriate school programming and long-term planning.

It is unfortunately all too frequent that one encounters a seriously developmentally disabled adult 30 to 40 years of age who lives with his or her parents until the parents become infirm or die. Many times there has been inadequate planning for legal, estate, and practical matters that suddenly become crises. The solution to these types of problems is prevention by ensuring that they are anticipated long before they become a reality.

EVALUATION OF THE CHILD WITH A DISABILITY
There are several objectives to be achieved in the evaluation of a child with a disability.

First, one should determine the type and etiology of the disability. This is based on a careful history and physical examination, including careful assessment of all areas of development from birth to present. Review of previous medical records and evaluations by other members of the rehabilitation team may be helpful. Next assess the child’s potential to benefit from rehabilitation services. In addition to biologic potential, outcome may be affected by the family’s adjustment to their child’s disability and their ability to support the therapy program.

Evaluation by the Physician
It is essential that the physician establish rapport with both the child and family members present. Young children are typically frightened by any encounter with health care personnel. The most critical factor in obtaining useful data from the initial history and physical examination of a child is the development of a cordial relationship with the child and family as early and rapidly as possible during the interview process. The environment in which the examination is conducted must be as nonthreatening as the demeanor of the physician. Toys and children’s artwork are useful tools in creating such an atmosphere. The physician should present as much of a friendly image to the child as possible. Attention to seemingly unimportant factors such as not wearing a white coat in the child’s presence and sitting while examining or interviewing the child can have a major impact on the development of a cordial relationship. Other helpful measures include smiling frequently and assuring the child that he or she is not in your office to receive shots from you. If size permits, use the parent’s lap as the “examining table.” In the history, a careful review of the pregnancy and birth history as well as the family history is essential. In children with acquired disabilities, detailed information about the illness or injury that produced the disability should be obtained. A detailed developmental history should be elicited. Comparing the development of the child with a disability to other healthy children in the family often facilitates this process. The administration of standard developmental assessment tests adds valuable information in identifying the specific areas of developmental delay in all functional areas.

The initial physical examination should be brief with as much of the examination as possible conducted by observation of the child. Much of this can be accomplished while taking the history. Everything the child does is a hint to their developmental levels in all areas of performance. Actual physical contact
can be kept to a minimum. Considerable data regarding basic motor skills such as head and trunk control, reciprocal creeping, standing balance, and gait patterns can be obtained in this manner. Quality of movement can be easily assessed and may actually be more reliable than when performed upon request (i.e., the “doctor walk” of the mildly affected CP child). Toy play may allow observation of fine motor skills as well as the ability for imaginative play and social interaction.

Assessment of the developmental reflex profile should be performed. Manual muscle testing is usually not possible in a child less than 5 years old. Observation of an infant’s spontaneous movements as well as of the child during play provides the examiner with information about antigravity muscle strength. Muscle tone should be assessed not only for the degree of tone (i.e., low, normal, or high), but also for the pattern of tone and what activities may trigger abnormal tone. Significant fluctuations in muscle tone should be noted along with a description of factors that seem to cause the fluctuation. For example, in children with spasticity, a marked increase in tone of a given muscle may be produced by a sudden stretch of that muscle or its tendon. In children with athetoid CP, tone fluctuations may be produced by startling the child with an unexpected touch or loud noise.

Deviations from normal range of motion in any of the limbs should be recorded. In cases where limitation of range of motion exists, the family should be questioned regarding their observations as to whether such tightness is static or progressive and approximately how long it has been present. Monitoring of height, weight, and head circumference is an important routine assessment. Use of the normative charts for comparison purposes is appropriate. Significant deviations in the pattern of growth should be noted and explained. Difficulties in growth may be the result of chronic illness, nutritional deficiency, or chronic gastroesophageal (GE) reflux, along with a host of other specific problems.

In children of school age and above, the sensory examination including vision and hearing can usually be completed successfully. In younger children and particularly in infants, the sensory examination will primarily consist of the child’s response to noxious stimuli such as a pinprick. The response is usually crying and/or withdrawal.

Sensorineural testing in children under 4 years of age is usually limited to simple screening examinations due to the inability of the child to provide consistent responses to testing. The child’s ability to follow objects with his or her eyes combined with observation of spontaneous eye movements or their absence can assist the physician in determining whether more definitive testing is needed. The child’s ability to respond or localize sounds can help to make similar determinations with regard to hearing. Vision and hearing testing in infants and younger children suspected of abnormality in these functions will require the use of sophisticated measurement devices by physicians specializing in those areas. Auditory and visual brain stem response testing has proven to be very useful in assessing these functions in children too young to respond to standard screening tests. Older children can be screened for visual and auditory function using the same testing techniques used for adults.

Instruments used in the neurologic examination such as reflex hammers and tuning forks should be shown to the child before they are used. It is often helpful to allow the child to play with them before the testing.

The data obtained at the initial evaluation may be synthesized using various tools described below and used as a basis for future comparisons during the treatment process as the child grows. Commonly used screening tests for developmental assessment include the Bailey Scales of Infant Development, designed for children from birth to 30 months and the Denver Developmental Screening Test. Both are easy to perform but are relatively insensitive to increments of developmental progress that may occur in children with severe disabilities.

Quantitative analysis of motor performance of children is accomplished by several strategies. First is the measurement of physical parameters, such as range of motion and strength, and physiologic parameters, such as heart rate and respiratory rate. Timed trials of specific activities, such as the Jebsen Taylor Hand Function test, may be useful as norm-referenced comparisons or sequential performance reassessments.

Quantitative descriptions of the functional activities of children with disabilities are essential for monitoring and planning rehabilitation programs. The few pediatric tools currently available include the WeeFIM (9), the gross motor functional measure (GMFM) (10), and the Pediatric Evaluation of Disability Inventory (PEDI) (11). In addition, the analytic tool for children with spina bifida described by Sousa and colleagues and the generally useful Tufts Assessment of Motor Performance (TAMP) may be used. The TAMP provides a method for structured quantitative description of developmentally oriented activities that are commonly performed by children with serious disabilities.

The Rehabilitation Plan
Once the initial evaluations are completed, rehabilitation goals and a comprehensive management program should be developed at staffing conferences, involving as many members of the rehab team as possible (PT, OT, ST, teacher, social worker, psychologist, nutritionist, nurse, etc. and especially the child and family). Often, however, services are delivered at distant sites and face-to-face conferencing may not be practical. Written, telephone, and even e-mail communication among treating professionals helps coordinate the therapeutic approach and avoids sending “mixed messages.” The involvement of the family in daily carryover of treatment strategies is vital.

Prediction of Outcome
One of the major responsibilities of the health care professional is the establishment of medical and functional prognoses for a child with a disability. From these estimates of outcome, the specific objectives and plans of management can be formulated that will guide the daily activities of the child and family. It is therefore most important to both accurately predict prognosis and at the same time be cognizant of the limits of predictability for any individual child. In general, it is important to emphasize the importance or cognition and communication over
ambulation. Discussion of the child’s strengths as well as weaknesses is vital. Even for the most severely impaired child, comfort and a stimulating environment should be the minimum achievable goals. The correct approach is to admit uncertainty when it exists, make cautiously optimistic predictions, and move forward with goals that at least provide for comfort and care for both the child and family.

**SPECIFIC CLINICAL CONDITIONS**

**Cerebral Palsy**

CP is a disorder of movement control and posture resulting from a nonprogressive lesion to an immature brain, occurring in utero, near the time of delivery or within the first 3 years of life. Although more than 50% of cases of CP have no known etiology, several factors, occurring at different points in time, are thought to be risk factors for future CP.

In the prenatal period, congenital infections (TORCH—Toxoplasmosis, Rubella, Cytomegalovirus, Herpes, and others), often clinically unrecognized in the mother, can cause a spectrum of involvement, from severe microcephaly, seizures and spastic quadriplegia, to mild diplegia. Gestational toxins include iodine that may lead to diplegia and organic mercury intoxication while may lead to quadriplegia. Intrauterine subdural hemorrhage may cause hemiplegia (12). During the perinatal period, complications of prematurity include birth weight under 800 g, Grade III and IV intraventricular hemorrhage, prolonged seizures, and an APGAR score of less than 3 at 20 minutes. In full-term gestations, abruptio placenta, placenta previa, nuchal cord, or meconium aspiration can result in neonatal asphyxia, although recent evidence suggests asphyxia may be secondary to underlying prenatal malformations that also result in CP. Finally, hyperbilirubinemia secondary to Rh disease, G6PD, or ABO incompatibility may result in kernicterus, with deposition of bilirubin in cranial nerve nuclei and basal ganglia resulting in athetoid (dyskinetic) CP.

During the postnatal period, bacterial or viral sepsis or meningitis, especially within the first 6 months, can cause residual motor impairments. Traumatic brain injury can be due to child abuse (“shaken baby” syndrome with subdural hematoma and retinal hemorrhage), fall from heights, and motor vehicle accidents. Near-drowning causes hypoxic ischemic encephalopathy. Stroke syndromes with hemiplegia may be caused by traumatic delivery as well as by cyanotic congenital heart disease (e.g., Tetralogy of Fallot), clotting disorders, and ruptured arteriovenous malformations (13). Heavy metal and organophosphate ingestion can cause quadriplegia.

Clinically, gross motor delay is seen in all such patients, with lack of sitting after 6 months being the most common initially recognized deficit. Poor head control may be recognized earlier in the more involved child while delayed ambulation to after 16 to 18 months is seen in the more mildly involved. Motor delay, however, demonstrates poor diagnostic accuracy since the overwhelming majority of children showing isolated gross motor delay will eventually develop normally.

Abnormal motor characteristics (quality of movement) are often mistaken for “early” milestones, but represent influences of abnormal tone on the child’s movement capabilities; for example, “rolling” at 2 months by opisthotonic posturing (Fig. 56-5), “handedness” at less than 1 year in hemiplegics, “walking” at 4 months by reflex steppage. Additional abnormal movement patterns include w-sitting (Fig. 56-6) and sacral
sitting with posterior pelvic tilt due to hamstring spasticity, bunny hop crawling, and coming to stand through symmetric extension of the lower limbs due to poor pelvic dissociation in children with diplegia, and buttock hitching in hemiplegics. As the child achieves ambulation, abnormal gait patterns (crouch gait, “jump” stance, etc.) are seen.

Alterations in tone are seen both chronologically and positionally. Early hypotonia, seen universally, gives way to later hypertonia. Hypotonia or normal tone seen when the child is supine gives way to hypertonia when posturally stressed (such as ventral or vertical suspension).

Reflex abnormalities include muscle stretch reflexes as well as primitive reflex abnormalities. Infants, however, tend to be relatively hyperreflexic normally, limiting the usefulness of this finding, unless associated with either asymmetry (suggestive of a hemiplegia) or excessive spread of the reflexogenic zone (seen in severe spasticity). Hypo or areflexia is useful however in the differential diagnosis of infantile hypotonia (early CP vs. neuromuscular disease).

Primitive reflexes are body postures dictated by head position and are seen at birth, become more elicitable in the first 2 to 3 months, and normally start to fade by 4 to 6 months. Their presence after 6 months is deemed abnormal. However, the ability to obtain the reflex on every attempt or the infant’s persistence in the reflex posture for more than 30 seconds is considered abnormal at any age. Commonly elicited reflexes include the asymmetric and symmetric tonic neck reflexes, tonic labyrinthine reflex, and moro reflex. Abnormal postures at rest are often manifestations of primitive reflex patterns. Examples include fisting with cortical thumb position, extensor thrusting of the lower limbs with scissoring, and asymmetric tonic neck posturing.

The differential diagnosis of CP includes syndromes of early hypotonia (14) and developmental delay but without marked facial dysmorphisms. Two common examples include the Prader-Willi syndrome of low birth weight, normal neonatal length, hypogonadism, small hands and feet with tapered digits and a deletion of the long arm of chromosome 15 and Sotos syndrome of large birth weight, macrosomia with megencephaly, large hands and feet, prominent forehead, and advanced bone age. The Rett syndrome of normal early development leading to loss of speech by 9 to 18 months, stereotypic hand movements, sighing respirations, and spasticity is seen in females only and is associated with a genetic abnormality on the long arm of the X-chromosome (15).

Metabolic disorders include endocrine (thyroid, infant of diabetic mother), amino acid (PKU), and storage disorders (mucopolysaccharidoses).

Congenital neuromuscular diseases presenting with hypotonia, hypo or areflexia, and delayed motor development can be confused with the early hypotonic phase of CP. Examples include infantile Werdnig-Hoffman disease (WHD), congenital muscular dystrophy (CMD), congenital myotonic dystrophy (CMdD), hereditary motor/sensory neuropathies, etc. These will be discussed in greater depth later in this chapter.

Progressive central nervous system diseases may mimic CP until the appearance of neurologic deterioration. Examples include metachromatic leukodystrophy, olivopontocerebellar degeneration, Friedreich's ataxia, and ataxia telangiectasia. HIV encephalopathy, however, is the most common intrauterine-acquired neurodegenerative disease.

A thorough history for predisposing factors as well as possible neurologic deterioration should be obtained. Physical examination should stress recognition of dysmorphic features (syndromes), muscle weakness, hyporeflexia or hyperreflexia, the presence of abnormal movement patterns, and the abnormal retention of primitive reflex postures.

Laboratory examination should include blood and/or urine screens for amino acids, organic acids, serum lysosomal hydroxylase enzyme battery, thyroid function, and creatine kinase (CK) as indicated. Cultured skin fibroblasts for metabolic assay may be necessary if serum or wbc testing is unrevealing.

If diseases of nerve or muscles are being considered, electrodiagnosis allows differentiation of neuropathy from myopathy from normal with 85% to 90% accuracy for neuropathy but with only 40% accuracy for myopathy. Muscle and/or nerve biopsy as well as DNA analysis may be indicated for confirmation. MRI and cranial ultrasound in the perinatal period demonstrating cystic periventricular leukomalacia as well as ventricular enlargement may be predictive of disabling CP (16,17). CNS imaging studies are also useful in ruling out AVM or tumor in hemiplegia of unclear etiology.

Classification of CP is by tonal type and part of body involved. Spastic CP comprises 70% to 85% of cases with further subdivision based on the topographic distribution of the spasticity. Hemiplegia, diplegia, and quadriplegia account for about 90% of spastic CP’s with each contributing equally. Triplegia and monoplegia account for the other 10%.

Hemiplegia may result from focal perinatal injury and has the highest incidence of CT/MRI abnormalities, usually in the distribution of the middle cerebral artery. The most common presentation is failure to use the involved hand, although hitching on the buttocks rather than crawling on hands and knees is also seen. The upper limb is invariably more involved than the lower limb. Sensory deficits are difficult to evaluate in children under 4 years old and are likely underappreciated (18). Ambulation is usually achieved by 2 years unless severe mental retardation is an associated finding. Growth retardation of the affected side and an associated parietal lobe syndrome are seen in 50%. Initial seizures may occur as late as 5 years old. Long-term disability is usually more cosmetic than functional.

Diplegia represents the most common type of CP seen in premature babies. There is disproportionate involvement of the lower limbs, although upper limb abnormalities, manifested as motor perceptual dysfunction, are very common. Most children with diplegia eventually ambulate, although both lower limb orthoses and upper limb devices may be required and distances covered varies.
Triplegia presents with relatively symmetric involvement of the lower limbs and marked asymmetric involvement of the upper limbs.

Quadriplegics, also termed total body involved have the highest incidence of significant cognitive disability with 25% severely involved, 50% moderately involved, and 25% mildly involved. Lower limbs are usually more involved than upper limbs with asymmetries not unusual. A brief period of hypotonia giving way to early spasticity is usually a poor prognostic sign for independent mobility.

Dyskinetic (athetoid/dystonic) CP accounts for 5% to 8% of cases. In the past, most cases were associated with kernicterus due to Rh disease. Athetosis, dystarthis, sensorineural deafness, and paralysis of upward gaze were the usual signs, while intelligence was often normal, since cerebral cortex was spared. Today, dyskinetic CP is more commonly seen as part of the picture of diffuse hypoxia and may be associated with spasticity, seizures, and retardation (19). The period of hypotonia is usually longer, from 18 to 36 months, before the abnormal movement disorder appears.

Rarer types include atonic (hypotonic) and ataxic CP. The latter is usually associated with hypotonia. If spasticity is present, progressive CNS diseases such as Friedrich’s ataxia must be ruled out.

**Prognosis for Ambulation in CP**

“Will my child walk?” is usually the question asked most frequently by parents of a newly diagnosed child with CP. In the discussion that follows, one must clarify not only distances involved (household vs. community) but also the quality of the gait and need for both orthoses and/or upper limb assistive devices. Several clinical factors are relevant. The clinical type of CP is important. All children with hemiplegia will walk as well as those with ataxia, while those with atonia usually will not. Those with quadriplegics, diplegia, and dyskinesia vary. Molnar has shown that if independent sitting occurs by 2 years, prognosis for ambulation is good (20). Badell reported that ability to crawl on hands and knees by 1.5 to 2.5 years was a good prognostic sign (21). Persistence of three or more primitive reflexes at 18 to 24 months is a poor prognostic sign (22). Recently, the ability to transition from supine to prone by 18 months was shown to be a predictor of independent ambulation in spastic diplegics (23). Marginal ambulators during childhood may lose functional ambulation during or after adolescence due to progressive orthopedic deformity, insufficient muscle power and/or capacity to accommodate increased height and weight and social/emotional problems. Combinations of independent ambulation and use of wheeled mobility for greater distances should be introduced. Power mobility as a means of independent function should not be delayed in favor of waiting walking as children as young as 3 years old can learn to “drive” (24). Parental negative feelings are rapidly replaced by positive ones once independent movement and its effect on the child are observed (25).

**Associated Disabilities**

Mental retardation of moderate to severe degree is seen in one third, of mild degree in another one third, with the remaining one third showing normal intelligence. In the latter group, fewer than expected fall into the superior range. Those with pure dyskinesia as a rule are the brightest, while those with atonia and quadriplegia are most involved. Retardation is usually mild in those with diplegia and hemiplegia and may be confused with learning disabilities. Preservation of receptive language skills is a better indicator of good cognitive function than expressive language. Hearing loss should be ruled out prior to making judgments regarding receptive skills. Early educational intervention is a reasonable, if unproven, “treatment” and has been mandated by federal law (Public Law 99-457, 1986).

Seizures are seen in up to 35% to 40% of children with CP (26) and are most common in postnatal CP, and about 60% to 70% of those with hemiplegia and quadriplegia. Seizures are seen in only 25% to 33% of children with dyskinesia and diplegia. Imaging may be considered to rule out structural lesions and EEG characterization is useful. Anticonvulsant management is similar to those with idiopathic epilepsy, although the risk of breakthrough seizures when medications are discontinued after being seizure free for 2 years is higher in hemiplegic CP (approximately 60%) than in non-CP children (40%). Children with diplegia have a lower rate (about 14%) of breakthrough seizures (27).

Abnormal vision is seen in 50% of children with CP. Muscle imbalances cause esotropia, exotropia, or hyperopia and are most frequent in diplegia and quadriplegia. Secondary amblyopia can occur. Homonymous hemianopsia can be seen in hemiplegia. Paralysis of upward gaze is seen in pure dyskinesia. Nystagmus is seen in ataxia. Defective tracking can be seen in all types of CP. Most importantly, the possibility of refractory errors as a cause of poor vision should not be overlooked. This is twice as common in spastic CP as in dyskinesia. Surgical correction of muscle imbalances is mostly cosmetic, while optometric “exercises” are controversial, with little evidence of lasting improvement.

Hearing loss may be conductive due to abnormal Eustachian tube function as a result of palatal distortion or sensorineural due to aminoglycoside treatment during the neonatorum. Early assessment is possible with ABRs, although this modality assesses high frequencies only and is subject to false positives in the first 6 months. Amplification for those with bilateral hearing loss is justified.

Dental problems include malocclusions and enamel dysplasias secondary to palatal distortions and abnormal oromotor reflexes (suckle/swallow, tonic bite, rooting, etc.). Inappropriate administration of tetracycline in the newborn nursery is an avoidable cause of enamel dysplasia. CP children are also at increased risk for caries due to poor handling of secretions and food, as well as due to chronic drooling. This can be a source of pain, leading to increased agitation, worsened spasticity, and greater difficulty in parental handling. The adverse social
implications for the child are obvious. Early dental treatment before 3 years old is therefore vital. Anticholinergic treatment, including scopolamine patches for drooling, holds promise with several studies showing consistent long-term improvement. Potential side effects include constipation, urinary retention, and mood changes, all of which reverse with discontinuation medication. Most recently, Botulinum toxin injection into the parotid glands have been tried with effects lasting up to 4 months reported (28).

Increased risk of respiratory infections is due to both extrinsic (abnormally high tone and poor control over chest muscles leading to poor sigh and cough mechanisms) and intrinsic (bronchopulmonary dysplasia) reasons. Early antibiotic treatment is justified.

Malnutrition is more common in moderately to severely involved children but can also affect the more mildly involved (29). Abnormal oral motor function affects ability to handle food with a resulting increased risk of aspiration. The effect of poor feeding skills on survival to adulthood is profound (30). Disordered GI motility with reflux proximally and poor transit time distally is seen. Video swallow studies with multiple temperatures and textures of food may define abnormal mechanisms, but the question of relevance to every day feeding situations has been raised. For many children, gastrosopy tube placement not only improves nutrition but resolves the tension associated with feeding, thereby improving social emotional conditions.

A “neurogenic” bladder can occur in the face of normal sensation and takes the form of either a disinhibited bladder or a spastic dysynergic bladder due to external sphincter spasticity coupled with uncontrolled spastic detrusor contraction. Abnormal urodynamic studies have been seen in more than 85% of symptomatic patients. Similarly, spasticity of the external anal sphincter can lead to difficult initiation of bowel evacuation and secondary constipation. Cutaneous stimulation and sacral cleft massage may be useful. Ditropan for the spastic detrusor may also be helpful (31).

Behavior disorders include true emotional lability as part of an organic pseudobulbar palsy consisting of dysarthria, drooling, and poor chewing. Attention deficit disorder with hyperactivity may be seen. Poor peer acceptance leading to a negative self-image, school phobia, depression, and anger may be exacerbated during normal periods of transition; for example, preschool to kindergarten, and early adolescence. The more mildly physically involved child may actually have more social and peer group difficulties than the severely involved child with wheelchair-bound quadriplegia—spastic hand posturing, lurching “effeminate” gait, etc. Limited participation in school results from both physical impairment and cognitive-behavioral abnormalities (32). Psychosocial support including psychotherapy is indicated as soon as a problem is identified.

**Therapy in the Management of CP**

The multidisciplinary management of the motor handicap in CP maximizes potential, but does not “cure” brain damage. “Physical” therapy consists of a hands-on approach by physical, occupational, and speech therapists to improve gross motor, fine motor, and oromotor function. Developed in the early 1960s by the Bobaths, neurodevelopmental therapy (NDT) (33) is the most popular “system” of therapy in use today. It emphasizes hands-on facilitation of movement and positioning to “normalize” tone and reduce the influence of abnormal postures (including primitive reflexes). This can be further assisted by the use of positioning aids, such as sidelyers, adaptive seat inserts, and prone standers.

Conductive education (34), developed at the Peto Institute in Hungary, has received recent support as a system capable of “treating” many children in a non-staff-intensive setting. “Educators” or “conductors” encourage spontaneous achievement of motor activity without regard for abnormal quality of movement. The “end” is more important than the “means.” Additional therapy systems include craniosacral manipulation, hyperbaric oxygen (35), and various electrical stimulation systems, both functional electrical stimulation (FES) (36) and threshold electrical stimulation (TES) (37,38). No system of therapy has been objectively shown to result in a greater degree of improved motor function than that expected with a program of passive range of motion only (39). The role of progressive resistive exercises (PREs) for strengthening has been questioned with the previously concern of worsened spasticity not seen. However, while strength can be increased, the extent of functional improvement is not clear (40).

Bracing goals include reduction of abnormal tone, avoidance of deformity, and facilitation of normal movement patterns. Lightweight plastics are widely used and include aquaplast (low temperature, softer, direct fabrication possible) and polypropylene (high temperature, more rigid, fabricated from cast mold). Where metal components are necessary, lighter weight alloys have replaced steel.

In the lower limbs, inframalleolar and supramalleolar orthoses (SMOs) are used primarily to control foot and talocalcaneal alignment with little direct tibiotalar control (41). Ankle foot orthoses (AFOs) add direct tibiotalar control and indirect control of the knee. Setting the ankle in neutral to slight dorsiflexion promotes heel strike and limits knee recurvatum. Articulated AFOs allow setting of a plantarflexion stop while allowing free dorsiflexion and promote active use of anterior tibialis as well as passive dorsiflexion necessary in stair climbing, crouching, and half-kneeling. The addition of twisters (either rod or elastic) attached to a pelvic belt, adds an element of control over hip internal or external rotation. Floor reaction AFOs, with the ankle set at neutral dorsiflexion and molded anteriorly to just below the patella limit crouch gait secondary to hamstring spasticity. Posterior leaf spring orthoses (PLSAFOs) are solid ankle orthoses, thinned posteriorly to simulate push-off at the end of stance phase, following passive dorsiflexion in early to midstance. Knee-ankle-foot orthoses (KAFOs) add direct control over knee flexion and extension as well as varus and valgus, but add bulk and weight. Similarly, hip-knee-ankle-foot orthoses (HKAFOs) add control over hip
position. Neither of the latter two braces significantly improves gait capability, but they do prevent deformity and may facilitate standing. Use of ratchet joints may allow for gradual reduction of flexion contractures over lengthy time periods.

The primary purpose of upper limb orthoses is to prevent fixed deformity. Minimal improvement in hand function is seen with use of upper limb orthotics. The cortical thumb loop orthosis, a simple fabric loop providing pressure into the thenar eminence, promotes abduction and extension of the thumb and facilitates thumb-opposed grasp. Wrist and/or elbow extension splints can be used during the day to extend reach, or at night to prevent flexion deformities. More recently, restrictive casts and slings have been used on the functional upper limbs of children with hemiplegia in constraint-induced therapy or "forced use" programs with early encouraging results (42, 43).

Medications for Spasticity

By limiting the effects of spasticity, deformity can be prevented, nursing care improved, bracing better tolerated, and function enhanced. Most commonly used medications include Baclofen (Lioresal), Diazepam (Valium), Dantrolene Sodium (Dantrium), Clonidine and Zanaflex (Tizanidine).

Baclofen acts at GABA-B receptors in the spinal cord. Dosage starts at 2.5 to 5 mg bid, increasing 2.5 mg per dose every 3 days. Effects can be expected by 1 mg/kg/day, although doses as high as 2 mg/kg/day may be well tolerated. Side effects include confusion, depression, weakness, GI upset, and lowered seizure threshold.

Diazepam acts at both the brainstem reticular activating system and the spinal cord. Dosage starts at 1 to 2 mg bid and is titrated up to the desired effect. Side effects of lethargy, urinary retention, and dependence may limit usefulness. Prompt withdrawal may cause seizures. Dantrium acts at the level of intra- and extrafusal muscle fibers resulting in decreased release of calcium from sarcoplasmic reticulum. Dosage starts at 0.5 mg/kg/day. Weakness, fatigue, lethargy, and diarrhea are potential side effects. Liver function tests and blood counts are monitored during use.

Clonidine, an $\alpha_2$-agonist originally used to treat hypertension in autonomic dysreflexia, was found serendipitously to reduce spasticity. Dosage starts at 0.05 to 0.1 mg bid, increasing as needed. Side effects include temporary sedation as well as hypotension. Delivery by transdermal patches in 0.1, 0.2, or 0.3 mg doses, each lasting 7 days, adds a measure of convenience. More recently, Zanaflex starting with 1 to 2 mg/day, has replaced oral Clonidine.

While medications have been shown to be capable of reducing tone, the doses required may cause significant side effects (most commonly sedative or GI) so as to limit the amount of functional improvement realized.

Intrathecal Baclofen (ITB) Infusion

Intrathecal infusion of Baclofen offers the advantage of bypassing the poor ability of Baclofen to cross the blood-brain barrier, leading to CSF concentrations 30 times higher than when given by mouth. The incidence of side effects is also reduced. The intrathecal catheter is inserted by lumbar puncture and positioned at the T10 spinal level. The dose is delivered by a programmable pump implanted in a subcutaneous abdominal pouch and varies from 30 to 800 $\mu$g/day. Changes in dose as well as mode of delivery (continuous infusion, bolus, or variable rate) can be made by transcutaneous telemetry. While tone reduction is most significant in the lower limbs, effects on the upper limbs and trunk are also seen by upward migration of infused Baclofen. Functional changes in gait and upper limb use can be seen (44). Potential complications are multiple and include CSF seroma, leaks, catheter kinking or dislodgement, infection, and pump failure. However, the rate for any single complication is quite low (45–47). Battery life is currently up to 7 years, but replacement requires a surgical procedure. The cost of the pump and its placement is about $25,000, with refills every 2 to 3 months costing about $750.

Injection Therapy in the Management of Spasticity

Previously used in various movement disorders (Blepharospasm, Spasmodic Torticollis, Spasmodic Dysphonia), Botulinum Neurotoxin Type A (Btx-A) has recently been used in the treatment of spasticity in children. Btx-A acts by irreversibly blocking presynaptic release of acetylcholine (ACh) at neuromuscular junctions (NMJs). Given as an IM injection, the onset of the clinical effect is delayed by 24 to 72 hours due to a complex process of neuronal uptake by endocytosis and subsequent binding to presynaptic receptor sites in the axon terminal. The toxin blocks binding and fusion of vesicles with the presynaptic membrane and interrupts ACh release. Clinical effects peak at 2 to 6 weeks and last about 3 to 6 months with reinnervation by terminal axon sprouting. Large muscles may require several injection sites. Common clinical practice allows injection of 12 to 14 U/kg divided among all injected muscles.

Advantages over nerve block include no need for anesthesia, no sensory side effects, and no apparent tolerance to repeated injections. Loss of effect due to the development of antibodies to Btx-A has been rarely described in adults receiving higher doses (>400 U) every 2 to 3 months. In one study of 27 patients, Btx-A significantly improved positioning in children with severe quadriplegia, including those with paraspinal spasticity (48), as well as improving gait function in those with hemiplegia and diplegia with severe gastrocnemius spasticity (49). In another recent study, Btx injected into thenar muscles reduced the CP cortical thumb and resulted in significant hand opening. Injection of biceps and forearm flexors reduces “flexion synergy” pattern in children with hemiplegia, but verification of functional improvement is still forthcoming. Typical injection patterns and suggested doses for individual muscles are available (50, 51).

Local injections of phenol or alcohol near a peripheral nerve or muscle motor point cause chemical neurolysis resulting in temporary decrease in tone and strength of the affected muscle. Distal regeneration from the site of injection results in loss of effect after 4 to 6 months (less for motor point blocks
than for nerve blocks). Potential benefits include prevention of deformity and improved function through facilitation of other therapies—such as physical therapy and tolerance of orthoses or positioning devices. In younger children, these blocks may buy time before requiring surgical intervention. Potential negative effects include temporary sensory dysesthesias (nerve blocks only), especially in the tibial nerve and upper limb nerves as well as permanent weakness leading to deformity (e.g., tibial nerve block may lead to calcaneovalgus foot).

In the lower limbs, obturator nerve blocks (anterior and posterior branches) reduce adductor tone, diminish scissored gait, and promote passive abduction as a means of protecting hip joint integrity. Sciatic branch blocks to the medial hamstring muscles (semitendinosus and semimembranosus) lessen crouch gait and internal rotation postures (52). Tibial nerve blocks (and more recently, tibial branch blocks to the heads of the gastrocnemius) diminish plantarflexion tone and allow better tolerance of AFOs. Femoral nerve block decreases spastic genu recurvatum. In the upper limbs, musculocutaneous nerve blocks promote elbow extension and facilitate reach. Median and ulnar nerve blocks are generally avoided due to the high risk of sensory dysesthesias. Motor point injections into the forearm, wrist, and finger flexors are preferred. Except for obturator block, Btx-A is the preferred agent of choice for affecting spasticity in the muscles supplied by the above nerves.

**Surgery in CP**

Orthopedic surgical intervention can be classified as either soft tissue or bony. Soft-tissue procedures are done at the muscle or tendon level and consist of releases, lengthening, or transfers. Recurrence of abnormalities tends to occur with the former two, especially if done before 4 years old (53). Transfers also tend to weaken at its new position but may balance forces across a joint, for example, medial hamstring transfer leading to less knee flexion, split posterior tibialis transfer leading to less ankle inversion, and rectus femoris transfer to sartorius transfers leading to less knee recurvatum (usually done in conjunction with hamstring lengthening). Bony procedures consist of either joint fusions (ankle or spine), (de)-rotations (of the femur or tibia), or angulations (of the femur). A list of specific procedures is included below (Table 56-5). The current trend is for multilevel soft-tissue surgery (rectus transfer, hamstring lengthening, and tendoachilles lengthening, TAL) and bilateral bony surgery (bilateral femoral varus osteotomies) to avoid imbalances and asymmetries. Computerized gait analysis measuring joint kinematics, kinetics, and dynamic EMG is recommended by many prior to undertaking complex orthopedic surgery (54,55).

Intensive postop rehabilitation is required to maximize gains and can be started while the child is still casted. Activities include upper limb strengthening, use of a tilt table to decrease postural hypotension and increase trunk control, and wheelchair and scoliosis cart ambulation to maintain cardiovascular conditioning. Once casts are removed, aggressive passive and active assistive range of motion, lower limb strengthening, transfer training, and gait training are begun. Generally, for complex procedures, up to 6 months may be required before the child’s preoperative functional status is regained, with further progress made for up to another year.

Selective posterior (or dorsal) rhizotomy (SPR, SDR) involves sectioning a variable percentage of sensory nerve rootlets after L2-S1 laminectomy. Theoretically, this results in a decrease in peripheral excitatory influences on the anterior horn cell in a spastic patient. Favorable patient selection criteria include lack of dystonia and/or atetosis, preservation of functional strength independent of spasticity, presence of selective motor control, younger age (3 to 8 years), lack of significant joint contracture and few previous orthopedic procedures. In addition, cognitive preservation, motivation, and positive family supports are important. Selection of “abnormal” rootlets is based on electrophysiologic measures (sustained contraction after a 1-second stimulation train, and spatial spread of contraction to different spinal root level muscles) and clinical (briskness and overflow of contraction) criteria at the time of surgery. It is preferable not to section all rootlets at any given level. More recent studies suggest contralateral spread of the EMG response or spread to the upper body may be the most valid abnormal criteria (56). Postop therapy early on avoids excessive hip flexion (>70 degrees), straight leg raising (>30 degrees) or trunk rotation, and passive flexion or extension is avoided in order to avoid stress on the operative site. By 6 weeks postop, full PROM can be started as well as functional strengthening and mobility training with or without orthoses. Small but statistically significant improvements in gait function as measured by the GMFM have been documented 1 year postop (57). Additional improvement in gait function has been documented for up to 2 years postop, but may not be significant when compared to physical therapy alone (58).

**TABLE 56.5 Orthopedic Surgical Procedure by Joint**

<table>
<thead>
<tr>
<th>Joint</th>
<th>Procedure Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Foot and ankle</td>
<td>TAL—for ankle equinus&lt;br&gt;SPLATT (split anterior tibialis transfer)—for inversion and dorsiflexion&lt;br&gt;Split posterior tibialis transfer—for inversion and plantarflexion&lt;br&gt;Subtalar arthrodesis—for calcaneovalgus. Often combined with lateral column (fifth metatarsal ray) lengthening</td>
</tr>
<tr>
<td>Knee</td>
<td>Hamstring lengthenings—for crouch, internal rotated gait&lt;br&gt;Rectus transfer (to sartorius or semitendinosus)—to balance hamstring weakness and prevent recurvatum&lt;br&gt;TIBIAL derotation osteotomy—for internal rotation</td>
</tr>
<tr>
<td>Hip</td>
<td>Psoas lengthening (intramuscular over the pelvic brim) for hip flexion&lt;br&gt;Adductor tenotomy—for scissored gait or early hip subluxation&lt;br&gt;Varus derotational osteotomy (VDO)—for hip subluxation&lt;br&gt;Pelvic shelf procedures (Salter, Chiari, etc.)—for subluxation with severe acetabular dysplasia</td>
</tr>
<tr>
<td>Knee</td>
<td>Medial hamstring lengthening, lateral hamstring lengthening, TAL, and subtalar arthrodesis— for calcaneovalgus</td>
</tr>
<tr>
<td>Foot and ankle</td>
<td>lateral column (fifth metatarsal ray) lengthening</td>
</tr>
<tr>
<td>Knee</td>
<td>TIBIAL derotation osteotomy—for internal rotation</td>
</tr>
<tr>
<td>Hip</td>
<td>Psoas lengthening (intramuscular over the pelvic brim) for hip flexion&lt;br&gt;Adductor tenotomy—for scissored gait or early hip subluxation&lt;br&gt;Varus derotational osteotomy (VDO)—for hip subluxation&lt;br&gt;Pelvic shelf procedures (Salter, Chiari, etc.)—for subluxation with severe acetabular dysplasia</td>
</tr>
</tbody>
</table>

**CHAPTER 56 | CHILDREN WITH DISABILITIES**

**TABLE 56.5 Orthopedic Surgical Procedure by Joint**

<table>
<thead>
<tr>
<th>Joint</th>
<th>Procedure Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Foot and ankle</td>
<td>TAL—for ankle equinus&lt;br&gt;SPLATT (split anterior tibialis transfer)—for inversion and dorsiflexion&lt;br&gt;Split posterior tibialis transfer—for inversion and plantarflexion&lt;br&gt;Subtalar arthrodesis—for calcaneovalgus. Often combined with lateral column (fifth metatarsal ray) lengthening</td>
</tr>
<tr>
<td>Knee</td>
<td>Hamstring lengthenings—for crouch, internal rotated gait&lt;br&gt;Rectus transfer (to sartorius or semitendinosus)—to balance hamstring weakness and prevent recurvatum&lt;br&gt;TIBIAL derotation osteotomy—for internal rotation</td>
</tr>
<tr>
<td>Hip</td>
<td>Psoas lengthening (intramuscular over the pelvic brim) for hip flexion&lt;br&gt;Adductor tenotomy—for scissored gait or early hip subluxation&lt;br&gt;Varus derotational osteotomy (VDO)—for hip subluxation&lt;br&gt;Pelvic shelf procedures (Salter, Chiari, etc.)—for subluxation with severe acetabular dysplasia</td>
</tr>
</tbody>
</table>

**CHAPTER 56 | CHILDREN WITH DISABILITIES**

**TABLE 56.5 Orthopedic Surgical Procedure by Joint**

<table>
<thead>
<tr>
<th>Joint</th>
<th>Procedure Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Foot and ankle</td>
<td>TAL—for ankle equinus&lt;br&gt;SPLATT (split anterior tibialis transfer)—for inversion and dorsiflexion&lt;br&gt;Split posterior tibialis transfer—for inversion and plantarflexion&lt;br&gt;Subtalar arthrodesis—for calcaneovalgus. Often combined with lateral column (fifth metatarsal ray) lengthening</td>
</tr>
<tr>
<td>Knee</td>
<td>Hamstring lengthenings—for crouch, internal rotated gait&lt;br&gt;Rectus transfer (to sartorius or semitendinosus)—to balance hamstring weakness and prevent recurvatum&lt;br&gt;TIBIAL derotation osteotomy—for internal rotation</td>
</tr>
<tr>
<td>Hip</td>
<td>Psoas lengthening (intramuscular over the pelvic brim) for hip flexion&lt;br&gt;Adductor tenotomy—for scissored gait or early hip subluxation&lt;br&gt;Varus derotational osteotomy (VDO)—for hip subluxation&lt;br&gt;Pelvic shelf procedures (Salter, Chiari, etc.)—for subluxation with severe acetabular dysplasia</td>
</tr>
</tbody>
</table>
Negative effects of SPR include hypotonia (usually transitory immediately postop, but occasionally lasting up to 6 months), weakness (unmasked by reduction of tone), sensory changes and bladder dysfunction (both usually of brief duration), hip dislocation (though exacerbated by sparing of L1 root leading to unbalance hip flexor spasticity) (59), and spinal deformity (no significant evidence of scoliosis is reported but lordosis may be associated with sparing of L1) (60). Some centers now start SPR at the L1 level.

There are recent reports of decreased plantar flexor spasticity as a result of sectioning rootlets at the S2 level. In one study, performing a “pudendal neurogram” (SSEP technique), and not sectioning rootlets carrying pudendal nerve responses, reduced bladder dysfunction from 24% to 0% (61). Results of SPR have been mixed and there is still great controversy regarding advantages of SPR over orthopedic surgery. Additional unanswered questions regarding SPR include the need for EMG monitoring during surgery (62) as well as the role of limited SPR, either by level (only L4 through S1) or percentage of rootlets cut (only 25% to 35%).

**Progressive Neuromuscular Disease**

**Early Diagnostic Features and Differential Diagnosis**

With the site of pathology within the motor unit, these disorders can be classified as originating from the anterior horn cell, peripheral nerve (either myelin or axonal), NMJ or muscle.

**Approach to Diagnosis**

In the history and physical examination, one should look for evidence of hypotonia with weakness, disproportionately delayed motor milestones compared to other developmental areas, and hyporeflexia or areflexia. Serum CK is a useful laboratory test for possible primary muscle disease, although consistent elevations are seen in dystrophinopathies and inflammatory muscle disease only. Serum electrolytes (particularly potassium), organic acids (lactate, pyruvate), and amino acids (carnitine) may also be useful. Genetic testing (high-resolution banding and DNA analysis) has allowed diagnosis of many diseases based on abnormal genetic loci (Table 56-6) and may avoid the need for biopsy as well as allow for prenatal diagnosis. Electrodiagnostic studies give about 90% positive correlation with neuropathy but only 44% with myopathy. Biopsy of either muscle or nerve, performed as open procedures, confirms diagnosis of nerve or muscle disorders. H&E stain of muscle shows gross structure—size of cells, uniformity, fatty or fibrous proliferation—while histochimical studies divide cells into subpopulations of type I (light) and type II (dark) cells (ATPase), show abnormal intracellular organelles (NADH-Tichrome) or accumulations of storage material (PAS). Electron microscopy gives additional resolution. Fluorescent antibody staining techniques can be used for dystrophin analysis in the dystrophies. Nerve biopsy assesses involvement of large versus small fibers and myelinated vs unmyelinated fibers.

<table>
<thead>
<tr>
<th>Locus</th>
<th>Disorder</th>
</tr>
</thead>
<tbody>
<tr>
<td>1q21.2q23</td>
<td>Charcot-Marie Tooth, type 1B</td>
</tr>
<tr>
<td>1q21q23</td>
<td>Nemaline myopathy, autosomal dominant</td>
</tr>
<tr>
<td>4q35</td>
<td>Facioscapulohumeral muscular dystrophy</td>
</tr>
<tr>
<td>5q13</td>
<td>Werdnig-Hoffman (infantile) SMA</td>
</tr>
<tr>
<td>5q13</td>
<td>Kugelberg-Welander (juvenile) SMA</td>
</tr>
<tr>
<td>15q</td>
<td>Limb-girdle muscular dystrophy</td>
</tr>
<tr>
<td>17p13.1</td>
<td>Charcot-Marie Tooth, type 1A</td>
</tr>
<tr>
<td>17q13.1q13.5</td>
<td>Paramyotonia congenita</td>
</tr>
<tr>
<td>19q13.3</td>
<td>Myotonic muscular dystrophy</td>
</tr>
<tr>
<td>19q12q13.2</td>
<td>Malignant hyperthermia syndrome</td>
</tr>
<tr>
<td>19q13.1</td>
<td>Central core disease</td>
</tr>
<tr>
<td>Xp21.2</td>
<td>Duchenne muscular dystrophy</td>
</tr>
<tr>
<td>Xp21.2</td>
<td>Becker muscular dystrophy</td>
</tr>
<tr>
<td>Xq13</td>
<td>Charcot-Marie Tooth, X-linked</td>
</tr>
<tr>
<td>Xq28</td>
<td>Emery-Dreifuss muscular dystrophy</td>
</tr>
<tr>
<td>Xq28</td>
<td>Myotubular (centronuclear) myopathy</td>
</tr>
<tr>
<td>Mitochondrial Mitochondrial</td>
<td>MERRF—myoclonic seizure, ragged red fibers</td>
</tr>
<tr>
<td></td>
<td>MELAS—encephalomyopathy, lactic acidosis, stroke</td>
</tr>
<tr>
<td></td>
<td>Cytchrome c oxidase deficiency—benign infantile myopathy</td>
</tr>
</tbody>
</table>

**Specific Disease of Infants and Toddlers—The “Floppy Baby”**

While the overwhelming majority of “floppy babies” will eventually show normal motor development (63), non-neuromuscular causes of pathologic hypotonia should also be considered (64). In early CP, clues in the history (prematurity, asphyxia, etc.) and physical exam (superimposed hypertonicity, hyperreflexia, abnormal primitive reflex profile) should arouse suspicion. One disease to consider is the Prader-Willi syndrome of hypotonia, small for gestational age, typical facies, failure to thrive early replaced by hyperphagia and marked obesity after 3 to 4 years, hypogonadism, small hands and feet, tapered digits, and a deletion at the long arm of chromosome 15-15q-. At the opposite end of the spectrum is Sotos syndrome or cerebral gigantism. Characteristics include hypotonia with macrosomia, megacephaly, large hands and feet, typical facies. Both syndromes are easily confused with congenital myopathies due to the hypotonia without obviously marked facial dysmorphism. CNS infections including bacterial and viral meningitis and sepsis as well as botulism cause acute presentations of hypotonia in a child who appears toxic.

**Anterior Horn Cell Diseases**

Spinal muscle atrophy (SMA) presents in two forms during infancy. Infantile SMA I or WHD is seen in 1/15,000 to 25,000 live births. Transmission is via an autosomal recessive inheritance pattern. The onset of hypotonia (Fig. 56-7) and global weakness with facial muscle sparing within the first weeks of life is
typical. Fasciculations are seen in the tongue only and there are no joint contractures. Stretch reflexes are absent but sensation is normal as is the serum CK. Electrodagnosis reveals normal motor conduction velocities but markedly decreased compound muscle action potential (CMAP) amplitudes. Sensory latencies and amplitudes are normal. EMG may demonstrate “chronic neuropathic” changes, although fibrillations may also be seen. Muscle biopsy shows rounded fibers, with areas of atrophy and compensatory hypertrophy. DNA analysis reveals a deletion on the long arm of the fifth chromosome (5q13) near the area of the survival motor neuron (SMN) gene (65). DNA testing is 95% reliable for diagnosis and can also be used for prenatal testing. The clinical course is progressively downhill to death in 90% by 2 years although long-term survival is now possible with the use of home ventilator support (66,67). Rehabilitative efforts should focus on provision of appropriately supportive adaptive seating and assistive technology for mobility and activities of daily living (ADLs).

Intermediate SMA II (also called chronic Werdnig-Hoffman disease) has the onset of progressive weakness and areflexia later in infancy, usually after sitting is achieved. A few patients become limited household ambulators during childhood with the assistance of AFOs and walkers. Fasciculations are more common and minipolymyoclonus is seen. Serum CK levels may be up to five times normal and increase with age. In addition to the nerve conduction abnormalities seen in type I SMA, large-amplitude motor units are more frequent on needle EMG. A tremulous baseline in the EKG represents cardiac muscle fasciculations and is almost pathognomonic for this disease. Muscle biopsy shows angular fibers and type grouping. Genetic findings are similar to SMA I. The clinical course is one of gradual progression with long-term complications of chronic wheelchair use (scoliosis, contractures, respiratory insufficiency) typically leading to death in early adulthood (Fig. 56-8). Contracture prevention, early spinal bracing, and surgical correction of scoliosis are necessary to allow optimum respiratory status to be maintained. As weakness progresses, noninvasive assisted ventilation may prolong life.

Diseases at the level of the NMJ include transient neonatal myasthenia gravis (MG) and congenital myasthenic syndromes. The former is seen in newborns of mothers with autoimmune MG due to transplacental transmission of AntiChR antibodies. The onset is usually within the first 24 hours after birth with feeding disorders, hypotonia, respiratory distress, weak cry, and facial weakness seen in decreasing order of frequency from over 90% to 50%. Ptosis is seen in only 15% of cases. Diagnosis is by either tension test or electrodiagnosis (see the text on autoimmune MG later). Symptoms are self-limiting and nonrecurring. Treatment is primarily supportive, although anticholinesterase (AChE) drugs (Mestinon) may be required in severe cases. Congenital myasthenic syndromes are an as-yet poorly delineated group of NMJ disorders that usually present at birth, but possibly not until years later. Most are autosomal recessive, but sites of pathology within the NMJ can be variably presynaptic, postsynaptic, or both. Diagnosis requires sophisticated electrodiagnostic and biopsy techniques and is available in few specialized centers only.

Peripheral nerve disorders include the hereditary motor-sensory neuropathies (HMSN) that are much more commonly observed in older children and will be discussed later.

**Congenital Myopathies of Infancy**

CMyD is an autosomal dominant disorder transmitted from an affected mother. Typical features include severe hypotonia at birth with respiratory distress often requiring prolonged ventilator support (68). Facial diplegia with a characteristic triangular-shaped mouth (Fig. 56-9), equinovarus contractures, and mental retardation are also seen. Clinical myotonia may not appear until 3 to 4 years age, while electrical myotonia is rarely seen at birth and may be absent until 2 to 3 years old. Clinical diagnosis is confirmed by examining the mother who invariably has the adult form of the disease. An abnormal
genetic locus on the long arm of chromosome 19 (19q13.3) results in a trinucleotide repeating sequence (CTG) (69). The number of repetitions, normally between 5 and 35, may be as high as 2,000 and is directly correlated with severity of disease, including cardiomyopathy (70). The genetic and clinical abnormalities tend to increase in severity with successive generations (termed genetic anticipation). Muscle biopsy is now rarely required. The clinical course is one of gradual improvement with hypotonia no longer clinically significant by 4 years. All will eventually become at least household ambulators, although soft-tissue surgeries and bracing may be required. Mild to moderate mental retardation, exacerbated by oromotor immobility affecting expressive language skills more than receptive skills, requires special education placement. Multisystem disease consisting of frontal baldness, cataracts, testicular atrophy in males, smooth muscle and cardiac muscle dysfunction, all typical of the adult form of myotonic dystrophy, is seen in CMyD adults and leads to a typically shortened life span, with death most commonly due to cardiac dysrhythmias.

CMD, a disorder of variable severity, is transmitted in an autosomal recessive pattern. Severe forms present early with hypotonia, proximal weakness with head and neck muscles involved early (Fig. 56-10), and congenital contractures usually seen distally only (ankles, wrists—“arthrogrypotic”). Cognition is usually normal. Milder forms may present with less severe gross motor delays and weak gait. CK is variably increased, while electrodiagnosis may or may not show “myopathic” patterns. Biopsy is similar to Duchenne dystrophy with variation in fiber size, fiber splitting, central nuclei, and fibrous proliferation. Severity of biopsy abnormalities is not necessarily correlated with clinical abnormalities. The clinical course can be static or slowly progressive, with scoliosis a common complication. Fukuyama CMD, a more severe form of CMD associated with seizures, mental retardation, and abnormal CNS imaging (71), is seen primarily but not exclusively in the Japanese population. With the discovery of the dystrophin-glycoprotein complex, specific component deficiencies have been associated with different forms of congenital dystrophy, the most common being merosin (or laminin) deficiency, which may also be associated with a peripheral neuropathy (72).

Other myopathies, once considered questionably distinct clinical entities, are named after histological changes seen in biopsy specimens treated with special staining techniques. Examples include nemaline (rod-body) myopathy, central core disease, centronuclear (myotubular) myopathy, and fiber type disproportion (73). In addition, each can demonstrate a broad spectrum of involvement (74). Recent discovery of abnormal genetic loci (see table) for many of these histologic myopathies supports the concept that these may indeed be true entities.

DISEASE OF OLDER CHILDREN

In older children, SMA III (Kugelberg-Welander disease) presents with proximal weakness during early childhood to young adulthood. Ambulation may be maintained into later years. Calf hypertrophy is seen in 25%, fasciculations in 75%. CK may be elevated two to five times normal. EMG shows a chronic neuropathic pattern (minimal fibrillations and positive sharp waves with large-amplitude polyphasic motor unit action potentials and diminished recruitment). Biopsy can show a mixed neuropathic/myopathic picture or can be nonspecific in 30%. The abnormal genetic locus on the long arm of chromosome 5 (5q13) is seen.

Several HMSN exist and have been classified by Dyck and Lambert. Most common are HMSN I (Charcot-Marie-Tooth—CMT) and HMSN II (neuronal CMT). HMSN III (Dejerine-Sotas) presents in early infancy, involves sensory nerves more often than in other types, and has more palpable nerve hypertrophy. Type IV (Refsum’s diseases) is associated with hearing loss and deficiencies in phytic acid. Type V (familial spastic paraplegia) may have upper motor neuron findings (75). Clinical findings in types I and II are similar.
Foot problems (pes cavus, claw toes, intrinsic atrophy) are common but can be clinically insignificant. In young children, pes planus is the most common presentation. Leg atrophy occurs later. The cavus foot appearance is secondary to progressive denervation leading to atrophy, fibrosis, and contracture in a distal to proximal direction. Type I can be distinguished from type II by electrodiagnosis. Type I gives a demyelinating picture—markedly slowed NCVs with few needle EMG abnormalities, while type II gives an axonal picture—relatively normal NCVs with a more abnormal needle EMG (76) and nerve biopsy. Several abnormal genetic loci have been found—the long arm of chromosome 1 (1q), long arm of chromosome 17 (17q), and an X-linked form. Orthopedic complications include hip dysplasia in about 6% to 8%, often asymmetrical (Fig. 56-11), and neuromuscular scoliosis in about 10%. Progressive cavus deformities and scoliosis may lead to the need for orthoses to improve ambulation and slow progression of spinal deformity. Resistant deformities may require surgical intervention (foot reconstructions, spinal fusions). As near normal physical activity as possible should be allowed.

Autoimmune MG occurs with a prevalence of 5 to 10/100,000 population and is four times as frequent in females. Clinically, abnormal fatigue occurs after activity and improves with rest. The onset is usually insidious although it may present precipitously after a febrile illness, allergic reaction, or emotional upset. The most common presenting sign is ptosis (ocular MG) that increases with reading and is accompanied by compensatory forehead wrinkling. Facial weakness may also be seen, presenting as a slack jaw, slurred speech or difficulty swallowing, and limb weakness, proximal more than distal with a positive Gower sign and Trendelenburg gait. Tensilon, a short-acting AChE, given intravenously, reverses weakness within 30 to 60 seconds and lasts 5 to 10 minutes. Electrodagnostic testing reveals a decrement of more than 10% on 2 to 3 Hz stimulation, while single-fiber EMG (SFEMG) shows abnormal jitter but is difficult to perform in children. Finally, antibodies to ACh receptors (Anti-AChR Abs) can be measured and are elevated in 85% to 90% of patients with generalized MG, but only 50% in ocular MG. Treatment is based on either improving NMJ transmission with AChEs or decreasing the lytic effect of Anti-AChR Abs by steroids, immunosuppressive therapy, or thymectomy. Commonly used AChEs include Prostigmine (neostigmine) or Mestinon (pyridostigmine). Steroids may be given in high doses once daily, low doses every other day or pulsed, each method having its proponents. Immunosuppression can be achieved by drugs (azathioprine, cyclosporin) IV gammaglobulin or plasma exchange (during crisis). The role of thymectomy is still controversial with optimal response seen when surgery is done within 5 years of onset and in relatively mild cases. Paradoxically though, thymectomy is usually performed only after other treatments have failed.

The presentation of facioscapulohumeral muscular dystrophy (FSHD) is often in adolescent or early adult years. Facial weakness often precedes shoulder girdle weakness. In later years, there may be progression to abdominal and pelvic girdle muscles. Associated abnormalities include high- and rarely low-frequency hearing loss and retinal abnormalities (telangiectasia, microaneurysms) in up to 75%. There are no cardiac abnormalities and cognition is preserved. CK elevation and muscle biopsy changes may be minimal. An abnormal genetic locus has been found at the 4q35 site. FSH and limb girdle dystrophy have many overlapping features: prominent facial involvement and autosomal dominant transmission favor FSH; proximal weakness and autosomal recessive transmission favor limb girdle. Both show variable severity and in their extremes, can be like Duchenne dystrophy with scoliosis, respiratory disease, and early death. The abnormal genetic locus is on chromosome 15.

Duchenne muscular dystrophy (DMD), a sex-linked recessive disorder, is seen in 1 to 3/10,000 male births. Gait deviations present after 2 years of age eventually lead to a Trendelenburg gait, pathognomonic of gluteus medius weakness. Patients come to stand through the Gower maneuver, pathognomonic of proximal pelvic girdle weakness. Gradually, a characteristic posture of tight heel cords, calf pseudohypertrophy (deltoid pseudohypertrophy also common but less visually obvious), a widely abducted stance, and hyperlordosis develops. Dull mentation is seen in about one third. CK is markedly elevated, maximally in the preclinical stage, with values greater than 10,000 not unusual. Values decrease with age as muscle loss progresses, but never approach normal. A “myopathic” EMG (BSAPPs, early recruitment) is seen. EKG shows cardiomyopathy with tall R-waves in the right precordial leads and deep Q-waves in the limb and precordial leads. These abnormalities are seen in 70% to 90% of patients and cardiomyopathy causes death in about 10% (77). Biopsy shows variation in fiber size, fiber splitting, central nuclei, fibrous/fatty replacement, and absence of type 2B fibers.

The abnormal genetic locus is at the Xp21 site, resulting in deficient or abnormal muscle protein. Dystrophin, assayed by immunofluorescent techniques in muscle biopsy specimens. Quantification of dystrophin now allows classification of the dystrophies based on the level of dystrophin deficiency. Less than 3% is characteristic of Duchenne while between 3% and 10% suggests severe Becker muscular dystrophy.
(BMD) and 10% to 20% suggests mild-to-moderate BMD. Levels above 20% suggest non-Becker/Duchenne dystrophies. Detailed DNA analysis on blood will soon be available using DNA probes, as will prenatal diagnosis with accuracy approaching 95%. Therapeutically, ongoing drug trials continue. Currently, both prednisone as well as androgenic steroids are used (78).

The clinical course is variable but inexorably downhill with ambulation typically lost by 8 to 12 years of age. Ambulation can be prolonged by bracing and/or timely surgical intervention to release contractures (79). Specifically, resection of the iliotibial band, TAL with or without posterior tibialis shortening or release can prolong ambulation an average of 2 years if done prior to initial entry into the wheelchair. Success depends on having a cognitively preserved patient, motivated for ambulation with a supportive family. Early remobilization is vital and only minimal orthotic support is needed. The wheelchair phase is characterized by progressive weakness, scoliosis in 80% (although 15% of these will not show significant progression and do not require surgical correction), and death by late teens to early twenties, usually secondary to respiratory failure on the basis of restrictive lung disease. Prolongation of useful life is possible by early surgical correction of scoliosis suggested once the curve exceeds 20 degrees and before lung vital capacity becomes less than 20%, and appropriate portable ventilator support (Fig. 56-12).

**BMD** occurs with an incidence only one tenth that of Duchenne. All clinical findings are the same as in Duchenne. Dystrophin deficiency is only partial (between 3% and 20%), rather than complete (<3%) as in Duchenne. The clinical course is protracted compared to Duchenne, with ambulation maintained into the 20s and survival to the 40s. Each form of dystrophy is genetically distinct.

**CONGENITAL BRACHIAL Plexus Palsies**

The reported incidence of congenital brachial plexus injuries is 0.6 to 2.6/1,000 live births. The mechanism of injury is traction to the brachial plexus. Most frequent is a neuropraxic injury from which more than 90% will recover fully. Axonal damage, rupture of the nerve, or nerve root avulsions result in variable residual damage. Risk factors include primiparous mothers, prolonged labor, birthweight more than 8.5 lb, shoulder dystocia (present in >50%), traumatic delivery with mid to high forceps, and breech presentation.

Injuries are divided and classified anatomically (80). Erb’s palsy involving the upper plexus (C5-7) accounts for about 80% of injuries. Klumpke’s palsy involving the lower plexus (C7, C8, T1) exclusively is now felt to be quite rare, with total plexus injuries accounting for 15%. Bilateral plexus injuries are also quite rare. Associated injuries secondary to use of forceps include facial palsies, cephalohematomas, and torticollis. Fractures of the clavicle, humerus and in worst cases, cervical spine can occur as can diaphragmatic paralysis. Horner’s syndrome (ptosis, miosis, anhydrosis) can be seen with lower plexus injuries and is secondary to stellate ganglion injury.

On initial examination, the position of the arm is related to muscle imbalances around the pivots (81). Erb’s palsies present with the typical “waiter’s tip” posture of shoulder internal rotation and adduction, elbow extension and pronation, wrist flexion and thumb in palm (due to loss of extensor pollicis) (Fig. 56-13). Klumpke’s posture is the reverse, with shoulder external rotation (abduction usually not seen due to gravity), elbow flexion and supination, wrist extension, and the intrinsic minus hand deformity due to loss of C8, T1 muscles...
strengthen deltoid, triceps, and wrist extensors. Active reaching in all directions is encouraged. Complete recovery occurs in more than 65% of children, while mild and moderate deficits are seen in 10% each. About 15% will have severe residual deficits. Poor grade elbow flexion with 0 to trace strength in wrist and finger extensors at 6 months old may be predictive of a poor outcome (84).

Since the mid-1990s, early neurosurgical repair of the injured brachial plexus has been advocated by some (85,86). Candidates for surgery have 0 to 1/5 strength in the biceps at 3 to 4 months or have plateaued at 2/5 for 4 months. The optimum age for surgery is less than 9 months since nerve growth factor is at a maximum before 1 year of age, and there is less nerve scarring and distal muscle atrophy at that age. Procedures performed include surgical neurolysis of scars, end-to-end anastomosis with microsurgical fascicular repair, and cable graft of nerve rupture. Intraoperative SSEPs may identify root avulsions not amenable to repair, while motor responses may distinguish neuropraxia from axonotmesis or neurotmesis. Postoperatively, up to 12 months may be needed for return of function to be seen. In one study, all muscles improved at least one grade by 9 months, with average strength in biceps and deltoids reaching 3/5. Surgical neurolysis also resulted in at least one grade improvement, while those patients requiring nerve graft showed the most dramatic improvement of between 2 and 3 grades. Others have questioned the criteria for early surgical repair, with recovery starting as late as 1 year seen (87). Additionally, soft-tissue procedures may be helpful (88). Subscapularis release improves passive shoulder external rotation, especially if performed before 4 years of age. Latissimus dorsi transfers can improve either elbow flexion or extension while shoulder abduction can be improved if coupled with trapezius transfer. Late deficits are largely cosmetic, with the presence of a high-riding, hypoplastic scapula with...
winging and loss of scapulohumeral rhythm. An elbow flexion contracture due to radial head subluxation and biceps contracture is almost always seen, and there is frequent shortening of the humeral segment. Finally, as with any disfiguring injury, the adverse effect on psychosocial development should not be underestimated, and appropriate counseling services provided.

**CHILDHOOD TORTICOLLIS**

The incidence of congenital muscular torticollis is about 4/1,000 live births, with 75% involving the right side (related to left occiput anterior presentation) (89). More recent studies have shown more equal presentation of right- and left-sided signs. A nontender, soft enlargement in the sternocleidomastoid (SCM) muscle, the so-called olive sign, may be seen within the first 6 weeks (90). It is mobile within the belly of the muscles, gradually decreases in size, and is usually gone within 4 to 6 months. Between 10% and 20% develop persistent torticollis and an additional 25% have mild SCM asymmetry. Secondary deformities including flattening of the ipsilateral face, contralateral occipital flattening, and orbital asymmetry can be seen (91). Ipsilateral hip dysplasia is seen in up to 12%. Typically, there are no spinal abnormalities or neurologic deficits. The differential diagnosis is quite extensive and is outlined in Table 56-7 (92).

The treatment of torticollis should be based first on correcting identified etiologic factors. Conservative therapy should include stretching of the tight neck muscles as well as functional strengthening of contralateral neck muscles by use of lateral and anterior head righting reactions. Directing the gaze toward the ipsilateral superior direction should also be encouraged. The use of skull-shaping orthotics has recently been advocated as an additional cosmetic corrective measure although much controversy exists (93). Surgical intervention should be considered when there has been no improvement by 18 to 24 months and consists of resection of a fibrotic SCM. Successful correction can be achieved as late as 12 years old (94).

**TABLE 56.7 Differential Diagnosis of Torticollis**

<table>
<thead>
<tr>
<th>Congenital postural torticollis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intrauterine crowding</td>
</tr>
<tr>
<td>No olive seen</td>
</tr>
<tr>
<td>Resolve with conservative treatment</td>
</tr>
<tr>
<td>Congenital vertebral anomalies</td>
</tr>
<tr>
<td>Segmentation failures</td>
</tr>
<tr>
<td>Formation failures</td>
</tr>
<tr>
<td>Combinations—Klippel-Feil, Sprengel’s</td>
</tr>
<tr>
<td>Congenital ocular torticollis</td>
</tr>
<tr>
<td>Strabismus</td>
</tr>
<tr>
<td>Nystagmus—spasmus nutans</td>
</tr>
<tr>
<td>Acquired structural torticollis</td>
</tr>
<tr>
<td>Traumatic</td>
</tr>
<tr>
<td>Rotary subluxation</td>
</tr>
<tr>
<td>Fracture/dislocation</td>
</tr>
<tr>
<td>Muscular injury</td>
</tr>
<tr>
<td>Basilar skull fracture</td>
</tr>
<tr>
<td>Infection</td>
</tr>
<tr>
<td>Cervical osteomyelitis</td>
</tr>
<tr>
<td>Tuberculosis</td>
</tr>
<tr>
<td>Ligamentous laxity and synovitis due to adjacent inflammation: pseudosubluxation, calcified cervical disc</td>
</tr>
<tr>
<td>Juvenile rheumatoid arthritis</td>
</tr>
<tr>
<td>Tumor of the cervical spine</td>
</tr>
<tr>
<td>Metastatic</td>
</tr>
<tr>
<td>Osteoid osteoma/osteoblastoma</td>
</tr>
<tr>
<td>Eosinophilic granuloma</td>
</tr>
<tr>
<td>Rare primary bone/muscle tumor</td>
</tr>
<tr>
<td>Atlanto-occipital subluxations</td>
</tr>
<tr>
<td>Associated with Down’s syndrome</td>
</tr>
<tr>
<td>Mucopolysaccharidoses and other conditions causing ligamentous laxity</td>
</tr>
<tr>
<td>Ocular strabismus (most commonly CN 4)</td>
</tr>
<tr>
<td>Infection</td>
</tr>
<tr>
<td>Retropharyngeal abscess</td>
</tr>
<tr>
<td>Vestibular infection (otitis media)</td>
</tr>
<tr>
<td>Cervical adenitis</td>
</tr>
<tr>
<td>Vestibular</td>
</tr>
<tr>
<td>Mastoiditis</td>
</tr>
<tr>
<td>Paroxysmal torticollis of infancy</td>
</tr>
<tr>
<td>Tumor</td>
</tr>
<tr>
<td>Midbrain tumors</td>
</tr>
<tr>
<td>Posterior fossa/intracranial tumor</td>
</tr>
<tr>
<td>Cervical cord tumor</td>
</tr>
<tr>
<td>Lymphoma</td>
</tr>
<tr>
<td>Acoustic neuroma</td>
</tr>
<tr>
<td>Orbital tumor (neurofibromatosis)</td>
</tr>
<tr>
<td>Miscellaneous</td>
</tr>
<tr>
<td>Migraine headache</td>
</tr>
<tr>
<td>Myasthenia gravis</td>
</tr>
<tr>
<td>Associated with GE reflux (Sandifer’s syndrome)</td>
</tr>
<tr>
<td>Syringomyelia</td>
</tr>
<tr>
<td>Habitual, hysterical</td>
</tr>
<tr>
<td>Oculogyric crisis (phenothiazine toxicity)</td>
</tr>
</tbody>
</table>
**Traumatic Spinal Cord Injury**

Pediatric traumatic spinal cord injury constitutes approximately 3% to 5% of all spinal cord injuries (95–97). There are several key anatomic differences between the child and the adult patient, and these predispose the child to SCIWORA (spinal cord injury without radiographic abnormalities). These differences include wedge-shaped vertebral bodies and horizontally oriented facet joints in the younger child; usually by age 8 to 10, the pediatric vertebral column resembles that of an adult. Paraspinal muscle weakness and ligamentous laxity are present as they are not fully developed in the younger patients. These factors favor ligamentous injury over bony injury (98,99). Young children also tend to have proportionally large heads compared to their body size. Therefore, the fulcrum of injury is higher resulting in higher cervical injuries than in adults (97,100) (Table 56-8). The incidence of SCIWORA varies in the literature, but SCIWORA accounts for 33% to 50% of traumatic spinal cord injury in children less than 8 years of age. SCIWORA was first termed by Pang in 1982. His follow-up article reviewing two decades of literature reported that epidemiological data suggest: (a) in the age group from birth to 9 years, there are fewer fractures and subluxations, (b) the neurologic lesion in a young child is more severe than that in an older child, and (c) because the cervical spine is inherently more unstable in the young child there are more upper cord lesions in children younger than 8 years and in older children and all of these lesions are severe (101). He also reported that the MRI findings could be used to help prognosticate outcome. Some authors/clinicians have raised the issue that term SCIWORA excludes injuries from penetrating trauma, electrical shock, obstetrical complications, and those associated with congenital spinal anomalies.

It is important to be aware of the presence of SCIWORA, as in the unconscious pediatric patient in the emergency room, a lateral neck x-ray does not rule out spinal cord injury. Further diagnostic tests such as cervical MRI and CT scans may be warranted.

The management of the pediatric spinal cord injured patient is similar to the adult with a couple of exceptions. The exam of the pediatric patient is usually limited by the child’s ability to cooperate. At approximately age 5, depending on the cognitive abilities of the child, accurate manual muscle testing can be performed. It is important to remember that when assessing the pediatric patient, the resistance during manual muscle testing is not the same as exhibited by an adult. In children who cannot participate in manual muscle testing, the grading of muscle strength is an estimation. Also in the young patient, especially infants, the sensory exam should be done caudal to cranial. The child should be calmed, and then pinprick testing can be done, with the onset of a cry as the indication that sensation is present. Grading the sensory exam in infants is virtually impossible and only an estimate of the sensory level can be obtained.

There are limited data on deep venous thrombosis in the pediatric population. A retrospective review by Radecki found 1/20 patients under age 13 had a deep venous thrombosis. For the immature adolescent and child, the issue of DVT prophylaxis is not clear cut and further investigation is warranted (103).

**Spinal Dysraphism**

There are several different forms of spinal dysraphism with several different terminologies used (Fig. 56-15). Spinal bifida occulta, a normal variant seen in 5% to 10% of the population, is the failure of fusion of the posterior elements of the spine with an intact thecal sac and normal spinal cord. There are usually no associated neurologic deficits although it is rarely associated with a sacral lipoma or a tethered cord. Meningocele is the failure of fusion of the posterior elements with cystic-outpouching of the thecal sac filled with CSF, but with no neural tube disruption. Therefore, there are no neurologic deficits. The skin may or may not be intact. Meningomyelocele is the significant disruption of all elements of the bony spine typically with an open malformed neural tube covered by a membranous sac. There are variable degrees of neurological deficits. Another term that may be seen is spina bifida aperta, which is any lesion that is open to the environment. Some authors also use spina bifida occulta as any lesion that is not open to the environment.

**Embryology**

Normally, neural tube closure starts in the third week of gestation from the midcervical level and proceeds in both cephalad

---

**TABLE 56.8 Distribution by Level of Traumatic Spinal Cord Injury**

<table>
<thead>
<tr>
<th>Injury Type</th>
<th>Group 1 (0–8 Years)</th>
<th>Group 11 (9–16 Years)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total cervical</td>
<td>49 (79%)</td>
<td>63 (53%)</td>
<td>112 (62%)</td>
</tr>
<tr>
<td>Upper cervical (0-C3)</td>
<td>33 (53%)</td>
<td>31 (26%)</td>
<td>64 (36%)</td>
</tr>
<tr>
<td>Lower cervical (C4-7)</td>
<td>16 (26%)</td>
<td>32 (27%)</td>
<td>48 (26%)</td>
</tr>
<tr>
<td>Thoracic (T1-11)</td>
<td>7 (11%)</td>
<td>16 (14%)</td>
<td>23 (13%)</td>
</tr>
<tr>
<td>Thoracolumbar (T12-L1)</td>
<td>2 (3%)</td>
<td>17 (15%)</td>
<td>19 (11%)</td>
</tr>
<tr>
<td>Lumbar (L1-5)</td>
<td>4 (7%)</td>
<td>21 (18%)</td>
<td>25 (14%)</td>
</tr>
</tbody>
</table>
and caudad directions. The defect of neural tube closure is thought to occur around day 26 and accounts for most lesions through midlumbar. The most caudal cell mass forms between days 26 and 30 eventually resulting in formation of a central canal in the embryonic tail. Caudal regression with rostral extension resulting in fusion with the neural tube results in the formation of the spinal cord by day 53. Lesions of the lumbosacral levels occur before day 53.

Demographics

The incidence of neural tube defects worldwide is approximately 1 to 2/1,000 births depending on the population studied. In 2005, the U.S. rate was 18/100,000 live births and there were 734 live births with spina bifida (104). However, it is postulated there may be underreporting of this entity. Differences have been noted between races and nationalities. The populations of Irish, Scottish, or Welsh individuals have higher risks. In the United States, Hispanics have a higher risk for neural tube defects whereas the risk is lowest among African Americans (105). The incidence is decreasing worldwide and is felt to be due to several factors including better prenatal diagnosis and termination of the pregnancy, better nutritional support, and the use of folic acid and fortification of foods with folic acid (106–108).

Prenatal Detection

Early detection can be accomplished via serum alphafetoprotein with levels peaking at the 16th week of gestation. False positives can be caused by incorrect dates, multiple gestations, maternal cancer, and GI or renal disease (maternal or fetal). False negatives occur with incorrect dates, intrauterine growth retardation, maternal diabetes, and Down syndrome.

Fetal ultrasound can be used to assist with diagnosis. Amniocentesis for AFP and acetylcholinesterase can be performed to confirm the diagnosis.

Etiology

The specific etiology of neural tube defects has not been elicited. However, the mode of inheritance appears to be multifactorial with genetic and gestational influences such as folic acid deficiency and fetal exposure to valproic acid (109) or Tegretol (110). Maternal hyperthermia such as the use of hot tubs or saunas during early pregnancy has been implicated as a risk factor (111). Maternal diabetes has also been implicated as a risk factor but supportive data are lacking. The risk of recurrence has been reported in different studies to be approximately 1% to 4% after the first affected child and as high as 10% after the second. The risk of an affected mother giving birth to an affected child is 4% to 10%.

The risk of recurrence can be reduced by up to 70% with periconceptual use of folic acid in doses up to 4 mg/day. Multiple studies worldwide using 0.4 mg to 5 mg/day, given 1 month preconception through the first trimester have been performed with the majority citing a 60% to 70% reduction in risk of recurrence of neural tube defects. The Centers for Disease Control (CDC) recommends that all women of child-bearing years consume 0.4 mg/day of folic acid. Folic acid intake should be kept under 1 mg/day unless under a physician’s care. Some physicians believe that higher doses such as 4 mg/day should be used if the woman has already had a child with a neural tube defect (112,113). The exact mechanism of folic acid in prevention has not yet been determined. There is an association between elevated homocysteine levels and neural tube defects. A variation (C677T) in the gene coding for 5,10-methylenetetrahydrofolate is associated with a 70% increased risk of spina bifida.

Neural Tube Closure

Sac closure is usually performed within the first 24 to 48 hours to reduce the risk of infection. Microsurgical closure with reapproximation of the neural tube and construction of a fluid-filled pouch to “bathe” the cord had reduced the incidence of early tethering of the spinal cord.

Newer neurosurgical techniques now involve repair. Tulipan and Bruner reported three successful cases of in utero repair (114). A case study of 29 patients repaired in utero suggests that intratrune repair of myelomeningocele decreases the incidence of hindbrain herniation and shunt-dependent hydrocephalus. However, there was an increase in prematurity (115–117). A further study strongly suggests that it reduces the incidence of shunt-dependent hydrocephalus compared to conventional treatment. It remains to be proven whether the risks of intratrune surgery outweigh the potential benefits: five patients in one study died due to premature delivery (118).

Associated Neurologic Abnormalities

The Chiari malformation type II, seen in 90% of patients, is defined as downward displacement of the inferior portion of the cerebellar vermis, the medulla, lower pons, and an elongated fourth ventricle through the foramen magnum. There is also an associated kinking of the medulla.
Hydrocephalus occurs in 80% to 90% of patients. A “cork in the bottle” phenomenon, related to the abnormal placement of the above-mentioned structures through the foramen magnum, aqueductal stenosis, and a defect in CSF uptake are all causative factors. More than 80% of children require ventriculoperitoneal shunting.

Symptoms of the Chiari malformation and/or hydrocephalus include laryngeal and pharyngeal paresis and may be associated with stridor and/or apnea during infancy and early childhood. Infants may be born with symptoms but typically develop them within the first 2 months with the peak mortality between 9 and 12 weeks. Mild inspiratory stridor with crying is usually the initial symptom due to paresis of the vocal cord abductors. Progressive cardiorespiratory problems including apnea and bradycardia may develop, as well as aspiration pneumonia. Approximately one third of infants develop symptoms and about one-third of these will die (119).

A tethered spinal cord may be due to a lipoma, scarring at the closure site, or diastematomyelia. Presenting symptoms include deterioration in bladder and bowel function, loss of strength or sensation, spasticity, low back pain or radicular pain and/or rapidly progressive scoliosis. It is felt that microinfarctions of the spinal cord occur with repetitive spinal flexion and extension leading to symptoms. Urodynamics may be helpful, especially since a bladder that is becoming more spastic on urodynamics, may be a symptom of a tethered cord. Treatment requires early surgical detethering. McClone reported that 93% of patients that were detethered, had improvement or stabilization of presenting symptoms (120). Spontaneous resolution of scoliosis has been reported after detethering. A tethered cord is present in almost all patients with meningomyelocele (121) though it may be asymptomatic. Symptomatic tethered cord requiring surgical intervention occurs in about 15% to 25% of patients, but in some studies it is reported as high as 50%. Retethering can occur.

Syringomyelia is felt to be due to high CSF pressure within the central canal. Presenting symptoms include deterioration of neurologic function, loss of sensory and motor function in the upper limbs, and spasticity resulting in progressive joint deformity. It may also be present in 25% of asymptomatic patients. Shunting may be necessary when the process is symptomatic.

**Orthopedic Issues**

There are associated musculoskeletal malformations and complications that can lead to further impairment. Two types of scoliosis can occur: structural and paralytic. Structural scoliosis is caused by vertebral body anomalies such as wedge or hemivertebrae, unilateral bars or block vertebrae, alone or in combination. Paralytic scoliosis is secondary to loss of truncal support and is seen in most thoracic-level patients. Kyphosis may be structural or paralytic, with higher spinal levels at greater risk for paralytic kyphosis. Rib anomalies are also seen including fused or malformed ribs. Contractures can occur at any joint due to muscular imbalance. Hip dysplasia and club feet also occur. Pathologic fractures occur below the level of the neurologic lesion, and may be caused by trivial trauma and present as painless swelling, warmth and/or redness. There are no good data to indicate that passive standing alone reduces the risk of pathologic fractures.

Scoliosis is typically managed initially with spinal bracing for curves greater than 20 degrees and followed with regular radiographic monitoring. Curves greater than 40 degrees are considered surgical curves. Rapidly progressive scoliosis should raise the suspicion of a tethered cord. A severe kyphosis (gibbus) may be corrected by a kyphectomy. Bilateral hip dislocations without restricted range of motion may be left untreated. Unilateral hip dislocation in low-level lesions should be corrected, but in high-level lesions, surgical correction may have no benefit (122). Unilateral hip dislocation may lead to pelvic obliquity though the relationship between pelvic obliquity and scoliosis has been questioned. Serial casting or soft tissue releases to improve range of motion may be needed to increase mobility, improve positioning for seating and/or bracing.

After surgical procedures postoperative immobilization should be limited as much as possible to prevent further osteoporosis and pathologic fractures.

**Genitourinary**

The majority of patients have a neurogenic bladder, with a flaccid bladder being more common than a spastic bladder. In addition, there are associated genitourinary abnormalities seen in up to 20% of patients including horseshoe kidneys, hypoplastic kidneys or renal agenesis, as well as ureteral duplications and posterior urethral valves.

A renal ultrasound should be performed early in infancy to delineate the anatomy. A voiding cystourethrogram (VCUG) can define bladder contour and determine the presence of ureteral reflux. Renal scans and urodynamics can be performed after 2 weeks of age. Since the physiology can change over the first year of life, serial ultrasounds should be performed every 3 to 4 months to monitor the status of the kidneys and detect any changes that would necessitate further intervention. Most patients will be managed with clean intermittent catheterization (CIC). The age at which this may begin is debatable. CIC may be started at birth if the child has a known neurogenic bladder. Alternatively, since infants are normally not continent, if bladder pressures are not elevated, there are no changes on ultrasound and the patient is free of urinary tract infections, then CIC can be started at the age when continence is normally achieved, perhaps 3 to 4 years old. However, once the first urinary tract infection occurs, then catheterizations must be started. The long-term goals of bladder management are to prevent renal damage by preventing infection and reflux and to prevent wetness between catheterizations. Routine renal-bladder ultrasounds should be performed every 6 to 12 months in order to properly monitor the urologic system.

Pharmacologic intervention may be needed based on the results of the urodynamics and the patient’s clinical status with anticholinergic medication such as Detrol or Ditropan being used for spastic bladders.
Surgical procedures include bladder augmentation for small-capacity bladders, urethral implantation for patients with reflux, and suprapubic vesicostomy to allow for an alternative method of drainage. The Mitrofanoff procedure is used to provide an alternative conduit for urinary catheterization. The artificial sphincter is an implantable device used to help achieve urinary continence. It can be considered in a motivated patient after all other conservative methods have failed. CIC is still required. Patients should be free of infection, have normal renal function, adequate bladder capacity, and low pressure bladders with no uninhibited contractions. Problems include rejection, foreign body reaction, and infection. Neural stimulation involving stimulation of the sacral roots (123) and pudendal nerve stimulation have also been reported. Children with good fine motor skills and good cognition can be taught self-catheterization at approximately age 5. The goal is to have the child independent so their personal space is not invaded routinely. Controversial issues include the use of prophylactic medications such as antibiotics and the benefit of vitamin C.

Bowel
The management of the bowel program in patients with meningo(myelo)cele is a constantly changing process. Bowel incontinence can cause terrible social stigmatization. Training in timed bowel regulation may be started by 2 to 3 years developmental age. Since peristalsis and the gastrocolic reflex are still intact, postmealtime evacuations are usually more successful. Diet is an important component of the bowel program. Adequate fluid intake is important as well as knowing which foods soften or harden the stools. Bulk additives, stimulant suppositories, and enemas may be needed. For patients with uncontrollable incontinence, the Malone antegrade continence enema (ACE) provides a surgical alternative, with a conduit using either the cecum or the appendix between the intestine and the abdominal wall. This is a catheterizable conduit that can be used to deliver tap water or saline enemas and allow for evacuation through the rectum.

Sexuality
Sexuality does not only relate to sexual functioning but also to one’s perception of self that may lead to problems of decreased self-esteem, anxiety, depression and lead to impaired social interactions. It is important to address these issues with appropriate supportive counseling when they arise.

Females may have relatively normal function in spite of sensory loss. Males have variable ability for erection and ejaculation, but fertility may be compromised by recurrent UTI, repeated mechanical trauma to the testicles, testicular hypoplasia, and decreased temperature control of the scrotum resulting in decreased sperm count.

Skin
Patients with spina bifida are at risk for decubitis sores due to impaired mobility and impaired sensation. Meticulous skin care is important with weight shifts being done routinely by the family/caregivers and then the child when he/she gets older and can follow through on his own. There are reports in the literature of patients requiring limb amputation due to severe decubitis. Occipital decubiti are more common in younger children that are immobile due to the large head size. Prevention is the best treatment.

Latex
Latex allergy should be considered in every child with meningo(myelo)cele. Numerous different types of allergic reaction have been reported including anaphylaxis. Protein in the latex is the primary allergen. The FDA estimates that 18% to 40% of children with meningo(myelo)cele and 6% to 7% of surgical personnel are latex sensitive (124). Niggemann reports that atopic disposition, number of operations, and presence of a shunt all increase the risk of becoming not only sensitized but also allergic to latex. Sensitization may occur at any time, so serum IgE levels may give a false sense of security (125). Therefore, latex should be avoided in all patients with meningo(myelo)cele. One study from the Cologne Children's Hospital found a significant reduction in latex allergy by practicing primary prophylaxis: latex was avoided for all surgeries, anesthesia, and on the wards (126). This also indicates that repetitive exposure plays a factor in latex allergy.

Rehabilitation
The rehabilitation program for children with meningo(myelo)cele begins in the newborn period. A working knowledge of normal development is important. A careful examination in the newborn period, often by observation of the position of the infant, can help discern the functional level of the patient that result from imbalance of muscular forces around major pivots. Thoracic-level lesions are most often associated with flaccid lower limbs and with frog-legged positioning. High lumbar (L1-2) levels tend to lay with hips flexed and adducted. The knees are often flexed because of intrauterine positioning. This position predisposes to early hip dislocation. Children with low lumbar (L3-5) levels lay with the hips flexed and adducted but with the knees in extension and possibly the ankles dorsiflexed. Hip dislocation may be a late complication. Individuals with sacral levels may assume normal postures, but pes cavus and clawing of the toes may be seen.

Children develop psychosocially by interacting with their environment. If motor limitations exist, adaptive devices will be needed to help facilitate this interaction. Adaptive equipment in infancy is designed to allow attainment of functional milestones at as near to normal time as possible. Proper seating with appropriate relief for deformities allows the infant to sit upright and view the environment. At about 1 year of age, standing devices can be considered. The data are limited in regards to passive standing alone and the prevention of osteoporosis. The devices used include a stander or parapodium for thoracic-level patients, although the parapodium is typically not used for mobility until at least 18 months of age. For thoracic-level patients, the parapodium can be replaced by HKAFOs with the use of a walker or Lofstrand crutches, depending on the child’s abilities. Ambulation in childhood,
even in high-level patients, is a reasonable goal, although by adolescence, wheelchair ambulation may be required as a result of the energy demands of upright ambulation. The reciprocating gait orthosis (RGO) allows for an upright energy-efficient gait pattern. The isocentric RGO is the newest version, using a bar with a central pivot (Fig. 56-16). The quality of gait is improved if the patient can assist with active hip flexion. The energy consumption of ambulation with RGO approaches that of wheelchair mobility. In the study by Cuddeford and associates, energy consumption and energy efficiency of HKAFOs and RGO were compared. Previous studies had shown less energy consumption with RGO than with a swing-through gait with HKAFOs. The children who ambulated with HKAFOs had significantly higher energy consumption than those with RGO, but they ambulated faster. Therefore, the energy efficiency was less for the HKAFO group. The authors believe that further investigation is needed to confirm the importance of energy efficiency during ambulation in considering a mode of locomotion with the child with meningomyelocele (127). High-lumbar-level patients may receive bracing as early as 10 to 12 months, usually with HKAFOs. For lower-level patients, the residual motor movements will dictate the type of orthotic needed. Wheelchair ambulation can be achieved by 18 to 24 months, and steering a power wheelchair can be achieved before 3 years of age.

There are multiple factors that affect ambulation, including level of paralysis, age when first braced (128) obesity, degree of scoliosis, lower-extremity deformities, weight and efficiency of the orthosis, child's motivation, parental support, cognition, and upper-extremity function. If a child with meningomyelocele wants to ambulate, she should be given every opportunity to ambulate. Carney and Melchionni, in a retrospective study, evaluated high-lumbar and thoracic-level patients and found that the significant factors determining ambulation were degree of mental retardation, whether the child received physical therapy for ambulation, and parental involvement (129). Typically, most sacral-level patients will ambulate as adults, and most thoracic-level patients will require wheelchair mobility.

A high proportion of these children will have learning disabilities, and the appropriate supports need to be provided in the school setting. Supportive counseling may be needed to help the child cope and effectively manage his/her disability.

**Prognosis**

There are several studies regarding follow-up of patients with meningomyelocele.

Worley and associates monitored 63 patients (32 boys and 31 girls) for 5 years from infancy. Ninety-eight percent had VP shunts and 23% had symptoms of brainstem dysfunction, with the median age of 3 months (1 to 22 months). Eleven of fifteen underwent brainstem decompression. In four of fifteen, symptoms resolved spontaneously. Mortality rate at 5 years was 14%; five of nine died as a result of brainstem dysfunction (130). Hunt and Poulton reported a 25-year follow-up of 117 patients born between 1963 and 1970. Fifty-two percent, or 61 patients, were alive in 1992. Seventy-nine percent lived to their first birthday, 63% to their tenth birthday, and 56% to their twentieth birthday. The most severely disabled patients had the highest mortality. Sixteen deaths were due to renal failure. CIC was not a standard of care when many of these patients were younger, however. Of the 61 survivors, only 33 were able to live independently (131).

**A Team Approach**

The rehabilitation of the child with spinal cord dysfunction is multifaceted and requires a team approach. The physiatrist can guide the team, which may include an orthopedic surgeon, neurosurgeon, urologist, psychologist, social worker, physical therapist, occupational therapist, speech therapist, and nutritionist. Finally, it is important to include not only the child as part of the team, but the parents as well. Parents must be involved for a successful rehabilitation program.
REFERENCES


107. Williams L, Mai C. Prevalence of spina bifida and anencephaly during the transition to mandatory folic acid fortification in the United States.  


110. Rosa FW. Spina bifida in infants and women treated with carbamazepine during pregnancy.  


112. Centers for Disease Control and Prevention. Recommendations for the use of folic acid to reduce the number of cases of spina bifida and other neural tube defects.  


116. Sutton LN, Adzick NS, Bilaniuk LT. Improvement in hindbrain herniation demonstrated by serial fetal magnetic resonance imaging following fetal surgery for myelomeningocele.  


120. Herman JM, McLone DG, Storrs BB. Analysis of 153 patients with myelomeningocele or spinal lipoma reoperated upon for a tethered cord.  


122. Fraser RK, Bourke HM, Broughton NS. Unilateral dislocation of the hip in spina bifida.  

123. Schmidt R, Royan BA, Tanayho EA. Neuroprosthesis in the management of incontinence in myelomeningocele patients.  


126. Cremer R, Kleine-Diepenbruck U. Reduction of latex sensitization in spina bifida patients by a primary prophylaxis programme (five years experience).  


129. Carney E, Melchionni J. Community ambulation by children with myelomeningocele and high level paralysis.  

130. Worley G, Schuster JM, Oates WJ. Survival at 5 years of a cohort of newborn infants with myelomeningocele.  

Adults with Childhood-Onset Disabling Conditions

This chapter will outline the most common congenital and childhood-onset disabling conditions and the associated secondary conditions encountered as adults. An effort has been made to be comprehensive while emphasizing a practical approach to caring for the adult patient. Medical transition from pediatric to adult care will be introduced as it applies to public health initiatives. We sought to emphasize the many ways that a preventive approach to health and wellness is appropriate in the comprehensive care of adults with childhood-onset disabilities. While we will highlight some aspects of care in this chapter, we want to direct the reader to a more comprehensive treatment of some specific conditions elsewhere in this book, specifically the chapters on osteoporosis, sexuality, women's health, and primary care.

INTRODUCTION

Children with complex medical and genetic conditions are living longer than was predicted in the past due to better treatments and overall care (1). According to 2000 nationwide census data, approximately 6% or 2.6 million children and youth aged 5 to 15 years have a disability. Currently, 90% of children born with special health care needs reach adulthood (2). Often, the success of excellent pediatric medicine results in young adults who require ongoing, sometimes intensive, medical supervision throughout their lives.

In 2002, the American Academy of Pediatrics, The American Academy of Family Medicine, and the American College of Physicians—American Society of Internal Medicine collaborated on a policy statement to ensure that young people with special health care needs were equipped to move from child-centered to adult-centered health delivery systems (3). In a report from the Maternal and Child Health Bureau (MCHB), published in 2001, three general barriers were identified to health care transitions for young adults: training of qualified providers, funding for health delivery, and bureaucratic obstacles.

In reality, the process of transition should be viewed as a continuum. There are many barriers that challenge families, patients, and physicians alike that greatly complicate the process. The physician receiving transitioning youth must have an understanding of what some of these barriers entail in order to deliver compassionate care to young adults attempting to make the leap into the adult health care system.

The Medical Home model has been viewed as having primary responsibility for taking charge of transition activities to ensure that young adults with chronic, childhood-onset illness continue to receive the quality of health care that characterized their pediatric years. Thus, the role and responsibility of a medical home has continued to expand with ideals that are difficult to attain without a thorough understanding of the team approach implied in the concept.

The MCHB defines children and youth with special health care needs (CYSHCN) as “those who have, or are at increased risk for having, a chronic physical, developmental, behavioral or emotional condition and who also require health and related services of a type or amount beyond that required by children generally” (4). In 1989, The Omnibus Budget Reconciliation Act (OBRA), Public Law 101-239, amended Title V of the Social Security Act to include the authority and responsibility of MCHB to fully address the needs of all children. In addition, the focus under MCHB for CYSHCN was to provide leadership in promoting a community-based system of services that is family centered, comprehensive, coordinated, and culturally competent, thus building services around a medical home approach. The goals for the Title V CSHCN Program include:

1. Families will partner in decision making at all levels and will be satisfied with services.
2. All CYSHCN will receive coordinated, ongoing, comprehensive care within a medical home.
3. All CYSHCN families will have adequate insurance to pay for the services they need.
4. All children will be screened early and continuously for special health care needs.
5. Community-based service systems will be organized so that families can use them easily.
6. All youth with special health care needs will receive the services necessary to make transitions to all aspects of adult life.

While not all these goals will be the direct responsibility of one clinician, the overall success of medical care delivered in a
The medical home concept at the hub of chronic illness care services they need when they need it, even outside regular business hours and includes having a way to pay for needed services. Family-centered care is health care that takes into consideration the requests of the family and respects them as partners in the health care decision making of the medical home concept. Cultural responsiveness considers the values and beliefs that the family and the patient bring with them into the medical encounter. Continuity of care implies communication among all health care providers, especially primary care physicians and specialists. Comprehensive care is medical care that includes primary medical and developmentally appropriate care, dental care, mental health services, as well as crisis and chronic illness management. Application of routine primary care guidelines is also a component of comprehensive care. Compassionate care considers the concerns of the family and how the diagnosis of the patient may affect others. Coordination of care embodies the day-to-day development of a health care plan, including the transition plan, for a patient and ensures its implementation.

With the medical home concept at the hub of chronic illness care, rehabilitation professionals will be challenged to provide leadership in long-term follow-up of our patients. Preserving wellness and independence is the ultimate goal, not necessarily episodic/acute care management. Preservation of quality-of-life indicators, as individually applied to each patient, while postponing functional decline is an important role.

As young adults with disabilities transition to more independence, physicians need to educate patients and their families about potential age-related changes and medical issues. There is a growing interest in the literature about aging issues and secondary conditions among persons with congenital and childhood-onset disabling conditions. Differentiating between primary, associated, and secondary conditions may help the clinician categorically organize a patient’s care goals.

The primary condition is self-explanatory. It is the diagnosis that initiated entering into medical care from the point of evaluation during gestation, infancy, or childhood. Secondary conditions are impairments, functional limitations, disabilities, diseases, injuries, or other conditions that occur during the life of a person with a disability, in which the primary disabling condition is a risk factor for that secondary condition or may alter the standard intervention for prevention of treatment of any health condition (Syracuse Conference, 1994). A disabling condition, such as a spinal cord injury (SCI) or cerebral palsy (CP), is not synonymous with illness or disease. People with disabilities are generally healthy, and one should not expect a decline in health and function as they mature. Therefore, physicians must have an index of suspicion for such conditions that may have onset in late adolescence or adulthood or be insidiously progressive. They may include progression of pathology or impairment, either through complications or through the aging process. Commonly reported secondary conditions include pain, contractures, recurrent urinary tract infections, pressure sores, and osteoporosis. Associated conditions or residual impairments are conditions that result from the defect, injury, or disease and often may be considered primary impairments depending on their severity. An example is neurogenic bowel/bladder in patients with spina bifida. Comorbidities are medical conditions unrelated to the primary disabling conditions such as adult-onset diabetes or hypertension in a patient with CP.

Health perception is an individual determination and is affected by personal expectations, experiences, sense of vulnerability, support, and locale. The World Health Organization (WHO) defines quality of life (QoL) as “an individual’s perception of their position in life in the context of the culture and value systems in which they live, in relation to their goals, expectations, standards, and concerns” (6). Health-related QoL is a broad concept affected in a complex way by an individual’s physical health, psychological state, personal beliefs, social relationships, and their relationship to salient features of their environment. Although it is widely acknowledged that this is an area of concern for adults with chronic disabilities, there are limited studies that have systematically assessed QoL in these individuals, particularly those with neuromuscular disorders (7). This may be due in part to the fact that QoL is perceived as a somewhat vague entity that many people are concerned about, but something that nobody clearly knows what to do about. These issues are important in examining disability in adults who are survivors of disabling childhood disorders.

What follows is an overview of common chronic conditions and specific management recommendations for care.

**SPINAL CORD INJURY**

**Background**

When SCI occurs at a young age, there are unique medical complications that can significantly impact the adult years. Transition into adult care requires knowledge of these issues
as well as a multidisciplinary treatment approach. Recent data suggest that adults with childhood-onset SCI have a slightly lower life expectancy compared with those who sustained their injury as an adult. However, those with childhood-onset SCI are still living long lives with life expectancy estimates of approximately 83% of normal for those with minimal deficit incomplete injuries compared with 50% of normal in non–ventilator-dependent tetraplegics (8).

According to the 2006 Annual Statistical Report for the Model SCI Care Systems, approximately 253,000 people are living with SCI in the United States, with 11,000 new cases reported each year. Less than 20% of SCI occurs in children and adolescents with approximately 5% of those less than 15 years of age (9).

Adults with childhood-onset SCI may experience acceleration of the aging process because of diminished physiologic reserves and increased demands on functioning body systems. Secondary medical conditions are extremely common in adults with pediatric-onset SCI and can negatively impact a satisfying adult life. The most common conditions include pneumonia, pressure ulcers, and genitourinary issues (10). Long-term management of adults with SCI also needs to include ongoing evaluation of psychosocial issues such as life satisfaction, living situation, employment, and education.

Problem-Based Approach to Management

Pulmonary

Respiratory complications are the leading cause of death in persons with spinal cord injuries. It has been found that 72.3% of deaths from respiratory causes are due to pneumonia. There is a well-documented association between increased risk of pneumonia and higher level of SCI (11). Inspiratory capacity is decreased in persons with higher level injuries, which can lead to microatelectasis, dyspnea with exertion, and (in those with more severe impairments) respiratory insufficiency. Muscles of expiration can be impaired, which can impact cough effectiveness and clearance of secretions with susceptibility to lower respiratory tract infections (12). The high incidence of sicoliosis in childhood-onset SCI may influence pulmonary function as adults, particularly when curves are large and not surgically corrected. Where appropriate, respiratory muscle strength training, cough-assistive devices, and noninvasive ventilation need to be considered. Vaccinations against influenza and Streptococcus pneumoniae are recommended. Smoking should be strongly discouraged due to the potential adverse affect on respiratory function.

Bladder

Renal failure was, in years past, the leading cause of death in persons with SCI, but this is no longer the case due to significant advances in urologic management. However, urinary tract infections continue to be a high cause of morbidity. UTIs have been reported in 74% of adults with childhood-onset SCI (13). Greater neurologic impairment has been shown to be a risk factor for UTI in both childhood- and adult-onset SCI (13,14). UTIs may lead to bladder and renal calculi, pyelonephritis, and vesicoureteral reflux. The use of prophylactic antibiotics has not been well supported (15). Nephrolithiasis is a common problem in persons with SCI and can contribute to renal failure. Risk factors that can contribute to UTIs and resultant stone formation are vesicoureteral reflux, chronic catheterization, detrusor-sphincter dysfunction, immobilization hypercalciuria, and prior kidney stones (16). The incidence of bladder cancer in SCI is 16 to 28 times higher than that of the general population. Risk factors for bladder cancer are chronic indwelling catheters, smoking, and renal and bladder stones. Gross hematuria should raise concern for bladder cancer (17).

Routine screening of the urinary tract is necessary to decrease morbidity, but there is a lack of consensus on the surveillance tool to use. Most physicians recommend a yearly renal ultrasound for evaluation of the upper tracts, whereas more than half routinely use urodynamic studies for evaluation of the lower tract (18). A survey of physicians found that the combination of clean intermittent catheterization (CIC) plus anticholinergic agents is the favored method for hyper-reflexic bladder management, whereas CIC alone is preferred for the areflexic bladder (18).

Bowel

There is a high prevalence and wide spectrum of gastrointestinal complaints in SCI.

Bowel incontinence is a commonly reported complication in adults with childhood-onset SCI (13). Bowel dysfunction tends to increase with duration and severity of SCI (19). Chronic constipation can lead to megacolon, which has been found more commonly in those with longer duration of injury and older age (20). Sixty percent of adults with childhood-onset SCI report a bowel program exceeding 30 minutes (21). Thus, incontinence and prolonged bowel program duration can greatly impact QoL.

Skin

Pressure ulcers are one of the most common secondary conditions experienced by adults with childhood-onset SCI and have been reported in 44% of persons (13). They are associated with significant morbidity and have great impact on adult outcomes including perceived health status, community integration, and life satisfaction (22). They are also one of the leading causes of rehospitalization (11). The financial burden of such a hospitalization can be tremendous. Many studies have shown that prevalence of pressure ulcers increases with time postinjury. Those with complete injuries, previous pressure ulcers, and older age tend to have a higher risk (23–25). Adults who have sustained an SCI as a child are more likely to develop hip subluxation, which could result in a pelvic obliquity and predisposition to pressure ulcers (26). Individuals need to be vigilant about skin checks and pressure relief. Seating systems and orthotics need ongoing monitoring and need re-evaluation in the event that a new pressure sore develops.

Pain

There are many potential sources of pain in adults with chronic spinal cord injuries. According to Vogel and colleagues in 2002, pain in adults with childhood-onset SCI has been
reported as high as 69% (22). Adults with SCI report pain of a musculoskeletal nature most frequently, followed by neuropathic and lastly visceral (27). Neuropathic pain has been found to be less common in persons less than 20 years at the time of injury (28). Many studies have shown that persons with SCI and associated pain have decreased life satisfaction (29,30). Thus, it is imperative that clinicians investigate and manage pain appropriately.

Musculoskeletal pain is common, particularly in the shoulders. Vogel et al. found that 48% of adults with childhood-onset SCI report shoulder pain (22). It has been shown that prevalence and intensity of shoulder pain in wheelchair users is significantly higher in persons with tetraplegia (31). Wheelchair users may be more prone to pain due to the repetitive use of their upper extremities during wheelchair propulsion, transfers, and self-care. There is also evidence that the prevalence of shoulder pain increases over time (32). A home exercise program of stretching and strengthening has been shown to be effective in decreasing pain and improving function and satisfaction in wheelchair users (33).

Peripheral nerves can be a source of pain. Carpal tunnel syndrome has been reported to be the most common upper limb neuropathy, and sciatic neuropathy the most common lower-limb abnormality. Persons with tetraplegia developed more frequent peripheral nerve lesions than paraplegics (34).

Pain can result in increased spasticity which can further increase pain. Management needs to focus on the underlying cause, but pain may improve with reduction in spasticity such as with antispasticity medications or focal injections.

Management includes evaluation of ergonomics and biomechanics of functional tasks, equipment, and lifestyle in addition to traditional interventions such as medications, manual medicine, trigger point injections, physical therapy, and tone management.

**Neurologic**

A syringomyelia must be considered when there is progressive neurologic deterioration after a stable period. It has been found to occur more often in those with complete injuries and is often associated with arachnoiditis (35). Clinical symptoms can include pain, weakness, sensory disturbance, and autonomic dysfunction. It can occur months to years after SCI. Referral to neurosurgery should be made for surgical considerations.

**Autonomic Dysfunction**

Autonomic dysreflexia (AD) is a potentially life-threatening complication of SCI in which uncontrolled sympathetic outflow occurs in those with lesions above T6 in response to a noxious stimulus such as bladder or bowel distension. The hypertension and compensatory bradycardia that result can be life threatening if not promptly recognized and treated. AD has been reported in 42% of adults with childhood-onset SCI (13). This is higher than that reported in adult-onset SCI (23). Women need to be educated about the risk of AD during pregnancy and nursing. Most can have successful vaginal deliveries but need appropriate monitoring and epidural anesthesia early in labor (36).

**Osteoporosis**

Fractures may be asymptomatic in persons with SCI and impaired sensation may delay treatment. Complications including impaired fracture healing, pressure sores, infection, and osteomyelitis present an additional cause of morbidity.

Osteoporosis is a well-known consequence of SCI. Fracture prevalence has been reported to increase with time, from 1% in the first 12 months to 4.6% in individuals greater than 20 years postinjury (37). As with other disabilities, fractures are frequently associated with only minor trauma. Bone loss occurs rapidly after an SCI with the majority of loss occurring in the lower extremities. Trabecular bone is more affected than cortical bone in the SCI population. There is dissociation between bone mineral density (BMD) in the hip and the spine. It has been found that there is a relative maintenance of BMD in the lumbar spine with significant demineralization of the proximal femur (38). This is consistent with the finding that fractures occur primarily in the lower extremities and not in the vertebral bodies.

The accrual of bone mass is accelerated during puberty, and this process is interrupted in childhood-onset SCI (39,40). In a study of adults with childhood-onset SCI, DXA studies showed that the BMD at the lumbar spine was normal but greatly decreased at the proximal femur region. Also, persons with higher level injuries (C2 to T6) had higher BMD (40). This reinforces the need for aggressive prevention and treatment of osteoporosis early on.

DXA is the gold standard for evaluating BMD in children and adults. If possible, DXA scans should be done using the same machine each year to avoid variations in test results caused by different equipment. Adequate intake of calcium and vitamin D is important. The use of alendronate has been shown to have a positive effect on BMD in persons with SCI, but more research is necessary (41). Functional electrical stimulation (FES) has not been shown to change BMD significantly (42). Consultation with an endocrinologist experienced in the management of metabolic bone disease can be extremely helpful.

**Cardiac**

Age is clearly an important risk factor for cardiovascular disease in the nondisabled population as well as those with SCI. However, it has also been shown that total cholesterol and LDL are higher while HDL is lower in those with increased duration postinjury (43). It has been postulated that SCI bestows an additional cardiac risk over that with age due to enforced sedentary lifestyle and that this increases with duration post injury (44).

With cardiovascular disease becoming a leading cause of mortality in SCI, discussing modifiable risk factors such as activity level, weight control, smoking, and alcohol consumption is extremely important.
Depression/QoL
Depression is a significant concern in adults with pediatric-onset SCI as depressive symptoms have been reported in 27% (45). Depression has been associated with lower QoL, less community participation, and more medical complications. When evaluating life satisfaction, there appears to be no significant difference between men and women, but those with tetraplegia are significantly less satisfied than those with paraplegia (46). Fortunately, for most, perceived QoL among persons with long duration spinal cord injuries is high and relatively stable over time (47).

CEREBRAL PALSY
Background
CP is the most common form of chronic physical disability of childhood. CP refers to a heterogeneous group of disorders that describe impairment in the development of posture and motor control as a result of a nonprogressive lesion of the developing central nervous system. No assumptions should be made with regard to cognitive ability based on motor functioning alone. Many adults with CP live and work independently in the community and lead full, productive lives (48). However, motor performance can deteriorate once adulthood is reached due to the impact of various secondary conditions. It has been found that follow-up with health and rehabilitation services is dramatically reduced once adulthood is reached (49).

The life expectancy in CP is reduced depending upon the severity and other associated health conditions. Several studies have demonstrated that survival is influenced by the presence of severe mental retardation and reduced mobility (50–52). The key predictors were lack of basic functional skills including mobility and feeding. Adults lacking these skills had much reduced life expectancies, as short as 11 years for the worst functioning groups. By contrast, survival of high-functioning adults was close to that of the general population. Life expectancies of adults of a given age can differ by 40 years or more, according to their functional level (51).

In the modern era of molecular genetics, any child or adult with a diagnosis of CP should be considered a candidate for periodic evaluation to better describe their underlying condition. The time of transition from pediatric medicine to the adult health care system is an ideal time to relook at the individual from a diagnostic perspective.

Problem-Based Approach to Management
Pain
Pain has been recognized as a serious secondary problem in adults with CP. Chronic pain has been reported in 30% to 67% of adults, with back and lower extremities being the most commonly affected sites (53,54). Adults with dyskinesia tend to report more cervical pain (55). Fifty percent of adults report pain in more than one body part (48). It is plausible that improper biomechanical forces due to spasticity, contractures, and physical stress could lead to overuse injuries, but this has not been widely studied in CP (55). Several studies have documented the impact of chronic pain on the lives of adults with CP. Many adults experience limitations in their activities, social withdrawal, decreased self-esteem, depression, loss of roles, relationship strain, and emotional distress that leads to a cycle of despair and hopelessness (56,57). Deterioration in function can also be seen as a result. Assessment of pain can be challenging in those with severe cognitive impairment; thus, changes in behavior, posture, spasticity, or range of motion should heightened awareness.

Spasticity has been shown to play a significant role in the development of chronic pain in persons with CP. Spasticity can result in muscular and joint pain, and spasticity often increases in response to the pain. It has been found to be associated with higher rates of osteoarthritis, dislocation, pain, and pressure ulcers (58). Tone management including antispasticity medications, botulinum toxin, phenol neurolysis, and intrathecal baclofen can be helpful. In a study of adults with spasticity, including CP, botulinum toxin combined with therapy improved function and activities of daily living (ADL) in up to 91% while pain was improved in 90% (59).

Treatment of pain is challenging and it has been found that the majority of adults with CP and chronic pain do not seek help from health care providers for pain management (60). One study demonstrated no significant change in pain intensity over 2 years despite an increased use of varied pain treatments (61). Often, once an adult develops chronic pain, there may be few treatments that will be effective (56). Multidisciplinary and cognitive behavioral therapies have proven beneficial to patients with CP-related chronic pain. Coping in this population has been shown to improve when the training is focused on teaching and encouraging patients to learn to maintain their daily tasks despite pain (56).

Musculoskeletal
It has been suggested that adults with CP may show musculoskeletal or performance changes typical of advanced aging earlier than their nondisabled peers (55). Disabilities that have a long duration can produce excessive wear and tear on the muscular, skeletal, and other body systems.

Contractures are common in adults with CP and may significantly contribute to joint pain, sitting intolerance, difficulty with ambulation and transfers, and pressure sores. Eighty percent of adults report contractures with 33% of these occurring in 2 to 3 joints (48). Hip adduction and flexion contractures and windswept deformities create pelvic obliquities and tend to be associated with hip pain as well as problems with perineal care as reported by Noonan and colleagues. Subluxed or dislocated hips can remain problematic in adults with CP. By adulthood, severity of progression can preclude many interventions that may have been suitable early on, but what, if any, management is needed remains debatable (62). The presence of a unilateral hip dislocation often creates a pelvic obliquity that can result in sitting imbalance, hip pain, and pressure ulcers. The reported incidence of hip pain in adults with varied severity of CP is 18% to 50% (58,62,63). It is
plausible that those with less severe involvement may be able to report pain more consistently and accurately. Radiographic evidence of osteoarthritis has been found to strongly correlate with dislocated and subluxed hips, but pain is not necessarily an associated factor (58). If a dislocated hip becomes painful in adulthood or adduction contractures interfere with ADLs or sitting, a proximal femoral resection can be performed with a good outcome (62). Total hip arthroplasty for painful, refractory, subluxed, or dislocated hip has been shown to relieve pain and improve function in 94% of adults with various severities of CP, but its lifelong efficacy is not known (64).

Scoliosis is a significant problem for many youth and adults with CP. Severe curves can be associated with pelvic obliquity, pressure sores, and functional decline (65). Few studies have evaluated the natural history of scoliosis in adults with CP, and it has been shown that there is a high risk of progression with curves of 40 degrees at skeletal maturity (65,66). Average progression rate has been shown to range from 3 to 4.4 degrees per year, depending on function (65). Cardiopulmonary compromise can also occur and is associated with severity of curve and age (67). Surgical correction in these adults may improve respiratory function but outcomes are variable (68). Close surveillance of pulmonary function is needed, and it is important to continually monitor seating, skin integrity, and comfort.

Cervical myelopathy is a potentially devastating secondary condition in adults with athetoid CP. Cervical disc degeneration tends to occur at a younger age and may be attributable to abnormal excessive neck motion in this population (69). Awareness should be heightened with onset of neck pain, significant change in function, or change in neurologic exam. Surgical treatment may prevent further, often catastrophic loss of function, but recovery may not be complete. Short-term surgical outcome is favorable; neurologic deterioration or development of kyphosis can occur greater than 5 years and occasionally greater than 10 years postoperatively (70). Long-term follow-up is essential and postoperative management planning should accommodate change in functional capabilities and care needs.

Osteoporosis
Many children and adults with CP have decreased BMD and a propensity to fracture with minimal trauma (71,72). Reduced BMD is associated with prior fractures, use of anticonvulsants, feeding difficulties, more severe involvement, and increasing age (71). Use of calcium and vitamin D should be standard in those at risk for immobility-related osteoporosis. Studies on pamidronate in children and adolescents with CP have shown benefit, but long-term effects are not known (73,74).

Bowel/Bladder
Constipation and decreased colonic motility have been shown to be related to ambulatory function in children (75), and it is plausible that these issues persist into adulthood. An increased incidence of gastrointestinal complaints has been reported in adults with CP (55,76).

A number of the adults have urinary complaints due to difficulties with toilet accessibility and possible neurogenic bladder (55). Those with recurrent UTIs or incontinence tend to have abnormal urodynamics, and these findings increase with higher Gross Motor Function Classification System (GMFCS) level (77).

Function
Aging and development of secondary conditions can impact motor function and ADLs in adults with CP. It has been shown that the GMFCS level observed around age 12 is highly predictive of adult motor function (78). By age 25, improvement in ambulation is unlikely and decline more likely (79). For those persons who are ambulatory when they reach adulthood, there is a significant decline in ambulation, especially in late adulthood, and few of the 60 year-olds who walk well maintain this skill over the following 15 years. Older persons often lose the ability to dress themselves. Speech, self-feeding, and the ability to order meals in public are often preserved. A majority of young adults live in their families’ home or in small private group homes, while only 18% of the 60 year-olds live independently or semi-independently, and 41% reside in medical care facilities (80).

SPINA BIFIDA
Background
Neural tube defects are the second most common type of birth defect with a birth incidence of 1:1,000 live births (81). Neural tube defects can range from spina bifida occulta to anencephaly. Spina bifida occulta refers to a posterior element defect of the spine covered by skin and no protrusion of neurologic elements. Generally, there are minimal neurologic deficits. Myelomeningocele is a defect in the posterior spinal elements with neural structures found within the thecal sac leading to significant neurologic deficits. Neural tube defects are complex traits with multifactorial etiology encompassing both genetic and environmental components (82).

Health management for adults with spina bifida has gathered increasing attention. The Spina Bifida Association of America has led an effort toward better education of patients and health care providers working with this patient population. Improved management of neurogenic bladder dysfunction and hydrocephalus has increased the survival into early adulthood to 75% (83,84).

Problem-Based Approach to Management
Bladder
The majority of adults require some type of bladder management program. Sixty percent of young adults in one study reported urinary incontinence regardless of the type of management used (85). This includes intermittent catheterization, use of indwelling catheters, urinary diversion, and artificial sphincters. It has been found that most young adults use intermittent catheterization (86,87). Anticholinergic and alpha adrenergic medications are a mainstay of conservative management, and up to 40% of young adults have reported using pharmacologic treatment (84,87).
Renal failure and its complications remain the most prevalent cause of morbidity and mortality in adults with spina bifida (88). The rate of renal failure in adults with spina bifida has declined significantly due to appropriate intermittent catheterization techniques. Risk factors include chronic urinary tract infections and nephrolithiasis. Peritoneal dialysis can be complicated by ventriculoperitoneal shunts or urinary stomata, but recent studies have shown that having a shunt is not an absolute contraindication (89). Renal transplantation has been shown to have good outcomes (88). Adults with spina bifida and neurogenic bladder have an increased risk of bladder cancer, and this may be influenced by bladder augmentation (90). Careful surveillance is required.

Annual renal function tests, blood pressure, ultrasound, and urodynamics are important tools to monitor urologic function (91). Any change in bladder function should be thoroughly evaluated in the adult patient, as a change could also indicate a neurologic abnormality.

**Bowel**

To some extent, all adult patients with spina bifida have some degree of bowel dysfunction. Young adults tend to report less problems with fecal incontinence compared to urinary incontinence (87). The management of neurogenic bowel changes over time as gastric motility is altered with age. Treatment of bowel dysfunction includes medications, dietary modifications, and surgical interventions such as cecostomy buttons and antegrade continence enema (ACE) procedures (92). It has been shown that those using ACE management reported significantly improved fecal continence compared with conventional management, negating the need for retrograde rectal enema and digital extraction (93). Poor compliance with the bowel program may lead to increased numbers of urinary tract infections, incomplete bladder emptying, improper draining of VP shunts, colon cancer, abdominal pain or cramping, chronic constipation, megacolon, skin breakdown, and limited social/professional experiences (94).

**Skin**

Nearly half of adults have some history of skin breakdown (95). The presence of insensate skin and changes in fat and muscle distribution contribute to the high incidence of pressure ulcers (92). Lower extremities are a common site for ulcers due to poorly fitting orthotics and improper biomechanics created by muscle imbalance and associated foot deformities. Those dependent on wheelchair mobility are susceptible to pressure in the ischial and sacral areas. Frequent pressure relief and as appropriate cushion are imperative, and pressure mapping may be necessary for more complex cases. Surgical closure or even amputation may ultimately be necessary. The high economic burden associated with pressure ulcers and osteomyelitis is significant (96). Burn prevention is also a lifelong issue given insensate skin. Patients should be advised to perform routine skin checks and personal hygiene should be emphasized as this is an important part of maintaining good skin health. Given frequent exposure to medical products containing latex, screening periodically for signs and symptoms of latex allergy is recommended. Adults may have a higher rate of reactions including anaphylaxis (97).

**Neurologic**

Nearly 90% of young adults with spina bifida have a ventriculoperitoneal shunt (83). Lifelong follow-up is necessary in persons with shunts. Even though the incidence of shunt malfunction is greatest during the first year after the shunt is placed, and the majority of shunt revisions occur in the first two decades of life, it has been reported that 4% of emergency shunt revisions are performed in the third decade of life (98). Symptoms may include insidious onset of cognitive decline or chronic headaches in addition to the typical neurological symptoms seen in children (99). Adults should have routine neurologic examinations as well as periodic computed tomography scans. Although shunt revision remains a common treatment, one center has found the use of neuroendoscopic third ventriculostomies to be more successful in adults than children (100).

Spinal cord tethering may occur at any age. Symptoms include any change in neurologic status or bowel or bladder function, increasing orthopedic deformities, or gait changes. In contrast to children, adults may have an inciting traumatic event, pain in sacral dermatomes, and more subtle cutaneous abnormality (95). Evaluation includes magnetic resonance imaging and possibly urodynamics. As in children, imaging alone does not justify detethering as the conus will be low lying in all previously untethered lesions.

Adults with a previously treated tethered cord can present years later with symptoms of Chiari malformation. Several studies have shown a high prevalence of the Chiari/hydrosyringomyelia complex in adults with spina bifida with symptoms including headaches, upper limb weakness, sensory changes, ataxia, and lower cranial nerve palsies. Respiratory and swallowing difficulties are not reported as commonly as in children (86). Evaluation includes magnetic resonance imaging. Treatment with decompression and or shunting should be done early to avoid further neurological compromise (86).

**Musculoskeletal**

Those children with spina bifida who are still ambulatory as teenagers tend to remain so into their adult years (83). It is described that deterioration of ambulation can occur in adults due to change in the neurologic level of lesion, spasticity, knee and hip flexion contractures, low back pain, lack of motivation, as well as major events such as stroke (101). Adults often develop improper biomechanics created by muscular imbalance that can result in hip and knee pain. Charcot joints can also develop due to these imbalances and impaired sensation. In contrast, those who are wheelchair dependent are susceptible to repetitive upper extremity movements associated with wheelchair activities and transfers. One might expect shoulder pain and peripheral nerve injuries to be common in adults with spina bifida, but pain in this population is not as well described as in those with SCI (32). One study found that 56% of children and adolescents reported experiencing pain in varied sites once a week or more often (102).
Osteoporosis

It is known that fracture rate in children with spina bifida is higher than in the general population (103), but it is less known how this translates into the adult years. In adults with spina bifida, neither fracture rate nor osteoporosis has been well studied. A recent study actually showed a higher incidence of fractures in adolescents than adults (104). One study examined the prevalence of osteoporosis and osteopenia in adults with myelomeningocele using DXA and found that 33% had osteoporosis in at least one site and secondary conditions were associated with lower BMD (105).

Treatment includes appropriate calcium and vitamin D supplementation. Data on bisphosphonates are lacking.

Nutrition

Maintaining an ideal body weight is of critical importance in adults with spina bifida when considering prevention of long-term effects of obesity. Complications of obesity include hypertension, diabetes, cardiovascular disease, hyperlipidemia, obstructive sleep apnea, and osteoarthritis. Nutritional counseling should not only include dietary recommendations but also include information on physical activity.

Sexuality

Adults with longer duration of disability tend to report significantly more positive feelings about their sexuality (106). The majority of young adults with spina bifida report being sexually active (107). Data on the impact of sexual function on QoL in this population are inconsistent (107,108). It has been found that urinary incontinence is a negative predictor of sexual activity and that women with hydrocephalus are less likely to be sexually active (107). Sexual health needs to be enforced in women with spina bifida. Annual Pap smears and mammograms are needed although access and attitudinal barriers can make this challenging. Information on contraception and STDs also needs to be provided. Women with spina bifida are able to carry pregnancies to term; however, awareness should be increased as distal shunt malfunction can occur during the third trimester (95). Infertility issues such as erectile or ejaculatory dysfunction need to be addressed by physicians. Erectile dysfunction in spina bifida has been well studied, and while 85% of men achieve psychogenic erections and 88% are able to ejaculate, only 54% experience rigidity at erection (109). Sildenafil (Viagra) has been shown to improve erectile function in 80% of men with spina bifida (110). Counseling regarding the increased risk of neural tube defects in the children of those with spina bifida is important.

NEUROMUSCULAR DISEASE

Background

Patients with neuromuscular disease (NMD) present many complex transition issues and problems. Many issues arise in early childhood and increasingly more children are living into adulthood. Duchenne muscular dystrophy (DMD) is the most common form of childhood-onset muscular dystrophy. It is estimated that DMD affects 1/3,500 to 1/6,000 live born males (111). DMD and Becker muscular dystrophy (BMD) form a spectrum of clinical severity, and together are referred to as dystrophinopathies (112). DMD is an X-linked recessive disorder that is characterized by a progressive weakness of skeletal muscle. The lack of dystrophin, a muscle structural protein, results in an unstable muscle fiber membrane and impaired muscle homeostasis (113).

Problem-Based Approach to Management

Pulmonary

Respiratory disease is a major cause of morbidity and mortality in DMD. Assisted ventilation has had a profound effect on the natural history. Most caregivers have changed to a more aggressive and supportive approach. The American Thoracic Society (ATS) has developed a consensus statement entitled: “Respiratory Care of the Patient with Duchenne Muscular Dystrophy” (114).

DMD is associated with a gradual loss of muscle function over time, leading to ineffective cough, decreased ventilation, and respiratory insufficiency. These complications are generally preventable with careful serial assessments of respiratory function. Routine evaluations should include oxyhemoglobin saturation, spirometric measurements of FVC, FEV1, maximal midexpiratory flow rate and maximum inspiratory and expiratory pressures, and peak cough flow (115).

Effective airway clearance is critical to prevent atelectasis and pneumonia. Cough peak flows correlate directly with the ability to clear secretions, and values below 160 L/min and/or maximum expiratory pressures below 60 cm H2O have been associated with ineffective airway clearance (116). Many techniques have been developed to overcome an ineffective cough in patients with neuromuscular weakness. These include manual techniques such as glosopharyngeal breathing, air stacking, intermittent positive pressure, and mechanical ventilation.

Mechanical insufflator-exsufflators simulate coughing by providing a positive pressure breath followed by a negative pressure exsufflation (117). This device has been shown to be effective and well tolerated with occasional side effects of transient nausea, abdominal distention, bradycardia, and tachycardia. The ATS Committee strongly supports the use of mechanical insufflation-exsufflation.

Persons with DMD have increased risk for sleep-disordered breathing, including hypopnea, central and obstructive apnea, and hypoxemia. Treatment of these pulmonary complications with noninvasive ventilatory support can improve QoL and reduce the high morbidity and early mortality (118). Nasal intermittent positive pressure ventilation with bilevel positive airway pressure generator or mechanical ventilation can be used in the treatment of sleep-disordered breathing and nighttime hypoventilation (119). Oxygen supplementation is not recommended without ventilatory assistance. Daytime ventilation should be considered when waking PCO2 exceeds 50 mm Hg or hemoglobin saturation remains less than 92% at sea level while awake (114).
Full-time ventilation can be provided when other device interfaces are poorly tolerated or the patient lacks sufficient oromotor and/or neck control to use a mouthpiece. The advantages of a tracheostomy include a more secure interface, ability to provide higher ventilator pressures in patients with intrinsic lung disease, or severe reductions in chest wall compliance. The potential complications include increased secretions, increased aspiration risk, and infection.

Cardiac
Cardiac involvement is the second most common cause of death in patients with DMD, with 10% to 20% dying of cardiac failure. The heart is almost always affected by progressive cardiomyopathy. All patients should have regular cardiac evaluations with annual electrocardiograms and echocardiograms (120). Dilated cardiomyopathy primarily involves the left ventricle. Right ventricle failure can result from respiratory failure and pulmonary hypertension. The patients are also at risk for ventricular arrhythmias.

Initiating therapy at the presymptomatic stage of left ventricular dysfunction is known to delay the onset of cardiac failure and improve prognosis. Combination therapy with an angiotensin-converting enzyme (ACE) inhibitor and nonselective beta-blocker is recommended (121). There are some retrospective data suggesting that cardiac involvement is less frequent in patients treated with deflazacort (122).

Nutrition
Malnutrition and obesity are common problems in young adults with DMD. Regular evaluations of ideal body weight and body mass should be provided by a nutritionist especially if the patient is on steroid medication. Swallowing should be monitored clinically and an evaluation for dysphagia performed if there is a history of choking or dysphagia (123). A video fluoroscopy may assist to prescribe safer swallowing techniques. When adequate nutrition cannot be maintained, gastrostomy tube placement and enteral feedings are highly encouraged.

Osteoporosis
Osteoporosis and fractures are common in patients with DMD especially those on steroid medication (122). Long bone fractures occurred in 38.7% during steroid therapy compared with 26.5% in the nonsteroid group (124). Of the fractures occurring in the steroid group, 84.9% were in the lower extremity. Patients with DMD should have dietary counseling and baseline DXA scanning (125). Calcium and vitamin D supplements are recommended. The role of bisphosphonates and other bone-sparing agents is encouraging (126).

Pain
Pain is a significant problem for most patients with NMD, although it is not typically a direct consequence of the disease (127). Most commonly, the pain is caused by immobility. This may lead to adhesive capsulitis, low back pain, pressure areas on the skin, and generalized myofascial pain (127). It has been found that the frequency and severity of pain reported in slowly progressive NMDs is significantly greater than that reported by the general U.S. population (128). A significant correlation has been shown between increased pain and lower levels of general health, vitality, social function, and physical role (128). Pain assessment needs to be a standard aspect of care.

End-of-Life Planning
End-of-life directives are a critical part of anticipatory care. Respiratory failure can occur suddenly or in association with a prolonged process. Education about ventilatory and palliative options should be discussed prior to ventilatory failure (129). Physicians have an ethical and legal responsibility to disclose all treatment options, including long-term ventilation. There is evidence that health care professionals underestimate the QoL of ventilator-dependent people with DMD. Patient/families choosing to forgo long-term ventilation must receive palliative care and guidance.

OSTEOPORSIS

Osteoporosis is a group of disorders characterized by fractures with minimal or absent trauma, dentingenesis imperfecta, and hearing loss. The spectrum of the phenotype has a broad range. At the most severe end of the spectrum, perinatal death may occur. Other manifestations include individuals with severe skeletal deformities, mobility impairments, and very short stature to nearly asymptomatic individuals with a mild predisposition of fractures, normal stature, and normal life span. Inheritance has been documented to be autosomal dominant or autosomal recessive. Mutations in COLA1 or COLA2 characterize the underlying molecular defect for OI types I to IV. Most mutations are de novo. Adults with OI need to cope with the musculoskeletal concerns associated with OI and also manage the same health issues as other adults (130).

Most adults with OI experience a decrease in the rate of fractures after puberty. Tendon, muscle, and joint problems may be aggravated with time, and hearing loss may become significant (131). Vigorous and consistent medical care remains as important as it was during childhood.

Problem-Based Approach to Management

Osteoporosis
Maintaining bone mass is a priority for adults with OI because fracture risk is, in part, related to bone density. An initial DXA test may be obtained at any age and then yearly thereafter. Bone density can decline as a direct result of OI, from immobilization associated with casts or limited weight-bearing activity, and from age-related changes in bone and the endocrine system. Treatments can include calcium and vitamin D supplements, drug therapies including oral or intravenous bisphosphonates,
diet, and exercise (132,133). There is limited experience with adults and use of IV bisphosphonates (134). Smoking, overuse of alcohol, and certain medications, including cortisone-like steroids, can also negatively impact bone health.

Musculoskeletal
Adults often report pain in their lower back and hips. This can be the result of compression fractures of the spine, scoliosis, or joint deterioration. Other problems can include nonunion fractures and low muscle strength. Knee pain, ankle instability, and rolling in of the ankle are frequent complaints of people with OI. These problems can be a result of joint laxity which subjects knee and ankle joints to pressure over the years. Leg lengths may differ due to a history of previous fractures. Exercise or orthotic devices to improve hip, knee and ankle alignment, and back or joint surgery may provide relief. Heel lifts and firm ankle supports can limit wear and tear on the joints and improve walking comfort (135). Joint replacement surgery may be a treatment option for some, but not all, adults with OI who have joint problems (136).

Pain
Adults may experience pain from old fractures, or compression fractures of the spine related to either OI or osteoporosis (131). Unstable joints may increase degenerative changes, which are the source of pain in many individuals. Pain management may include lifestyle adjustments to protect the spine, medications, and alternative treatments (136). Adults should be wary of increasing the strength of pain medication to the point where it depresses respiration or consciousness. Reduced consciousness can increase the risk of falls and fractures.

Pulmonary
Type III and type IV OI and those individuals with significant curvature of the spine may have restrictive lung disease due to decreased chest volume, chronic bronchitis, and asthma. Rib fractures and muscle weakness also may contribute. Sleep apnea is a related problem for some adults with OI and can be diagnosed with an overnight sleep study. Exercise to promote deep breathing, regular testing of pulmonary function, and use of supplemental oxygen can help manage pulmonary function. BiPAP may help with sleep apnea or related pulmonary insufficiency (82,130). Aggressive treatment of all upper respiratory infections in adults with OI is recommended (131). Asthma and chronic bronchitis may contribute to impaired pulmonary function and should be treated with bronchodilators, inhaled corticosteroids, and antibiotics when appropriate.

Cardiac
A small number of adults with OI seem to have valvular problems. The most common is mitral valve prolapse. Dilation of the aorta also may occur but is rare. High blood pressure is as common among adults with OI as in the rest of the population (131). High cholesterol and related lipid disorders that may occur in families can contribute to cardiovascular disease. Medical management of these disorders includes appropriate diet and drug therapies and regular monitoring by a primary-care doctor. Drugs such as statins can be very helpful along with diet in controlling lipid problems. Coronary artery surgery has been successfully performed on people with OI, although precautions are necessary because of tissue fragility.

Hearing
Approximately 50% of all adults with OI will experience some degree of hearing loss during their lifetime (130). Hearing tests and MRI examination of the hearing canals can help delineate the involvement of the bones of the ear. Treatment for hearing loss usually begins with hearing aids. Some adults are candidates for either stapedectomy or cochlear implant surgery.

Vision
The connective tissue problem in OI can extend to the eyes. Eye examinations are recommended every 2 to 3 years. OI can affect the shape of the lens and the strength of the sclera (131). For this reason, adults with OI should consult with an ophthalmologist before using contact lenses. Also, laser lens surgery is not recommended for people with OI.

Gastrointestinal
Gastric problems are not uncommon in OI. These include gastric acid reflux, which is aggravated by a decreased length of the chest cavity, and chronic constipation. Short stature and frequent use of various pain medications can contribute to the problem. Constipation symptoms should be treated aggressively.

Basilar Impression
Also known as basilar invagination, this is a special problem for adults with Types III and IV OI. Basilar impression (BI) involves pressure from the spinal column on the base of the skull. Symptoms can include headache, muscle weakness, and tingling or numbness of hands and feet (130). Evaluation by a neurologist, including MRI examination of the cervical spine and base of the skull, is necessary. This condition may be progressive with age.

ACHONDROPLASIA

Background
Skeletal dysplasias are a group of genetic disorders resulting in short stature. Achondroplasia is the most common and affects approximately 1/15,000 to 1/25,000 live births. It is an autosomal dominant disorder but 75% of cases represent new mutations. Achondroplasia is caused by a mutation in the gene that codes for the fibroblast growth factor type 3 (FGFR3), which is located on chromosome 4. The average adult
height is approximately 4 ft for women and men (137). Intelligence is usually normal. Life expectancy for achondroplasia is approximately 10 years less than that for the general population (138).

**Problem-Based Approach to Management**

**Cardiac**

It has been found that mortality related to cardiovascular disease is increased more than ten times that of the general population and that males have a higher risk (138). Cardiovascular-related deaths were increased twofold, but associated heart disease risk factors were not significantly increased (138). Screening for cardiovascular risk factors is, therefore, an important part of health maintenance.

**Nutrition**

Obesity is a significant health problem in achondroplasia and has a prevalence of 13% to 43%. It begins in early childhood and occurs at all ages (139). Maintaining an appropriate weight is important given that obesity is a preventable risk factor for cardiovascular disease.

**Musculoskeletal/Neurological**

Musculoskeletal complaints are prevalent in adults with achondroplasia. In one study, almost 50% of those over 20 reported leg pain and 70% of those over 50 reported back pain (140). Deaths from a neurologic cause were increased eightfold and occurred at all ages in a 42-year follow-up study, including sudden death in infants and children from foramen magnum stenosis (138). Lumbar spinal stenosis with neurologic compromise is a common complication in achondroplasia and one third will require surgery. It does not typically develop prior to age 15 and is often associated with increased kyphosis (141). Symptoms include weakness, pain, numbness, and bladder incontinence. Surgical decompression is required if conservative management fails. Postural kyphosis is a common occurrence in children but usually disappears with onset of walking. It may persist into adulthood in more than 10% of persons as a result of structural wedging of the vertebra (142). Posterior spinal osteotomy with instrumentation in those with neurologic deficits has shown to be successful in this population (143).

**Women’s Issues**

Fertility is usually not affected. Women with achondroplasia may have an increased risk of uterine fibroids, therefore, oral contraceptives should not be used long term. All women require caesarian delivery due to small pelvic outlet. Spinal anesthesia is not recommended due to possible spinal stenosis. Premature menopause has been reported (144).

**Quality of Life**

Persons with short stature can experience many social and emotional challenges in an “average-sized world.” Adults with achondroplasia tend to report lower QoL and self-esteem levels (145). Little People of America has played a significant role by providing a network in which information can be shared and support provided to those with skeletal dysplasias.

**CONCLUSION**

Adults with childhood-onset disabling conditions are generally healthy and can be expected to lead meaningful, active lives. Not all adults have serious health problems, and many now recognize the aging process as a natural course of events, as it in the general population. The most common age-related changes and secondary conditions involve physical performance and the musculoskeletal system. Prevention strategies require anticipation of potential secondary conditions, communication of risk, recognition of changes that alter function and require intervention, and an understanding of interventions that have an impact on function. This requires that adults with childhood-onset disabling conditions have access to knowledgeable health care providers, and the medical systems necessary to implement a medical home approach to management of chronic illness and disease. Environmental, communication, attitudinal, and systems barriers must be overcome.

It is time to reconsider the model of illness and disease for persons with lifelong disabilities. Particularly in the realm of mobility, a health and wellness model should be developed. Use of prevention strategies must be considered in childhood and adolescence to address the more frequent secondary conditions. Programs of fitness and exercise have been proven beneficial in nondisabled groups and disability groups alike. Health promotion strategies should be employed for adults with childhood-onset disabling conditions. The inclusion of preventative measures into the model of chronic disease management will further solidify the team approach implied in the medical home concept and ensure that this population receives the high-quality health care that every person deserves.

**REFERENCES**


INTRODUCTION

In increasing numbers, women with disabilities are finding their way into public offices, educational institutions, health care organizations, corporate boardrooms, science laboratories, the arts, and sports arenas. They are defying the old adages of “can’t do” and invisibility and becoming authors and sculptors of their lives. With wisdom and resilience born out of years of hardship, exclusion, and prejudice, women with disabilities are finding their voices and increasingly shaping the world. They are declaring that no longer is it shameful to inhabit a body that is atypical in its anatomy or physiology or to need personal assistance services (PAS) (1). They are challenging us to question the wisdom of the drive toward homogeneity and independence with counter images and enhanced notions of diversity and interdependence as integral girders to the society of the future.

Nowhere are those voices more loudly heard than in the health care setting. Without question, in the last 20 years many social barriers have eroded. Curb cuts, accessible transportation and communication systems, and greater access to public spaces including higher education and the work place are increasingly expected “norms” in the community. Ironically, health care, a critical resource in the lives of many people with disabilities, has been one of the slowest domains to change and adapt (2–5). But the pressure is on. Women with disabilities are no longer accepting second-class health care, tolerant of a denied mammogram or Pap smear because they cannot stand or get onto the examination table. No longer are they willing to endure exclusion from reproductive health care services, based upon the mythology that they, as disabled women, are asexual, childlike, and unlikely interested or able to become wives and mothers (Fig. 58-1).

In this chapter, the phrases “women with disabilities” and “disabled women” are both used. Language is constantly evolving and in disability communities, “… a controversy has raged over preferred linguistic usage” (9). The term “people with disabilities” underscores “the importance of the individual in society and disability as being something not inherent in the person” (9). The term “disabled people” emphasizes minority group status and pride in disability identity. Consistent with this emphasis on disability as a social minority identity and not only a medically defined deficiency, the authors of this chapter have chosen to use these terms interchangeably.

BARRIERS TO HEALTH CARE

People with disabilities are no strangers to barriers (Table 58-1). The most overt barriers are physical such as stairs, narrow doorways, curbs, and inaccessible bathrooms (6,10–12). Another type of physical barrier is inaccessible medical equipment, such as examination tables, bariatric scales, and mammogram machines. Communication barriers include lack of sign language interpreters, materials in Braille, large print, or on computer disk. Programmatic barriers encompass lack of trained physical assistance, inflexible scheduling, and transportation difficulties. Economic barriers also play a significant role in preventing people with disabilities from accessing community services, such as health care (13). The most insidious barriers, however, are erected by ignorance and negative social attitudes
about life with disability (9). Women with disabilities, in particular, are disproportionately affected by discriminatory practices in employment, education, vocational services, economic programs, access to benefits and services, health care, and parenting activities (14,15). With the passage of civil rights laws such as the Americans with Disabilities Act (ADA) in 1990, and expanding clinical services targeting the needs of women with disabilities, this situation is beginning to improve (16,17). We are a long way, though, from full integration, where a woman with a disability can go to a community health center with the expectation that it will be fully accessible with wheelchair-adapted equipment and knowledgeable staff trained to assist women with a variety of disabilities in a manner respectful of their womanhood (Fig. 58-2).

The unequal representation of women in health care research is another barrier to health care facing women with disabilities (18,19). Historically, women have often been excluded from medical research for a variety of offered reasons ranging from methodologic issues about the menstrual cycle to liability concerns related to potential pregnancies (20). As a result, we have had little information about the causes and most effective treatment for major diseases of women such as coronary artery disease, the number one killer for both genders (21,22). The Women's Health Initiative of the National Institutes of Health, launched in 1991, was a major effort to address these deficiencies with a 15-year commitment of funding for large multicenter studies on topics such as breast and colon cancer, osteoporosis, and heart disease (23). Unfortunately, women with disabilities were not identified as a subpopulation for study; thus, despite the richness of the information, it did not help answer particular questions about the similarities or the distinctions of these health issues in the context of disability.

### DEMOGRAPHIC FACTORS

Based on the American Community Survey conducted by the U.S. Census Bureau in 2007, it is estimated that 15.5% of females over the age of 5 are disabled, 6.5% have one type of disability, and 9.0% have two or more types of disability (24). These figures are for civilians and noninstitutionalized persons. When considering the entire U.S. population, estimates are likely higher, with the most recent population-based data indicating that 21.3% of the female population has a disability, slightly higher than men (19.8%). They also tend to have more activity limitations than men with disabilities (25). Women with activity limitations are less likely to be married compared to other women and men with or without activity

<table>
<thead>
<tr>
<th>TABLE 58.1 Access Barriers</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Physical</td>
</tr>
<tr>
<td>• Environmental</td>
</tr>
<tr>
<td>• Equipment</td>
</tr>
<tr>
<td>• Communication</td>
</tr>
<tr>
<td>• Programmatic</td>
</tr>
<tr>
<td>• Financial</td>
</tr>
<tr>
<td>• Attitudinal</td>
</tr>
</tbody>
</table>
limitations. There is evidence that severity of disability plays a role. Only 44% of women with “severe” activity limitations (unable to carry on a major activity for their age group such as work, independent living, attending school, etc.) are married, and these women have higher rates of divorce than women with less severe limitations (26). About a third of women with disabilities have children at home, compared to about a quarter of men with disabilities (27).

In general, women’s education level is inversely related to degree of disability (25). Women with disabilities are more likely to live in poverty, be unemployed, or if they are employed make less than either women without disabilities or men (Table 58-2). Women with disabilities are as likely as nondisabled women to have health care coverage but are more likely to have publicly funded coverage (as opposed to private insurance) than nondisabled women (13).

**HEALTH DISPARITIES**

The burgeoning literature on health disparities underscores that women with disabilities are subject not only to economic barriers but also to the lack of access to health care (2,5,28). The data on access is not straightforward, however. For example, in a recent analysis of a national probability sample (29), researchers examined health care access using eight measures including sources of care, insurance status, satisfaction, Papanicolaou test, and breast exam. They conclude that, “although the women with disabilities had similar or better potential health care access than nondisabled women, they generally had worse realized health care” (29). More specifically, women with disabilities were more likely to postpone care and medications and were less satisfied with their medical care than women without disabilities, although they had higher rates of a “usual source of health care.” In a longitudinal study of medical expenditure survey data (30), disabled women were less likely to receive Pap smears, mammography, and other cancer screening services but more likely to receive influenza immunization, and colorectal and cholesterol screening. Furthermore, there were significant socioeconomic and racial differences in access. The authors conclude that disability is a barrier to preventive clinical services and “a key issue for improving women’s health care is to identify those who are at risk for specific measures of preventive care and also recognize subgroup disparities in care” (30).

These findings regarding health disparities in reproductive health care screening (e.g., Pap smears and mammograms) have been replicated by others including a study of women with major mobility impairments who were compared to their nondisabled counterparts (31). In addition, it appears that the severity of disability inversely correlates with the likelihood of receiving Pap smears and mammograms for women with physical disabilities (32). Furthermore, disparities for women with severe impairments appear to be exacerbated by a variety of factors including lack of knowledge on the part of health care providers, systems issues, ineffective communication, and lack of access to health information and services (33).

There is growing evidence of health disparities in the treatment phase of an illness and not just in health care screenings. In a retrospective observational study of women with stage I to IIIA breast cancer, those with disabilities were much less likely than other women to receive radiotherapy following breast conserving surgery than nondisabled women (34). Many reasons are posited for such disparities, ranging from absent or inadequate health insurance, patient preferences, lack of physical access or transportation resources, and attitudinal bias. Further research is needed to uncover the causes for these health disparities.

**REPRODUCTIVE HEALTH CARE**

Women with disabilities overwhelmingly report difficulties obtaining balanced information about reproductive health care issues, techniques for managing menstruation, birth control, risks associated with pregnancy, techniques for labor and delivery, information about sexual functioning, dating, gender identity, etc. (1,35,36). Consistent with the rehabilitation model, it is our recommendation that the physiatrist works with a team of HCPs, including obstetrician/gynecologists,
urologists, anesthesiologists, and allied health professionals, who are committed to delivering knowledgeable, respectful, and accessible reproductive health care services to women with disabilities. What follows are some general suggestions and tips to help optimize the access to needed reproductive and preventive health services for women with disabilities.

Preventive Health Care Services: Pap Smears and Gynecologic Examinations

Basic preventive health and gynecologic screening for women with disabilities is often overshadowed by more obvious physical or neurological problems, which may require more immediate focus (37). Thus, screening for diabetes, hypertension, hyperlipidemia, and thyroid imbalance, all of which are common concerns for women, may be neglected in women with disabilities. Many women avoid gynecologic care because of difficulty in obtaining an accessible, comfortable, and dignified examination (1,36,38). As a consequence, treatable early-stage problems may escalate and become much more difficult to manage. These concerns include Pap smear screening as well as breast evaluation.

Performance of the pelvic examination must be tailored to a woman’s physical impairments often through the use of creative positioning techniques (39,40). An accessible examination table that lowers to wheelchair height and has security features such as hand rails, boots, and straps can be indispensable (39). Leg adjustments should be performed slowly and gradually to minimize pain and spasticity. Liberal application of lidocaine gel to the perineal area can be helpful in minimizing spasticity or in preventing episodes of autonomic hyperreflexia (AH) in some spinal cord injured women with high lesions (above T6) (41).

Preventive Health Care Services: Breast Self-Exam and Mammograms

As of 2004, breast cancer afflicted 126.4 per 100,000 women/year, making it the most prevalent gynecologic cancer occurring in women (42). Currently, it is estimated that one in eight women will develop breast cancer. As a screening tool, mammography detects 80% to 90% of breast cancers. While there is general agreement that all women of age 50 and older should have yearly screening mammograms, there is still no clear consensus regarding women between the ages of 40 and 49 (43). Women with disabilities should, at the very least, obtain screening mammography at rates comparable to nondisabled peers.

Unfortunately, many facilities utilize equipment unable to easily accommodate mobility-impaired women. Though universal design mammography equipment is available, not all mammography centers or technicians are knowledgeable about how to accommodate women with disabilities. There are ongoing efforts to educate breast health providers about how to ensure that women with disabilities are accommodated (44–46). Women with disabilities should ask their health care provider for the names of facilities in their area best able to accommodate their disability and should contact the center in advance to notify them of any special needs they might have. Although many authorities dispute the value of the breast self-exam and have muted prior calls for its monthly performance (47), many women informally check their breasts after their menses or at a regular interval in order to report any concerning changes (48). Women with disabilities, particularly those lacking manual sensation or dexterity, might need to rely on a partner or personal attendant to assist with this examination.

REPRODUCTIVE AND HEALTH ISSUES FOR ADOLESCENT GIRLS AND YOUNG ADULTS

Management of Menstruation

Menarche is a symbolic moment in most women’s lives, marking the transition from “girlhood” to “womanhood,” with the attendant procreative possibilities. For young girls growing up with their disabilities, it can also bring a host of stresses—from the practicalities of managing menstrual hygiene to the parental anxieties and concerns raised by menstruation (49–52). Parental discomfort can encompass a wide range of issues: the need to assist with menstrual hygiene; concerns about menstrual cramps and painful periods (particularly for those daughters with intellectual or communicative disabilities); maturation and secondary sexual characteristics of their daughters with awareness of vulnerabilities for sexual abuse; interest in dating and sexuality; and potential pregnancy. The recent Ashley X case, which involved a novel growth attenuation treatment for a young girl with “static encephalopathy,” also raised the issue of the parental concerns about their daughter’s future care, particularly as they contemplated the maturation of her body from a lightweight young girl easily portable by family members to an adult woman’s body (50–52). The report of the growth attenuation treatment sparked much controversy about whether a treatment should be offered to both surgically and hormonally alter the body of children with developmental disabilities in order to make their care easier (with the belief that such an action will improve the person’s quality of life) (50). A number of disability activists and organizations have subsequently come out with strong statements against the practice (53–55).

When HCPs are faced with parental requests for sterilization or hysterectomy as a means of managing their minor child’s menstrual hygiene or concerns about future pregnancy, it is helpful to have some resources and guidelines. In general, the indications for hysterectomy and sterilization should be the same as for the nondisabled population, and other less invasive interventions should be considered to address the identified problem. The American College of Obstetrics and Gynecology Committee have issued an opinion on “Sterilization of Women, including Those with Mental Disabilities” (56), and present guidelines for consideration of these requests for both women with capacity to consent and those with cognitive disabilities who are not able to provide adequate informed consent. Given the permanency of the procedures and the possibility of coercion and abuses, many
states require that the courts be involved if a decision is made to proceed with sterilization of a minor child or woman with cognitive disability. Though the hysterectomy in the case of Ashley X was not done with the primary intent of sterilization, the case was investigated by Washington Protection and Advocacy (57). In their report, Washington Protection and Advocacy laid out some ground rules for future decision making around similar cases which included:

- Need to get a court order/permission before doing a sterilization procedure or growth attenuation treatment in a minor child with disability
- Need to have a disability advocate involved in ethics decision making around such cases

Though the ruling is applicable to the state of Washington only, it is important to know the laws in the particular state of residency regarding sterilization procedures of minor children or disabled adults who lack decision-making capacity.

When possible, HCPs need to proactively address these issues with both girls growing up with disabilities and their parents, introducing options for managing menstrual flow if possible, before the onset of menstruation (58–61); such discussions offer opportunities for girls with disabilities to develop a sense of control over their emerging sexuality and evolving images of themselves as women. For parents of girls with disabilities, the discussions offer opportunities to address their fears and concerns with balanced information, creative options, and the perspectives of the lived experience of women who grew up with their disabilities.

For women with acquired disabilities who are still menstruating, management of menstrual flow should be addressed shortly after the onset of disability, preferably as a part of a rehabilitation program. Some women may be able to work with a nurse or occupational therapist to develop a system for managing menstrual hygiene. Other women may find that switching to sanitary pads if they had previously used tampons may be all that is needed. Still other women may elect to work with a personal assistant to manage their menstrual hygiene. Some women, particularly those with more extensive physical disabilities, may find these options impractical and choose to look for pharmacological options (i.e., hormonal interventions) to regulate or curtail the menstrual flow (62–67). While there are a number of options used by women in general, there is little data on the safety and usage of these treatments in women with disabilities. Unfortunately, some physical disabilities and chronic disease states are inherently associated with menstrual irregularities, leading exactly to the unpredictable flow that is so unwelcome.

**Access to Information**

It is estimated that girls with disabilities (age 21 and younger) comprise 8% of female youth (25). While girls and young women with disabilities share similar health care issues with their nondisabled peers, their health encounters are likely to be more frequent, intense, and complex (68). In addition to routine adolescent health needs, disabled girls may encounter a regime of disability-specific repeated hospitalizations, health visits, and medication schedules that can intrude into their family, school, and social lives. Sensitivity to the medicalization of these girls’ lives is important and should be demonstrated by actively engaging the girls in the decision-making process as it affects their health issues. This can be accomplished in partnership with a parent or caregiver as long as the girl is not coerced into unwanted decisions and supported with information in a user-friendly format that allows her to be self-determining.

The promotion of self-determination in girls and young women with disabilities serves to decrease the incidence of negative health encounters while optimizing the likelihood of compliance with healthy life choices.

Lack of access to information can prevent girls and women with disabilities from achieving self-determination in health and other aspects of life. Research suggests that "adolescents with disabilities are likely to be participating in sexual relationships without adequate knowledge and skills to keep them healthy, safe, and satisfied," and are twice as likely to experience sexual abuse (69). Their exclusion from sources of information about sexuality and reproductive health options has been assisted by inadequate disability accommodations, negative disability attitudes, or a combination of both types of barriers (70). Many girls with disabilities grow up outside the usual channels for learning about sexuality. For example, they may be denied admission to their neighborhood schools and recreation programs due to access barriers and discriminatory enrollment practices (71). Consequently, they are denied important informal opportunities to exchange information and experiences regarding sexuality with their peers. Instead, they might be placed in a sheltered and closely supervised “special education” school where even formal sex education may not be offered. Their parents are less likely to talk to them about sexuality and reproduction than to their nondisabled sisters (68).

Girls and young women who are blind are rarely given tactile anatomical models so they can safely explore and understand the form and function of human body parts. Those who are deaf or hard of hearing may miss conversations about “the facts of life” whispered between hearing friends. Girls and women with cognitive disabilities may not fully understand spoken and written information about sexuality and health matters unless it is presented in an illustrated, uncomplicated manner. There are some materials that are helpful in preparing girls and women with disabilities for their first pelvic and breast exams, such as the video and workbook “Let’s Talk About Health: What Every Woman Should Know” designed for girls with developmental disabilities (72).

In adulthood, information barriers interfere with health maintenance for women with disabilities (6,11). Blind women and those with low vision confront clinic signs, instructions, release forms, and educational materials that have not been translated into accessible formats, such as Braille, audiotape, or computer disk versions. Deaf women are not offered interpreters to insure accurate doctor/patient communication. Women with cognitive disabilities are not allowed adequate time and explanation to ensure comprehension and informed
With others (relatedness)” (73). For women with disabilities, goals (competence), and to form cooperative relationships. Sheldon (74) focused on three psychological traits, characteristic adaptations, life narratives, and culture. Specifically, there are some universal, cross-cultural, and historical aspects of human life which have evolved over time such as adaptation and cooperation. Sheldon (74) focused on three psychological needs: “need to sustain a basic sense of self (autonomy), to manipulate the environment in order to achieve instrumental goals (competence), and to form cooperative relationships with others (relatedness)” (73). For women with disabilities, the issues of autonomy (or self-determination), competence, and relatedness are as fundamental to their flourishing as they are for health care providers and family members who each bring their own perspectives to the medical encounter.

The second principle is the importance of dispositional traits. Dispositional traits include what has classically been referred to as the big five factors of personality: openness, conscientiousness, extraversion, agreeableness, and neuroticism (75,76). Debates within personality psychology have focused on dispositional traits. The concept of self-esteem has the qualities of a personality trait and a psychological state (78) and refers to an appraisal of one’s worth. Other terms such as self-regard and self-respect have also been used in the literature. Recently, scholars have focused on specific aspects of self-esteem. For example, Crocker et al. have defined “contingencies of self-worth” or domains from which individuals derive their perception of self-worth. They describe the domains of virtue, God's love, support of family, academic competence, physical attractiveness, gaining others’ consent. There is also evidence that physicians’ perceptions of a woman's disability may result in the withholding of information about her health options. One of the first systematic studies to document disabled women's health service experiences (38) revealed that physicians’ inclination to offer disabled women patients information regarding sexuality and contraception was mediated by such factors as the age of onset and extent of a woman’s disability.

Some reproductive health agencies, such as Planned Parenthood and Regional Family Planning, have taken important steps to ensure that their information and services are accessible to women with disabilities. State or city government offices specializing in disability services can guide physicians in providing accessible information and communication to patients with disabilities. A useful rule of thumb in working clinically with any woman with a disability, particularly early onset disability, is to ask if she would like information or guidance about sexuality, reproductive health options, or any other aspect of health. Some women with disabilities are sophisticated about their health options and the workings of their bodies; others lack basic information about anatomy and human sexuality. Many are experts regarding their particular disability but have gaps in general health maintenance knowledge due to the information barriers discussed above. It is important to explore with each woman her level of knowledge and desire for more information before concluding any health service contact.

**Personality/Self-Esteem**

Girls and women with disabilities are as tremendously diverse as any other category of human being. Indeed, the development of personality characteristics and life histories will influence a woman's reaction to medical intervention, her adjustment to changes in functioning, and her coping mechanisms. In this section, we will begin by briefly reviewing personality factors—the relatively stable dispositional and relatively malleable situational variables that impact human functioning. For rehabilitation professionals, however, the research literature on experiences of marginalization, disempowerment, self-esteem, and implicit messages about what it means to be a woman with a disability are important to consider. After a general discussion of personality and self-esteem, we will summarize some research literature focused specifically on women with disabilities.

Personality psychology is a diverse area of study. McAdams and Pals (73) outline five big principles in personality theory and research: evolution and human nature, dispositional traits, characteristic adaptations, life narratives, and culture. Evolution has impacted human nature, resulting in fundamental ways in which human beings are the same. Specifically, there are some universal, cross-cultural, and historical aspects of human life which have evolved over time such as adaptation and cooperation. Sheldon (74) focused on three psychological needs: “need to sustain a basic sense of self (autonomy), to manipulate the environment in order to achieve instrumental goals (competence), and to form cooperative relationships with others (relatedness)” (73). For women with disabilities, the issues of autonomy (or self-determination), competence, and relatedness are as fundamental to their flourishing as they are for health care providers and family members who each bring their own perspectives to the medical encounter.

The second principle is the importance of dispositional traits. Dispositional traits include what has classically been referred to as the big five factors of personality: openness, conscientiousness, extraversion, agreeableness, and neuroticism (75,76). Debates within personality psychology have focused on dispositional traits. The concept of self-esteem has the qualities of a personality trait and a psychological state (78) and refers to an appraisal of one's worth. Other terms such as self-regard and self-respect have also been used in the literature. Recently, scholars have focused on specific aspects of self-esteem. For example, Crocker et al. have defined “contingencies of self-worth” or domains from which individuals derive their perception of self-worth. They describe the domains of virtue, God’s love, support of family, academic competence, physical attractiveness, gaining others'
approval, and outdoing others in competition (78). Therefore, self-esteem is contingent on the extent that an individual bases her self-worth on a specific domain, and either succeeds or fails in that domain.

The majority of the research on self-esteem has focused on the level of self-esteem, high or low, and the assumption has been that high self-esteem is associated with positive outcomes. Recently, researchers have suggested that the relationship may be weaker than previously assumed (79). Crocker and Park (80) further argue that “the pursuit of self-esteem interferes with relatedness, learning, autonomy, self-regulation, and mental and physical health. Pursuing self-esteem can be motivating, but other sources of motivation, such as goals that are good for the self and others, can provide the same motivation without the costs” (80).

When the HCP focuses on perceptions of self-esteem, he or she may be limiting the creative solutions to goal attainment. For example, if a woman reports low self-esteem and views this as a global construct about her worth, and the HCP is complicit in this belief, alternatives may not be suggested. Focusing on specific activities, on areas of strength, on relationships and interdependence, rather than individualistic perceptions of self-worth, may provide alternative ways of positive adaptation.

In a study focused specifically on the self-esteem of women with physical disabilities, Nosek et al. (81) assert that it is not the disability itself but the contextual, social, and emotional impact that affect the women’s sense of self. The authors state, “we learned that their self-evaluation was strongly influenced by feedback from parents and other family members, teachers and schoolmates, friends and intimate partners. The impact of this feedback about their value as women with disabilities was predominantly negative” (81). The authors also discussed the importance of activities and the various barriers, environmental and attitudinal, that negatively impacted the women. In their study comparing 475 community-dwelling women with disabilities to 406 women without disabilities, they found that women with disabilities had significantly lower self-efficacy, quality of intimate relationships, employment, and less education. They had higher social isolation and overprotection during childhood. The authors further state, “The social environment in which women with disabilities live is primarily a hostile one, transmitting stereotypes that have existed for millennia, resulting in stigmatization and exclusion” (81).

Some young women with disabilities, however, including those with extensive physical limitations, have demonstrated resilience in the face of invalidating messages. They have developed positive self-concepts as women and have established enduring intimate relationships despite all predictions (8). Many give credit to perceptive family members who affirmed their gender roles and who facilitated their social contacts. Others, surrounded by discouragement, refused to internalize it. Seeking strong disabled women role models for support or simply finding validation from within, they have integrated their disabilities into sound and positive identities as women (68,71). Accepting disability as a familiar and “normal” part of their lives, they seem to set their own standards for what is valuable, sexually appealing, and womanly.

Self-efficacy, or one’s beliefs about her capabilities to attain goals, is also related to the construct of self-esteem. In a qualitative study of women and men that examined the perceptions of self-efficacy and rehabilitation with patients with traumatic brain injury (TBI) and stroke (82), the authors identified 11 themes that patients described as helping them engage effectively in rehabilitation. In the first domain, self; the first five themes were self-reliance, determination, independence, recognizing own improvements, and pushing one’s limits. In the second domain, others, the next three themes were external reassurance, vicarious experience, and working with professionals. In the third domain, process, the final three themes were setting goals, information needs, and making time for rehabilitation. The authors suggest that examining self-efficacy is one way to “maximize the benefit gained from rehabilitation” (82). In the outpatient setting as well, focusing on task mastery and use of vicarious experiences (e.g., peer support) may be an effective means for increasing self-efficacy and psychosocial outcomes.

Self-determination is also influenced by self-esteem. Self-determination is distinct from independence, in that someone can be physically dependent on others but be self-determining in their choices and preferences. In the psychological literature, Self-Determination Theory (SDT) (83) refers to a theory of human motivation that encompasses the concepts of perceived competence, autonomy, and relatedness. More specifically, SDT differentiates between autonomous and controlled motivations. “Autonomous regulation involves experiencing a sense of choice, a sense of full volition,” whereas controlled regulation “involves people feeling pressured or controlled by an interpersonal or intrapsychic force” (84). Autonomous motivation leads to lasting changes and positive behaviors while controlled regulation results in temporary changes. For women with disabilities, as with any other patient, HCPs have an important role to play. For example, HCPs can provide relevant feedback to encourage mastery and work toward a strong sense of relatedness which involves trust and connection. This can lead to higher self-determination and willingness on the part of the patient to learn and adopt new strategies for health improvement (84). In general, when physicians support the concepts of self-determination in their practice, there are positive outcomes such as adherence to medications and improved physical and psychological health (85).

**REPRODUCTIVE AND HEALTH ISSUES FOR WOMEN OF CHILDBEARING AGE**

**Menstrual Irregularity and Fertility in Women with Physical Disabilities**

Women with disabilities need to be empowered to not only understand their bodies and how they work but also to direct their choices regarding reproductive health care. For some women, this may involve decisions regarding childbearing.
For most women with physical disabilities, fertility potential is preserved, and menses resemble the patterns of women without disabilities (86). In some cases, though, menstrual irregularity and fertility problems can occur. The most common hormonal imbalances found in women in general are disorders of prolactin secretion or thyroid function. Disorders of neuroendocrine function are not uncommon after TBI, though there is no clear data on the frequency with which it affects the menstrual cycle or fertility (87). In an analysis of 13 pooled studies with 809 people with TBI, the prevalence of hypopituitarism in patients at least 5 months post TBI was found to be 27.5%; with a greater prevalence found in populations with more severe injuries (88). Some groups now recommend routine screening for hypopituitarism of patients with moderate to severe TBI (88,89).

For women with spinal cord injury (SCI), almost all women will resume their normal cycle within the first 3 to 6 months post injury (90); about 25% of one population of women studied report increased autonomic symptoms (sweating, headaches, flushing, gooseflesh) around the menses (91). Occasionally, medications can also cause menstrual irregularities. Phenytoin and corticosteroids may affect thyroid function and ovulation; tricyclic antidepressants, antipsychotic medications, and some antihypertensive medications may also cause menstrual irregularities by affecting prolactin levels (86,92,93). Treatment of menstrual irregularities varies with age and medical condition. Pregnancy should always be considered in evaluating menstrual irregularity. Thyroid function tests and prolactin levels can be helpful in evaluating women with irregular menstrual periods (94,95). Correcting hormonal causes of menstrual irregularities can often regulate the menstrual cycle successfully and improve fertility. In short, for women who desire to conceive and are having difficulties, they need to undergo an evaluation identical to those without disabilities (96). The male partner needs to be involved in the process by providing a semen specimen for analysis, as about 40% of all couples who present with infertility problems will have a male factor contribution (96).

Abnormal bleeding can also occur with structural problems such as uterine or endocervical polyps, fibroids, cervical pathology, and vulvovaginal lesions (97). Menstrual irregularities become more common as women approach menopause. Careful gynecologic evaluation is required to determine an appropriate course of treatment. Although gynecologists frequently prescribe low dose combination oral contraceptives and other hormonal agents either to regulate menstrual flow or control vasomotor symptoms, there is no data derived from studies of women with disabilities to guide practitioners (62–67).

Contraception
The history of contraception for women with disabilities has at times been coercive and even frankly oppressive. Dating back to the early 1900s, fears of women with disabilities producing children with disabilities led to some social policies encouraging sterilization and criminalization of marriage (98,99). Other fears have checkered the reproductive history of disabled women, including the perceived inability of women with disabilities to be “good” mothers and exaggeration of health risks for women with disabilities who choose to bear children (100–102). Concerns about “physician-controlled” contraceptives, such as Depo-Provera (Upjohn, Kalamazoo, Michigan), still percolate in segments of the disability community as a result of this history and the fear that a woman’s reproductive choices might be curtailed by physicians, parents, or guardians who make decisions on her behalf (103,104). Sensitivity to this history is essential in establishing patient-centered, trusting relationships between health care providers and women with disabilities. Women with disabilities have the same right to knowledgeable information about contraception as nondisabled women and should participate in their reproductive health care decision to the fullest extent possible. Contraceptive choices may be colored, though, by the nature of the woman’s physical impairment or chronic disease condition (105). For example, a woman with impaired hand function may have difficulty using barrier methods such as the diaphragm, sponge, or spermicidal product without assistance. Problem solving with an occupational therapist, or working with the woman’s partner or personal assistant, may result in an adequate solution for the use of barrier methods. Condoms, of course, always remain an effective method if the male partner is in agreement and provide the dual benefits of contraception and protection from sexually transmitted diseases (106).

There are a rapidly growing array of hormonal options for contraception that include a variety of combination pills, progestin-only mini-pills (62), injections (e.g., Depo-Provera—depo medroxyprogesterone acetate), subdermal implants such as Implanon (Schering Plough, Kenilworth, New Jersey), transvaginal rings, such as Nuvaring (Schering Plough, Kenilworth, New Jersey), the levonorgestrel intrauterine system (Mirena; Berlex Laboratories, Inc., USA), and transdermal patches such as OrthoEvra (Ortho-McNeil Pharmaceutical, USA). For women who desire hormonal methods of contraception, information about the risks and benefits of various hormonal contraceptive options in women with disabilities is limited. For example, some women with mobility limitations (such as women with spina bifida, SCI, and multiple sclerosis [MS]) might have an increased risk for developing thrombotic events when using combination oral contraceptives, but definitive statistics are not available (105). Woman with a history of thrombotic or thromboembolic stroke should probably avoid estrogen containing contraceptives (105). There is some indication that there may be a protective effect between oral contraceptives and rheumatoid arthritis, but the results thus far are inconclusive (105). The progestin-only contraceptive pills are associated with an increased risk of abnormal uterine bleeding and are less effective than the combination pills (62). Depo-Provera, an injectable form of progesterone that is effective for at least 12 weeks, is highly effective, and patients are often pleased with the decreased menstrual flow or amenorrhea that results from this method (107,108). Many women experience some weight gain and have reduced estrogen levels, which can lead to osteopenia (109). Indeed, Cromer et al. (110)
documented a 3% loss in bone mineral density among adolescent girls who should typically have experienced a 9% increase in bone mineral density. Concerns regarding how prolonged use of Depo-Provera might impact peak bone mineral density and subsequent risk for osteoporosis later in life prompted Food and Drug Administration (FDA) to issue a “black box warning” in 2004 (111). For women with disabilities who often confront an elevated risk of both osteopenia and osteoporosis due to chronic immobilization, this consideration is particularly important and providers should use caution when prescribing this for more prolonged periods of time.

Implanon (Schering Plough, Kenilworth, New Jersey), an etonogestrel containing subdermal implant, has been available worldwide since 1998 and in the United States since 2006. Unlike Norplant, a six-rod levonogestrel implant previously available in the United States, Implanon consists of only one 4 cm, plastic rod implanted at the biceps-triceps junction of the nondominant arm. Placement consists of a minor procedure that resembles receiving an injection. Removal is far less problematic than encountered with Norplant. Approximately 10% of women discontinue Implanon due to irregular bleeding and many other patients note occasional breakthrough bleeding (112). This might present a hygiene issue for women with chronic immobilization impairments. Limited data are available to assess pharmacokinetic changes among Implanon users taking other medications that induce hepatic microsomal enzymes. Although pregnancy associated with Implanon is rare, several pregnancies reported with its use occurred in women using carbamazepine or other enzyme-inducing agents (113). Fortunately, the implant exerts little impact on bone mineral density and contains no estrogen.

OrthoEvra is a transdermal patch applied weekly on the buttocks, abdomen, upper torso, or outer arm. Approved in the United States in 2002, OrthoEvra offered a novel method of delivering hormones, contained in combination oral contraceptives, for patients unable to swallow a pill or who could not remember to take a daily pill. FDA issued “updated labeling” in 2005 following reports of venous thromboembolism (VTE) associated with its use (114). Risk of VTE typically correlates with estrogen dose, and 3 weeks’ of patch use exposes patients to roughly 71% more estrogen (“area under the curve”) than ingestion of a low-dose oral contraceptive (115). Although the absolute risk of VTE remains small among users, patients at elevated risk of VTE, including chronically immobilized patients, should likely forego its use until more data becomes available.

Nuvaring is a flexible, transparent ring placed in the vagina 3 out of 4 weeks/month. Indications and contraindications to the ring resemble traditional oral contraception. Unlike the OrthoEvra Patch, Nuvaring exposes women to less estrogenic exposure than either the patch or pills while providing superior cycle control (115). Women with disabilities might find it difficult to check for the ring’s presence in the vagina, particularly after intercourse when the vagina physiologically dilates and may need the assistance of their partners to do so. Women with denervation atrophy and relaxation of the pelvic floor might also have difficulty retaining the ring and approximately 5% of women discontinue the ring due to “ring-related” issues such as vaginal discharge.

Intrauterine devices (IUD) represent the most popular form of reversible contraception throughout much of the world but remain underutilized within the United States. Despite contraceptive efficacy that approaches that of surgical sterilization, many clinicians still fear infectious and other risks previously attributed primarily to shield type IUDs (116). Recently, clinicians in the United States have begun to rethink an unjustified bias against intrauterine contraception, liberalizing prescribing practices to include groups such as nulliparous patients and adolescents (117). No literature specifically evaluates the risk of IUD use among women with disabilities, though physicians should clearly exercise caution in some circumstances. Women with impaired pelvic sensation, for example, might be less able to detect symptoms consistent with pelvic infection as well as spontaneous expulsion of the IUD. Other women will have difficulty palpating vaginally for the IUD string, one sign that the IUD remains properly placed, and again may require the assistance of their partners or personal assistants to do so. Although the FDA has not approved IUDs for noncontraceptive indications, progestin-bearing IUDs may offer the opportunity to control menometrorrhagia without resorting to more invasive, surgical methods (116). This might offer significant hygienic advantage to women with mobility impairments who experience heavy or irregular menses—regardless of their need for contraception.

Pregnancy and Prenatal Preparation

When possible, as with any nondisabled woman, it is best to discuss issues and concerns about pregnancy and motherhood before becoming pregnant (6,118,119). Preconception counseling allows the woman to carefully consider any disability-specific factors which may raise medical concerns or issues for the woman, her partner, and her health care team to consider. Counseling about health behaviors such as diet, potential adverse effects of smoking, alcohol and illicit drug use on the fetus, and the benefits of various vitamin supplementations (such as folic acid), all may maximize the woman’s ability (disabled or not) to make choices and be an active participant in managing her pregnancy (120).

As many women with disabilities are on multiple medications, such as antiepileptic and antispasticity agents, careful consideration of the potential effects of the medications on a developing fetus is warranted (121–125). Although many physicians continue to rely on FDA’s Use-in-Pregnancy Ratings when prescribing medications during pregnancy, organizations such as the Teratology Society (Teratology.org) have recommended abandoning use of this system, since it fails to realistically balance the likelihood of fetal risk against maternal benefit (126). The benefits, for example, that many disabled patients derive from antiepileptic, antispasmodic, and other medications might justify some fetal risk in order to assure that the patient avoids seizures, worsening immobilization or other maternal compromise that unduly jeopardizes the health of both mother
and baby. Obstetricians enrolling disabled prenatal patients can now utilize numerous web-based resources, particularly regional and state teratogen information services, to obtain current data regarding fetal risk. As the intrathecal baclofen pump becomes an increasingly common tool for managing severe spasticity, it is inevitable that more disabled women who become pregnant are likely to have a pump. Though there is limited experience with the use of the pump in pregnancy, labor, and delivery, the information thus far has been reassuring for both mother and fetus/baby (127–130).

For some women, information about their health risks with pregnancy may be useful in facilitating a decision to pursue biological motherhood. A few disabilities may progress with pregnancy or in the postpartum phase, perhaps irreversibly, or require higher-risk interventions and support (131–139). For example, a woman with multiple sclerosis may expect that her MS could worsen postpartum though pregnancy itself does not seem to have a significant long-term impact on disease course (140–144). Similarly, a woman with spinal muscular atrophy may need ventilatory assistance as the gravid uterus pushes up on her diaphragm and impedes ventilatory excursion (134,138,145–147). Women with spina bifida, particularly those with shunts for hydrocephalus, urinary diversions, kyphoscoliosis, contractures, and pelvic obliquity may be at increased risk for pyelonephritis, shunt malfunctions, and caesarian sections (139). Pregnancy for patients with systemic lupus erythematosus (SLE) may pose health risks for both the mother and the fetus, though the prognosis has improved with better understanding of the pathogenesis of the disease and improved treatment options (148).

Women with physical disabilities may experience a temporary decrement in their functional abilities as pregnancy progresses, necessitating more physical assistance or different mobility devices (91,149,150). Most women with mobility impairments will be affected to some degree by changes in their center of gravity with increasing weight gain. Women who had previously managed their mobility with braces or assistive devices may find their stamina and balance affected as the pregnancy progresses and elect to temporarily use a manual wheelchair (150). Women who had previously been independent with manual wheelchairs may notice that they fatigue much more quickly and are unable to navigate the environment independently; these women may choose to use a power wheelchair in the latter stages of pregnancy. Some of these women will also appreciate the option of using a power mobility device postpartum to transport their child. Anticipating these changes will allow the woman and her health care provider to determine needs that might arise during the pregnancy and plan proactive interventions (such as referrals to physical therapy in the second trimester) that can be instituted to lessen the impact of some of these changes (150).

Information about other potential risks associated with pregnancy may assist in developing a plan of care to obviate or lessen these risks (106,121,134,139). For example, for women who use indwelling catheters and are chronically colonized with bacteria, the risk of bladder infections and pyelonephritis may be heightened (91,151–153). Prompt recognition and prevention of urinary tract infections is extremely important in preventing preterm labor (152). Women who have been accustomed to intermittent catheterization may find their bladder capacitance decrease, with increased leakage or frank incontinence, particularly during the latter stages of pregnancy. Increasingly, urinary diversion is being offered as a means to manage the neurogenic bladder for people with physical disabilities (154–157). This population poses particular considerations in the managements of pregnancy, labor, and delivery, especially if a caesarian section is required.

Another potential complication of pregnancy for a woman with mobility impairment is skin breakdown. Excellent skin care and vigilance are extremely important in preventing pressure ulcers during pregnancy for any woman with areas of insensitive skin (121,150). A pressure-relief program that had previously been adequate in preventing pressure sores may no longer be adequate with the increasing weight gain and hormonal changes of pregnancy. Counseling about more frequent pressure reliefs, evaluating the adequacy of the wheelchair cushion, and increasing the frequency of skin checks are good prevention strategies. Counseling with the labor and delivery staff about padding and positioning is also crucial in preventing this complication.

Constipation and hemorrhoids are potential problems in any pregnancy with changes in diet, hormonal influences on the intestinal tract, iron supplementation, and the gravid uterus pushing on pelvic veins, but these potential complications can be particularly difficult for a woman with a neurogenic bowel dysfunction who must use a bowel program (158). Vigilance, modifications in diet and the bowel program, and proactive bowel hygiene can go a long way in preventing potential bowel impaction and hemorrhoidal complications.

Pregnancy is a time of hypercoaguability for all women, and may pose increased risk for deep venous thrombosis for women with mobility impairments (106,159). Frequent examinations and counseling about signs and symptoms of deep vein thrombosis (DVT) are extremely important in helping women and care providers quickly recognize and seek assistance in managing this potentially life-threatening complication. Counseling regarding periodic elevation of the legs, watching the salt intake, and the use of compression stockings may help minimize dependent edema.

For women with respiratory impairments from SCI, neuromuscular disorders such as spinal muscular atrophy, or extensive scoliosis with restrictive lung disease, respiratory function may also be affected with lowering of the functional residual capacity as the uterus grows (106,121,160). Pregnant women frequently hyperventilate and develop a mild degree of respiratory alkalosis, believed to be an effect of progesterone (160). Women with limited respiratory reserve may be particularly affected by these changes and require some modification in their activities, sleeping positions, and occasionally, ventilatory assistance to manage the later stages of pregnancy, labor, and delivery (138). There have even been reported cases of women who had no autonomous ability to breathe on their
own and successfully used noninvasive ventilation throughout pregnancy (161).

A few potential complications of pregnancy can be life-threatening for a woman with disability (162–165). One of the best recognized and frequently written-about complications is AH in a woman with tetraplegia or high paraplegia (usually T6 or above) (121,163–167). Not only can AH signal an issue requiring immediate attention such as bladder distention or infection, bowel impaction, or labor, but also the signs and symptoms can be misinterpreted as preeclampsia and lead health care providers down an incorrect path of diagnosis and treatment. Untreated AH can result in intracerebral hemorrhage and even death. For women who are at risk for AH, consultation with an anesthesiologist well before delivery may be prudent (162–165,168). A team approach to pregnancy, with a specialist in SCI working closely with the obstetric staff, can facilitate prevention and management of these potential complications (35,106,121). Women with disabilities other than SCI also benefit from consultation with an anesthesiologist prior to labor and delivery, including women with lung and cardiac disease, skeletal abnormalities and scoliosis, and MS (169).

Pregnancy offers an excellent opportunity for the woman with disability to begin to prepare for her future role as a mother (150,170,171). If a woman has an extensive mobility impairment, referrals to an experienced occupational therapist and physical therapist may help her to develop a variety of anticipated skills to maximize her ability to parent (150, 172–175). For example, how will she carry the baby? How will she feed the baby? If she plans to breast-feed, what positioning techniques will work best for her? What about diapering, dressing, and bathing? A home visit may be helpful in setting up the environment with adaptive equipment to facilitate her functional abilities. If the woman will require or chooses to use assistance in managing some of the babycare tasks required, working out a plan prior to delivery is important.

Many women with disabilities also yearn to talk to other women with similar disabilities about their experiences as mothers. Peer support and referrals can be extremely helpful in facilitating successful adjustment for both mother and baby and in anticipating the needs of each subsequent developmental stage (150,176). If a woman uses a wheelchair, how will she manage her toddler and maintain safety and discipline? How will she work with her child around the child’s growing perception of her disability (177–180)? Women also report that continued contact with child care specialists from occupational therapy and psychology can be extremely important in adapting to each new developmental stage. More detailed information and resources on mothering with a disability are presented later in the chapter.

**Labor and Delivery**

Preparation for labor and delivery is an important consideration for the woman with a disability, though, depending on her disability, the traditional childbirth classes may not be adequate for her and her partner (150). Information to assist the woman in proper recognition of labor involves understanding expected signs and symptoms in the context of her particular disability (121,181). For example, a woman with extensive sensory impairments may not be able to sense pain from uterine contractions but have to rely on manual or electronic detection of the contractions. If she has high paraplegia or tetraplegia, periodic headaches from recurring AH (occurring with each contraction of the uterus) may be the first symptom she recognizes. It is important that the pregnant woman be instructed on the early signs of labor, including recognition of amniotic fluid leakage, and know when to seek immediate care. Of course, frequent regular checkups with her obstetrician will also be helpful in predicting impending labor.

As previously mentioned, a discussion of anesthesia before labor and delivery is important in establishing a care plan (162,163). For example, consideration of an epidural catheter may be important for a woman at risk for AH (e.g., not to manage pain but to block the afferent signal from the uterus to the spinal cord triggering the sympathetic response (121,162,165)). For a woman with respiratory compromise, a plan for providing ventilatory support should be discussed if this becomes necessary. If general anesthesia is needed for any reason, pertinent knowledge about disability-related issues that could influence the safety of general anesthesia should be considered. For example, a woman with a SCI should not receive succinylcholine as a depolarizing agent secondary to the risk of hyperkalemia (162,182). In addition, if a woman has had a cervical fusion of her neck, oral intubation may be difficult.

Working with the labor and delivery staff is also important in anticipating the needs of a woman with disability. If she is at risk for skin breakdown, proper padding and frequent position changes are important. If she has significant contractures and spasticity, tips for managing these issues with range of motion, positioning, and occasionally, medication may be advisable (150). An increasing number of women are being recognized with latex allergies, in particular a large proportion of women with spina bifida (183), and this should always be mentioned to the staff as well. It is optimal if the woman (and her partner) has an opportunity to tour the facility where she will deliver, to have an opportunity to ask questions about accessibility (e.g., bathrooms, showers, etc.), and to discuss her wishes for her peripartum care with her obstetrician and the nursing staff.

**MOTHERING WITH A DISABILITY**

A growing number of women with disabilities are choosing to become mothers, defying stereotypes of asexuality and material incompetency historically ascribed to them by a society that stigmatizes disability (Fig. 58-3). They and mothers who later acquired disabilities are seeking recognition and support for their capabilities. This surge was generated in the 1970s by the independent living movement and in the 1990s by the passage of the ADA and is forcing health care providers to take a closer look at ways to assist mothers with disabilities to
achieve self-determination over their parenting activities. Best estimates suggest that of the 6.9 million disabled parents in the United States, approximately 60% are women, and about a third of these women, compared to 25% of disabled men, have children in the home (27).

Although mothers with disabilities share the same basic needs and concerns as nondisabled mothers, their capabilities to function aptly in a parenting role are greatly enhanced with the use of adaptive equipment and techniques. Also vital is having regular access to peer support, and knowledgeable health care providers who understand the concept of self-determination, as well as the sociopolitical barriers confronting women with disabilities (184,185).

Mothers with disabilities have more possibilities for support and resources available to them than ever before, for example, through the Internet, specialized resource centers, and rehabilitation specialists. Yet, many women with disabilities live in fear of their children being taken away from them if they do not live up to other people’s expectations, and thus feel they have to go to great lengths to present themselves and their children as managing “normally” in order to be accepted as “ordinary” mothers (186). This is a realistic fear, as historically the lay public and HCPs have questioned the capacity of women with disabilities to be care providers, based upon the perception and presumption that those who receive care (e.g., PAS) could not then provide care to others (186). Findings from a national study of parents with disabilities confirm the prevalence of discrimination that mothers with disabilities face regarding their rights to bear and raise children (27,187–190). Of the more than 1,000 parents with physical, sensory, and/or cognitive disabilities surveyed, 15% reported that others attempted to remove their children (27).

Research

While research on disabled parents is relatively scarce, some significant strides have been made in the past 25 years, spurred by the dedicated work of a national center. In 1982, Through the Looking Glass (TLG) was founded in Berkeley, California, as a disability community-based organization, a product of the independent living movement. It is dedicated “to create, demonstrate, and encourage nonpathological and empowering resources and model early intervention services for families with disability issues in parent or child which integrate expertise derived from personal disability experience and disability culture” (191). In 1998, TLG became the first National Center on Parents with Disabilities, funded by the National Institute of Disability and Rehabilitation Research, U.S. Department of Education. In addition to performing research, TLG provides information, referral, publications, training, and consultations regarding parenting with a disability (including counsel on custody, adoption, and adaptive babycare equipment). The research is conducted from a disability cultural perspective, an approach grounded in the view that disability does not reside solely within the mother, but rather is socially constructed in her experiences of discrimination and marginalization.

TLG remains on the forefront of conducting research that documents the capabilities of mothers with disabilities in an effort to combat pervasive stereotypes that emphasize the negative effects of a mother’s disability on the well-being of her children. Its early research was driven by data showing that parents with physical, visual, and hearing disabilities “were offended that professionals had questioned their abilities to care for their babies, implying that the parents had never considered the babycare implications of their own disabilities” (188). From 1985 to 1988, TLG explored basic infant care tasks including feeding, lifting, bathing, diapering, and dressing by videotaping the interactions between ten mothers and their infants (192). They were able to capture the gradual process for mutual adaptation that emerged between mother and baby during their interaction (192). For example, the study documented the process of reciprocal cooperation between a mother with paraplegia and her 1-month-old infant who would help his mother lift him by balling up his legs and head on his mother’s lap and remaining still until his mother completed the lift. The videotape documented the mother’s cuing process that facilitated the infant’s adaptation through a series of strategies that included positioning the baby on his back and a variety of tugs on his clothing that signaled her intention to lift him. Also, mothers with hemiplegia report that their children knew at an early age to always approach their mothers on their stronger side and to refrain from extending arms to be picked up when their mothers were standing or not in a more stable, seated position (192). This study offered practitioners an expanded working sense of range of “good enough
parenting” to include issues of physical difference in mothers (188). The findings demonstrated that mothers with physical disabilities were naturally resourceful in creating their own adaptations to baby care and that adaptation in infants was often easily facilitated by their disabled mothers.

TLG’s videotape evaluation and analysis methods can be an invaluable tool for clinicians and health providers called upon to assist parents with disabilities involved with child protective services systems (188). In one renown incident that involved Tiffany Callo, a mother with cerebral palsy who lost her two sons to foster care, the TLG video based analysis was able to demonstrate that Callo’s diapering skills were within the range of “good enough” (188). Furthermore, when Callo’s lack of a mutual gaze with her infant was interpreted by social service professionals to be pathological, TLG researchers directly challenged this assessment with the introduction of a lap tray and well-positioned pillows that allowed Callo to establish a mutual gaze with her infant, which she was previously unable to achieve when left solely to her own abilities. TLG determined that Ms. Callo’s experience and that of many others underscored the necessity for ensuring the provision of adaptive baby equipment when attempting an accurate assessment of a disabled woman’s full range of mothering abilities. This is especially important in situations that involve a disabled woman’s legal rights as a parent or in custody decisions.

TLG continues to spearhead research that helps improve the quality of family life for individuals parenting with a disability. Three important initiatives conducted by TLG that are of particular importance to rehabilitation clinicians include (a) the development and evaluation of babycare equipment, (b) a national survey of parents with disabilities, and (c) a study examining the effects of mothers with disabilities on their children.

Research Initiative on Babycare Equipment

TLG has undertaken some pioneering work in babycare equipment including the design and evaluation of lifting harnesses for use by mothers with physical disabilities (193–197). By lessening the disabled mother’s level of physical stress and fatigue, these studies demonstrated that the provision of adaptive equipment and adaptive techniques can help prevent secondary disabilities such as back or repetitive stress injuries. In addition, the equipment was found to help with “alleviating depression associated with postnatal exacerbations of disability such as in MS” (188).

Although accessible babycare equipment is more commercially available now than ever before, it can be difficult to find accessible cribs and playpens. Some equipment needs to be customized to be truly accessible, which can pose liability issues. It can also be difficult to find skilled and affordable rehabilitation engineers to make the necessary adaptations. Sometimes, creativity and nontraditional equipment can be used (198). Irrespective of whether a mother is seeking to locate commercial baby equipment or equipment customized by rehabilitation engineers or fabricating the adaptations herself with tips from parenting groups on the Internet, it is essential that all equipment be safe for use by both the mother and child, which can be determined by the evaluation of a skilled occupational therapist.

Research Initiative on National Parenting Survey

In the largest national study of its kind (27), 1,175 parents with disabilities were surveyed on their needs and experiences that included the barriers they encounter in their role as parents. Noteworthy are findings that reveal that parents with disabilities face significant social barriers that impede their parenting activities such as the lack of accessible transportation (87%) and lack of accessible housing (69%). Fifty percent of respondents also reported they did not know how to locate adaptive parenting equipment, but a predominance of respondents felt their independence, efficiency, stress, and fatigue would improve if it were available. Forty-eight percent reported difficulties in affording adaptive equipment, even if it were available.

Research Initiative on the Impact of a Mother’s Disability on her Family

Research on the predictors of problem parenting were the same as those for nondisabled mothers (e.g., history of physical, sexual, or substance abuse in the mother’s family of origin) (188) and challenge the common stereotype that children of mothers with disabilities will be “parentified” or obliged to care for the emotional or physical needs of their disabled mothers at too early of an age (188). In fact, contrary to being parentified, this study found that children of disabled mothers were often spared the responsibilities of common family chores, such as taking out the garbage, when the mothers felt that the request was in any manner “necessitated” by their disabilities (188). An additional study by other researchers who looked at mothers with SCI found that there were no significant differences between children raised by mothers with SCI and children raised by mothers without disabilities in regards to individual adjustment, attitudes toward their parents, self-esteem, gender roles, and family functioning (180). It is commonly reported that among the multitude of positives that disabled parents offer their children are the ability to learn compassion and openness toward others.

Facilitating Self-Determination

Although the introduction of adaptive equipment and parenting resources early in the mothering experience can minimize problems and significantly improve a disabled mother’s capacity to provide “totally independent care” for her baby, this is not feasible for all mothers. Financial constraints, difficulty finding people who can fabricate equipment, scarcity of customized equipment, and physical limitations may still prevent many women from becoming fully self-sufficient in managing the care of a child. In fact, over half (57%) of disabled parents report that they used personal assistance for help with parenting tasks (27). It is thus essential for mothers and health care providers alike to recognize that a disabled woman’s lack of physical independence does not necessarily diminish her mothering skill. Rather, it shifts the focus to the value of
facilitating self-determination instead of self-sufficiency in the parenting role. This distinction is a crucial one.

A caregiver can be a family member, spouse, friend, or hired personal assistant. Caregivers are most beneficial when they enhance the special bond between a parent and child (199). Health professionals and occupational therapists can help maximize the value of hired caregivers by facilitating childcare training to teach the caregiver how to assist the disabled mother without taking over the role of the parent (175). It is also important to help the mother find her role during the activity. For example, as the caregiver diaper the baby, the mother can engage the baby in play activities.

Recently, TLG devised an assessment tool, Baby Care Assessment for Parents with Physical Limitations or Disabilities, administered by an occupational therapist that is used to evaluate a disabled mother’s full range of parenting capabilities (200) (Fig. 58-4). The tool emerged from the disability community and incorporates perspectives of parents in this community. For instance, independent performance is only valued as much as the parent needs or desires it. The tool brings together the parent’s experience, including a self-assessment, and the evaluation of such factors as cognition, ergonomics, environmental barriers, and mental health issues, and then analyzes how those aspects effect the child care activity. It includes tips and questions that guide assessment and a screening of parent/child relationship issues as well as appropriate referrals. It has also been used successfully, as one aspect of evaluation, in mandatory assessment situations involving family court or child protection systems. An alternative instrument, though less comprehensive than the TLG tool, can also be used as a screening tool for mothering tasks (Fig. 58-5). On the basis of the assessment, the occupational therapist, mother and spouse, or caregiver collaborate to determine tasks the disabled mother needs simplified, is comfortable performing independently, or will delegate to helpers.

The lack of disability-aware social institutions and services can pose real threats to the self-determination of mothers with disabilities. Not uncommonly, states exclude child care from the list of activities of daily living for which a disabled person can get personal assistant support. Unfortunately, women with disabilities who have a low income and are unable to afford private child care have lost custody of their children because the state refused them assistant services and then deemed them incompetent in parental functioning (100). Full knowledge of relevant state policies on the provisions of personal assistants will create realistic expectations on the part of prospective

---

CHAPTER 58 | EMPOWERING WOMEN WITH DISABILITIES TO BE SELF-DETERMINING IN THEIR HEALTH CARE

Item Name: Care of Others (Child Care)

Definitions:

a. Physical daily care—Selects proper equipment, uses correct positioning, and performs the following:
   - Feeding—Breastfeeding, bottle feeding (includes set-up and clean-up), spoon feeding (includes opening jars)
   - Changing a diaper—Donning and removal (specify type: cloth with pins or diaper cover or plastic with tape) and managing hygiene
   - Dressing/Undressing—Donning and removal (specify type: one-piece jumpsuit, t-shirt, plastic pants, dress with snaps or buttons, socks, shoes, booties)
   - Bathing—Holding, washing, drying off (indicate location: tub, sink, positioning device)
   - Preparing for bed—Changing clothes, lifting and lowering side of crib, moving child from crib to lap or wheelchair, covering child with blanket
   - Carrying/lifting—Lifting to and from floor and transporting 50 feet
b. Playing/maintaining—Selects appropriate toys/activities, uses age-appropriate interaction, and is aware of various intervention strategies for behavior management and facilitation emotional growth
c. Illness/first aid—Takes temperature, uses nasal aspirator, manages suppositories, and gives medication. Is aware of symptoms and is able to perform minor emergency procedures (cuts choking, CPR, poisoning), “baby-proofing” of home, and seeking assist in an emergency
d. Community access/resourcing—is able to select and access day care/preschool, child supplies, physician, babysitter. Is able to instruct babysitter. Is able to manage stroller and car seat (includes placement and removal from car, placing child in/out of devices), diaper bag, back pack.

Note: Select and note activities appropriate to the patient’s lifestyle/abilities.

SCALE POINTS

NO CAREGIVER REQUIREMENT:
1. Complete independence: Patient can perform task safely and consistently within a functional amount of time in any environment without specialized equipment, modifications, or verbal/physical assist.
2. Modified independence: Patient can perform task safely and consistently without caregiver assistance but requires specialized equipment, an adapted environment, and/or excessive time to complete task.
REQUIRES CAREGIVER ASSISTANCE
5. Set-up/in frequent assistance: Patient consistently requires caregiver presence to initiate and/or complete task (e.g., with set-up and/or clean-up), but patient safely and consistently completes interim components independently, or patient requires caregiver’s assistance for unpredictable occurrences only.
4. Minimal assistance: Patient performs more than three quarters of the task safely and consistently. Because of physical or cognitive impairment, patient requires caregiver to provide physical or verbal assist to complete one quarter of the task.
3. Moderate assistance: Patient performs half to three quarters of the task safely and consistently. Because of physical or cognitive impairment, patient requires caregiver to provide physical or verbal assist to complete on quarter to half the task.
2. Maximal assistance: Patient performs one quarter to half of the task safely and consistently, or may be physically unable to perform any part of the activity but can accurately direct the caregiver. Patient requires caregiver to provide physical or verbal assist to complete the remainder of the task.
1. Dependent: Patient performs less than one quarter of the task consistently. Patient is unable to direct the caregiver and is dependent of the caregiver’s physical or verbal assistance for completing more than three quarters of the task.

FIGURE 58-5. RIC functional assessment scale. The occupational therapy department would like to recognize the contribution made by Sharon Gartland, OTR/L, and Sheila Gupta, OTR/L, in the development of the definition for child care. (Used by permission of the Rehabilitation Institute of Chicago.)

parents and better aid practitioners in educating their clients and peers about ways to reduce the risk for unnecessary removal of children from their disabled parents.

Child Protective Services: Empowering Mothers With Disabilities and Their Advocates in Court

Despite the increase in parenting resources and the fact that there is no scientific evidence for a correlation between a parent having a disability and inadequate parenting, mothers with disabilities can encounter discrimination in the legal process and find themselves in circumstances where their right to parent is unfairly challenged (201). Such legal challenges require that disabled mothers defend their parental right to the companionship, control, and custody of their child (201). When the state threatens to remove a child from his or her mother, the situation can be frightening and complex for the entire family. It can involve documentation and testimony from rehabilitation clinicians who need to be prepared to challenge the misconception that women with disabilities, by virtue of their disability, should not have children and are unable to care for them. When preparing to advocate on behalf of a mother’s parental right, it is critical that the advocate understands the rules specific to their state or county that determine the state’s definition of “unfit” (201) (Fig. 58-6).
All advocates or mothers with disabilities who have questions or concerns about parental rights for mothers with disabilities are encouraged to contact the National Resource Center for Parents with Disabilities to help determine state-specific rules and the procedures involved regarding the procurement of an attorney or preparations for a Dependency Court experience (1-800-644-2666). In some states, once clear and convincing evidence has proven a parent unfit, the state must yet also prove to the court that it is in the best interest of the child for the parent's rights to be terminated. The age and sex of the child, mental and physical health of the parents, emotional ties between the parents and the child, as well as the parent's ability to provide the child with food, shelter, clothing, and medical care are among the many factors that may be considered by some courts in determining the best interest of the child (201).

Certainly, a goal of all rehabilitation clinicians working with prospective or current mothers with disabilities should be to help assess the mother’s ability to adequately parent with the supports of appropriate baby equipment, parenting techniques, and social support as mentioned above, but also to offer guidance on local resources to help the parents negotiate the challenges of creating a child-centered home that improves their child’s well-being, and is appropriate for the developmental stages. Disability-savvy social workers, psychologists, and pediatricians, in addition to physiatrists, physical therapists and occupational therapists, and peer support are all potential resources that can assist a new mother with disability to develop her parenting capacities. When possible, the earlier the process of parenting evaluation and support can be initiated, the better. A skilled occupational therapist can help the mother identify commercially available, safe adaptive babycare equipment, learn childcare techniques, and create a plan for how to make the home safely accessible and childproof (193,197). It is important that women develop a plan for how any childcare needs will be met, if an assistant will be required, and to have the opportunity to network with other mothers and learn about social and community resources. Continued contact is important as the child grows and develops to anticipate the future needs of the child (150).

**Adoption**

Stereotypes depicting women with disabilities as having to be “cared for” rather than as caregivers present ongoing challenges to those women wishing to adopt children (202). Given the lack of a universal standard for good parenting, agency staff may impose arbitrary standards onto disabled individuals applying for adoptive services (203). Despite these disincentives, many disabled men and women are taking advantage of the advocacy services of disability service groups and the supports of knowledgeable physicians and health providers to successfully negotiate the social and attitudinal obstacles to become adoptive parents (204). When screening for the parenting capabilities of a prospective parent, it is not uncommon for an adoption agency to require a physician’s certification documenting the individual’s good health and ability to raise a child (27). Before rendering a determination, rehabilitation and medical practitioners should carefully consider the prospective parent’s support network and full spectrum of parenting abilities. To focus on the disability as a negative to the exclusion of a woman’s full range of beneficial attributes may result in the denial of a child’s placement with a caring and competent parent. Disabled women report that an unfavorable medical determination adds to the frustrations of an already stressful adoption process and can reinforce a host of self-doubts and low self-esteem (205). In one case a physician, familiar to a disabled couple, surmised that their cerebral palsy represented a potential hazard to raising a child. The mother later on documented the couple’s challenges and successes in becoming adoptive parents by engaging in a national writing and speaking campaign to enlighten disabled women and their service providers on strategies for maximizing adoption options (202). Among the suggestions that she offers medical professionals who can often make or break a couples’ chances for adoption (205) are the following:

- Read the literature on adoption for people with disabilities available via the TLG Web site www.lookingglass.org.
- Remember that the same screening standards should be applied in adoption screening of all potential parents, whether they have a disability or not.
- Learn about the disabled parent’s support system, what resources are available to them, that is, family, friends, neighbors, and childcare assistance?
- Examine your own concerns and be honest about your discomfort. Consider whether you are the best person to handle this issue.

When a clinician is asked to be involved in evaluating parental rights in the legal process, it is important that the clinician attempts to obtain an evaluation of the prospective mother’s parenting capacities and need for adaptations and supports (by an occupational therapist using an assessment tool such as those referenced above). If a judge refuses to finalize an adoption or a caseworker declines a consent to the adoption and the prospective adoptive parent believes these refusals are a result of discrimination, she may consider filing an appeal of the decisions or grievances with the adoption agency, state bar
association or in the case of judge based discrimination, with the Department of Justice (203).

Peer Support for Parenting
Facing the onslaught of negative messages concerning one’s reproductive and parenting capacities can reinforce self-doubts in even the most confident of disabled women. Advocates and experts around the world are finding that disability consciousness and culture emerging from the power of peer-to-peer contact are the tried-and-true remedy for counteracting the isolation and devaluation women experience when parenting with a disability (184,185). From each other, many women are hearing for the first time that their decision to become mothers is a valid one, a cause for celebration. They rely on one another for a wholeness and strength that come not from independence and self-sufficiency but from interdependence and self-determination. As disabled mothers increase in number, so are the national and international parenting networks for people with disabilities, many of them online. Most independent living centers can assist health providers in identifying the peer support groups and networks for disabled mothers and women in their area (http://www.ilusa.com/links/ilcenters.htm).

MENOPAUSE AND AGING WITH A DISABILITY

More than half of women with disabilities in this country are greater than the age of 50 and pose unique issues in preventive health care services (206,207). Menopause is the permanent cessation of menses that occurs after the cessation of ovarian function and is frequently preceded by oligoovulatory bleeding. Menopause usually occurs in women between 45 and 55 years of age, with a mean age of 51, and is suspected after 6 months of amenorrhea. The diagnosis can be confirmed by findings of low estradiol and high follicle-stimulating hormone (FSH). These tests can be performed if the diagnosis is in question. The age of onset of menopause can occur earlier in women with chronic disease states such as thyroid disease, diabetes, and in diseases such as MS (208,209). Symptomatology may also vary in its manifestation and effects on women with disabilities. For example, women with neurologic conditions that are particularly heat sensitive, such as MS, may have difficulty distinguishing hot flashes and vasomotor instability of menopause from symptoms commonly attributed to their disability (210).

Women with physical disabilities may be at higher risk than their nondisabled counterparts for the consequences of estrogen deprivation such as osteoporosis or difficulties with skin integrity. For example, a woman with weakness or paralysis from a neurological disability may have extensive immobilization osteoporosis that occurred around the onset of the disability (211,212). She is still at risk for age-related and menopause-related bone loss but will be entering these phases with a lower bone mass to begin with. If she is at increased risk for pressure sores from immobility or loss of sensation, she may note increasing difficulty with maintenance of her skin integrity and turgor around the time of menopause with the loss of estrogen.

Many women begin experiencing menopausal symptoms, particularly vasomotor symptoms, before cessation of menses. Provided they have no other contraindications to its use, low-dose combination oral contraceptives often provide the dual benefit of cycle control and amelioration of hot flashes. Other noncontraceptive benefits, such as reduction in risk of ovarian carcinoma, endometrial carcinoma, and preservation of bone mass, prompt some women to continue them until menopause is confirmed by evaluating gonadotropin levels. Though hormone replacement therapy (HRT) has been believed to be a safe and effective agent for many postmenopausal women and has also been used in the hope of preventing cardiovascular disease, osteoporosis, and even cognitive decline, recent studies have questioned the wisdom of routine use. The Heart and Estrogen/Progestin Replacement Study Follow-up (HERS II) found that HRT was not effective in preventing coronary artery disease and in fact, increased the risk of biliary tract disease and VTE (213–215). The Women’s Health Initiative (WHI) arm that evaluated estrogen plus progesterin was stopped early, as this large randomized controlled trial found an increased incidence of coronary events, strokes, pulmonary emboli, and invasive breast cancer in those receiving the combination HRT versus those receiving placebo. Data from the estrogen-only arm of WHI demonstrated similar findings. The WHI study did indicate some potential benefit in those receiving combination HRT for a lower risk of colorectal cancer and hip fracture (216). Physicians no longer use HRT as primary preventive treatment of heart disease or osteoporosis, as many other more effective treatments are effective (215,217). Aspirin, β-blockers, angiotensin-converting enzyme inhibitors, statins, and other lipid lowering strategies are important agents to consider as preventive strategies for coronary heart disease and cerebral vascular disease (215). A number of agents, including bisphosphonates and reloxifene, are useful in preventing fractures in women with low bone mineral density. Patients experiencing acute symptoms related to estrogen deprivation, including vasomotor symptoms, should realize that relative risks of major cardiac events and invasive breast cancer were low and only demonstrable after 4 years of therapy in the WHI trial. Immediately following release of the WHI trial data, The American College of Obstetricians and Gynecologists (ACOG) formed a task force to review and revise previous recommendations regarding HRT. ACOG continues to recommend that hormone therapy be individualized and that it “can still play a role in the treatment of menopause, provided you use the medications for appropriate reasons and after weighing the benefits and risks” (218).

Similarly, the North American Menopause Society (NAMS) position statement stresses the importance of individualization of therapy noting that providers cannot assume that all benefits and risks of HRT apply equally to all populations, at all ages, and for all durations of use (219). Given the remarkable heterogeneity of disability and its clinical sequelae, the need to specially tailor hormone replacement is even greater than among women without disabilities.
DEPRESSION

It has been well established that women are troubled by significant depression at a rate approximately twice that of men (220–222). Among the causes hypothesized to explain this disparity have been physiology, particularly hormones, social oppression including restricted choices and narrow standards for appearance and behavior, and socially imposed powerless-ness. The National Institute of Mental Health Web site elabo-rates “…many women face the additional stresses of work and home responsibilities, caring for children and aging parents, abuse, poverty, and relationship strains. It remains unclear why some women faced with enormous challenges develop depres-sion, while others with similar challenges do not” (223). The reverberating effects of undiagnosed and underdiagnosed emo-tional disorders are hard to estimate. Moreover, most depres-sive disorders are treatable and HCPs should be aware of the signs, symptoms, and range of treatments available.

It has also been established that women with disabilities experience depression at higher rates than women without disabilities (224). The high comorbidity of depression and anxiety (225) also makes it likely that a range of clinically significant emotional disorders differentially impact women with disabilities. The multiple minority status of women with disabilities can elevate their risk for stress and mental health problems (226). Nosek et al. (224) discuss the antecedents of depression in women with physical disabilities. Factors include socioeconomic disadvantage, functional limitations, health disparities and secondary conditions, health behaviors, violence, self-esteem, sexuality, stress, and environmental barriers. The complexities and pressures of daily living with a disability command a balancing act of finding transportation, making appointments, filling out forms, negotiating with service agencies, managing personal assistants, worrying about equipment breakdowns, and coping with discrimination. For women, who are more likely than men with disabilities to confront poverty and social isolation, the lack of adequate resources can compound the routine pressures of disability, leading to depression and feelings of hopelessness.

Furthermore, access to intervention such as psychotherapy can be limited for women with disabilities due to lack of health insurance coverage and lack of practitioners with disability knowledge. Women who live in rural areas may be particularly affected by under diagnosis and lack of treatment. In one study of rural centers for independent living in nine different states, the majority of the participants reported symptoms of moderate to severe depression, with almost one in five reporting suicidal thoughts (227). Moreover, 36.6% of the women in the sample had not received any treatment for depression in the 3 months prior, and nonwhite women were more likely to be untreated. The authors conclude that social factors impact the causes of and treatments for depression. “Mobility limitations and pain can have a severe negative effect on a woman’s ability to engage in social activity. Diminished social networks and increased social isolation can increase the risk of depression, abusive relationships, and illnesses” (227).

A survey study of women with disabilities in Canada found a significant association between suicide, depression, and abuse (228). Of 371 women with a variety of disabilities, over 60% had contemplated suicide and over 45% of the contemplators had attempted suicide at least once. Approximately half had experi-enced physical abuse or sexual abuse, and two-thirds had experi-enced emotional abuse. The more types of abuse a woman had experienced, the more likely she was to have considered suicide. Aside from abuse, the two factors mentioned most by the women as contributing to their suicidality were poverty and isolation.

Treatable depression in disabled women remains undiagnosed when health service providers incorrectly assume that depression is a natural concomitant of disability. While depression is an expected feature of the initial disability adjustment process, it should be noted and addressed with support, counseling, and other interventions appropriate to situational distress. Beyond the adjustment phase, depression in women with disabilities calls for a thorough exploration of specific causes (whether directly related to the woman’s disability or not) and a review of potentially effective interventions. Peer support can be a powerful adjunct to professional treatment.

Unfortunately, many women’s clinics and community mental health centers are inaccessible to women with dis-abilities due to physical, economic, and communication bar-riers. Residential treatment programs for substance abuse and psychiatric disorders commonly refuse to admit women who use equipment or need physical assistance with activities of daily living. Access is limited for those seeking sexual counseling, psychotherapy, and suicide intervention, from therapists who are adequately informed about the stressors and options resulting from the interaction of being a woman and being disabled. A concerned physician can be an important ally in a disabled woman’s journey through depression, from identification of the problem to support in surmounting the barriers to adequate mental health services. The management of depression is a critical facet of disabled women’s health care, and the consequences of failing to address it can be grave. Psychosocial support can also mitigate symptoms of depression. Physicians can help link their patients to community resources for peer support and opportunities for self-expression. Physicians can acquaint themselves with independent living centers, a national network comprising hundreds of centers run by people with disabilities to provide disability-related information, services, and advocacy to others with disabilities. Most cities have one or more such centers, and many offer peer support groups and domestic violence programs as well as information on housing, health services, transportation and in-home PAS. Additionally, in many cities, there are disability rights and consumer organi-zations that can be contacted for resources helpful to women with disabilities. State, county, and municipal government offices routinely have departments addressing the needs of citi-zens with disabilities. By contacting these organizations and programs, physicians and the women with disabilities they serve can inquire about local recreational, artistic, employ-ment, and peer support programs that welcome women with disabilities and provide potential psychosocial resources.
VIOLENCE AND ABUSE

Despite the increase in research and public health campaigns on domestic violence over the past decade, domestic violence remains a complex social problem that is frequently invisible to health providers and the public. It is a crime of power that utilizes fear to control a victim and is an everyday reality in the United States. Most critically, violence and abuse are issues that pose severe health problems to women with disabilities (229–232).

Seminal work by Sobsey (233) long ago demonstrated a dynamic that intrinsically linked violence to disability, both as an outcome and as a means for establishing societal perceptions of women’s vulnerability. This association creates a cycle of disability and violence, begetting more disability and violence, that easily entraps but proves difficult for a woman to escape (233) (Fig. 58-7). Recent evidence remains conflicted, but many experts agree that women with disabilities experience twice the rate (or more) of all forms of abuse and violence compared to other women (229–232,234). In one in-depth qualitative study of 181 women, women with physical disabilities were more likely to experience intimate partner abuse if the women perceived themselves as more sexually unattractive than women with lesser impairments. Negative perceptions of low body and sexual self-esteem, coupled with the woman’s desire to be partnered, increased a woman’s vulnerability to getting into and staying in abusive relationships (231).

The Center for Research on Women with Disabilities (CROWD) at Baylor College of Medicine defines abuse and neglect for people with disabilities (235). These definitions are also echoed by research on abuse involving people with disabilities which found that female gender, people who were younger, unemployed and uncoupled, were at risk for abuse (229). The reliance upon others for personal assistance (236) poses additional risk and makes it harder for a woman to escape the situation. Removing adaptive equipment or withholding needed care may be one way the abuse is experienced in the context of these relationships.

Researchers from CROWD identified that women with disabilities were more likely than their nondisabled peers to experience abuse from a significantly greater number of perpetrators that included attendants and disability service/health providers. Other offenders included spouses, ex-spouses, family members, strangers, and dates. Secondly, the CROWD study found that women with disabilities were more likely to experience abuse for longer periods of time (235) than women without disabilities. Environmental and attitudinal barriers may play a critical role in accounting for significantly longer periods of abuse (235). These two findings were also affirmed in a qualitative study that investigated care-related abuse in women with SCI (231).

Women revealed that they often refused to confront abusive PAS providers because of fear of retribution and to avoid having to find a new provider (237). Other studies on PAS abuse conducted by the Center for Self-Determination at Oregon Health and Science University participants described several barriers to addressing abuse: “(a) difficulty recognizing it and having their experiences validated; (b) feelings of shame or embarrassment; (c) lack of PAS providers and emergency backup services; (d) fear of institutionalization or loss of their children if they reported the abuse; (e) lack of access to abuse services such as crisis services, support groups, and domestic violence shelters” (236). Findings from this study also indicated that women found professionals of little help in helping them in identifying or addressing PAS abuse, but indicated that they would welcome being asked about these experiences and being assisted to discover management options (236).

Probing deeper into the nuanced differences between the experiences of disabled victims of intimate partner abuse and
other abused women, Copel (230) studied the varied forms of disability-specific abuse in conjunction with the stresses women with physical disabilities identified as triggers to the abuse episode, their responses to the abuse, and the aftereffects of the episode (Fig. 58-9). When the stress level in the home exceeds the perpetrator’s ability to cope, the likelihood of abuse increases. Such factors which can contribute to the stress may include financial problems, complex health issues, shortage of resources, a breakdown in support systems, and the perception of unmet sexual and emotional needs of the perpetrator. In each of the 25 transcripts analyzed by Copel, the woman’s disability was mentioned as a stressor, a burden blamed for the conflict. The study noted that as the male partner’s stresses increased and his ability to cope decreased, his verbal abuse escalated, sometimes exploding into acts of physical violence. While the start and progression of the violent process was similar to the patterns of violence experienced by nondisabled women by their intimate partners, Copel found that unlike their nondisabled counterparts, women with disabilities experienced no separation, or honeymoon periods from the violence. Rather than a period of separation that was followed by remorse, apologies, contrition, or reconciliation, women with disabilities after the abusive episode, experienced separation but indifference and learned to take their cues from their abusive partners and act as if nothing happened. The women rationalized the violence and lack of remorse from their male partners by commenting that their disability made them “less than” or “different from” other women. The women often stated that the male abusers regarded them in derogatory or dehumanizing terms as “nothing but a disability” or “physically damaged,” which decreased the quality of the couple’s life. The unexpected finding that the aftereffects of violence differ significantly between women with disabilities and nondisabled women, caused investigators to construct the “Proposed Model for Partner Abuse in Physically Disabled Women.”

Copel also found that the women were generally unsuccessful in their ability to stop the conflict or in their desire to find peaceful solutions, and it was only after the violence escalated and culminated into crisis that they were able to establish a connection with a health care provider or other women to help them recognize and label the partner’s behavior as abuse. It was the crisis that was the main link to the health professional. Up until this point, it was difficult for the women to understand how to empower themselves or extricate themselves from the abuse (230).

Implications of this research for rehabilitation providers, clinicians, and advanced nurse practitioners are profound and cannot be overstated. Providers need to be proactive in working with women with disabilities and their partners to help them connect to resources that can alleviate stressors. Having options for disability-sensitive, therapeutic and social support interventions, for example, such as support groups or counseling, may help reduce social isolation, enhance self-esteem, educate the woman about what constitutes abuse, and help her implement health boundaries in relationships. Of course, interventions to enhance the woman’s self-reliance and self-efficacy, though skills training, assistive devices, etc., can be extremely helpful. It is also important to assist women in creating safety plans when working with home or family care providers, so a plan is in place in case it is ever needed.
It is essential that providers routinely assess for abuse among women with disabilities, talking to them in private without the personal assistant or partner present, and help them recognize the various forms abuse can take. Yet, a survey conducted by CROWD found that less than 20% of rehabilitation providers screen for abuse and violence (235). Despite the low percentage of disability providers who screen for abuse, there is evidence revealing that disabled women would welcome providers asking these questions (236) and that asking abuse screening questions can increase the detection of violence and abuse from 0.8% to 7% (238). The Abuse Assessment Screen-Disability (AASD) screening tool developed by researchers at the CROWD (Fig. 58-10) is one excellent tool for screening women with disabilities for general and disability specific abuse.

Finally, providers and clinicians must distinguish the subtle differences between the exacerbations of the woman’s health condition, or the effects of her disability, from evidence of abuse such as severe anxiety, panic attacks posttraumatic stress disorder, depression, etc. Since the passage of the ADA, domestic violence shelters in some states have improved their services to victims with disabilities and have committed to implementing disability-specific abuse screening protocols and disability-specific safety plans (239). But continuing limitations regarding the shortage of disability-friendly resources such as accessible housing, battered women’s shelters, transportation, quality personal care attendants, disability-aware law enforcement officers, and accessible courtrooms can cause women with disabilities who are victims of partner or caregiver abuse to feel revictimized and frustrated in their attempts to obtain help (236,240).

The Disabled Women’s Network of Canada were among the first to determine that the abuse experiences of women with disabilities are often not reported to law officers because of their fear and dependency on the offender (241). Anecdotal evidence suggests that, like other abused women, those who are disabled often resist reporting their experiences of abuse because of denial, guilt, self-contempt, and fear of loss of financial security. However, some experts believe that disabled women are

---

**FIGURE 58-10.** CROWD Abuse Assessment Screen-Disability (AASD).
often discouraged from disclosure because they are commonly regarded as less credible and easily discredited as witnesses by virtue of having a disability (242). Threats from abusers to withdraw assistive devices, such as wheelchairs and respirators, as well as personal assistant services add to the fears of disabled women and further inhibit their disclosure of violent incidents. These threats and fears also serve as disincentives to leave abusive relationships with caregivers and family members and reinforce the cycle of violence. Other experts speculate that a lifetime of negative encounters with law enforcement, judicial and social systems, coupled with a fear of becoming institutionalized or other losses of self-autonomy, may prevent victims with disabilities from disclosing encounters of abuse and violence (243–245).

Past research exploring sexual abuse among women with physical disabilities, compared to women with developmental disabilities, showed that sexual assault among women with physical disabilities is commonly perpetrated by male strangers rather than acquaintances or family members (as is often the case with women with developmental disabilities) (236). However, a recent study of 16,672 women survivors of sexual assault that studied sexual assault patterns among women with and without disabilities, which included a sample of 1,711 women from a range of disability specific groups, including mental health disabilities, physical, sensory, and cognitive disabilities, revealed that the majority of assailants for all groups were known assailants. Other significant findings included that the survivor’s home was the most frequently reported setting for assault among all groups and that former and current partners “represented significantly greater proportion of assailants for women with physical disabilities, or hearing impairments, and the least for women with cognitive disabilities.” An unexpected finding of the study underscored that the average time from assault to reporting the abuse does not differ significantly between women with and without disabilities or between disability-specific groups. However, the researchers speculate that women with disabilities may be more likely to tolerate more abuse before they seek support (246).

Health care providers are cautioned to discriminate between the signs of sexual abuse and those of disability. Frequently, symptoms of abuse are confused with the sequelae of a woman’s disability, or they may be overlooked completely. Common symptoms of sexual abuse include pain, bruises, and bleeding in the genital area, behavioral disturbances, loss of weight, fear of sexual intimacy, and fear of intervention (8). Practitioners are also urged to be mindful in recognizing and treating the symptoms of posttraumatic stress syndrome (PTSD), which can manifest in survivors of violence. When she has had previous brain trauma acquired through violence, the compounded effects may overwhelm the woman and require physicians to invoke the expertise of mental health workers familiar with such cases.

Moreover, physicians must take a supportive and facilitative role in identifying and treating violence and abuse. Many states have mandatory reporting laws that require health professionals to report incidents of abuse detected in their disabled patients or clients. Since the specifics of these laws can vary from one state to another, it is essential that health providers educate themselves about their responsibilities as mandated reporters. If a provider is mandated to report on abuse of a woman with a disability, without her knowledge, the implications for the woman can be uncertain. Some women may risk institutionalization or the loss of custody of their children if they are not engaged with the provider/reporter in a problem-solving process before the report is made to the authorities. It is therefore important for providers to become familiar with local domestic violence or disability advocates who can help collaborate and advocate with the disabled abuse victim in an empowering fashion that promotes legal compliance, safety, and self-determination. As pioneers of domestic violence advocacy remind us, domestic violence requires practitioners to move beyond the biomedical model and collaborate with advocacy groups committed to ending violence against women. They can assist by working to facilitate access to high-quality personal assistants and to increase living options beyond nursing homes and other institutional facilities. Health care providers can help prevent sexual abuse in women with disabilities by collaborating with other social services to develop comprehensive outreach strategies that are accessible to women with physical, cognitive, sensory, and communication disabilities, including those living in institutions.

Offering high-quality health services to disabled women oblige providers to make their services welcoming, with accessible equipment, signers, and TTYs (teletype writers). It also dictates that they become knowledgeable about local accessible shelters and their policies regarding personal assistants and children. Shelters and support services are accessible only if they can be accessed by a woman at any time regardless of her physical, sensory, or communication disabilities. Physicians should be aware of their role in breaking the cycle of violence and disability. They are encouraged to work with disabled women’s centers and social services to prevent isolation to the fullest degree by involving the expertise of disabled women in the design and execution of outreach and safety planning strategies. For example, transportation is a key component in any emergency. It should be factored into all safety plans. For practitioners who are interested in grasping the details of comprehensive safety planning procedures and solutions, a protocol has been developed by The Washington State Coalition Against Domestic Violence for the following disabilities: cognitive, sensory, mental health, mobility disabilities, and hidden disabilities (243). Most important, health providers can assist in breaking the cycle of violence and disability by implementing a system of detection, reporting, and prosecution for cases of assault. Assaults will continue unless practitioners “represented significantly greater proportion of assailants for women with physical disabilities, or hearing impairments, and the least for women with cognitive disabilities.” An unexpected finding of the study underscored that the average time from assault to reporting the abuse does not differ significantly between women with and without disabilities or between disability-specific groups. However, the researchers speculate that women with disabilities may be more likely to tolerate more abuse before they seek support (246).
the help of, but often in spite of powerful social institutions such as health care. How is that almost two decades after passage of the ADA, women with disabilities often still feel they do not belong, shuttled between primary care, reproductive health and rehabilitation providers, each feeling ill-prepared to address the woman’s health care needs in the context of her disability. The reasons given are many. It may be the inaccessible office, the poorly trained staff, the discomfort with one’s lack of knowledge and experience, but the result is the same: exclusion. To address the resultant health care disparities, poor quality care, preventable secondary morbidities, and our systems of care and training programs must change, and HCPs need to lead the charge. Health care institutions and educational programs need to recognize that disability is part of life, must be taught and planned for, and that universal design principles should be the standard. Researchers need to gather information about disability-specific variables when studying general health conditions, in order to further our understanding of common medical conditions in the context of disability. This does not mean that all care providers need to provide care to all. Interdisciplinary and collaborative care in our ever specialized world of medicine is here to stay. But it does mean recognizing that women with disabilities are first and foremost women and need to have their basic health care needs addressed in a competent and inclusive fashion. The narrative of exclusion in the health care setting is a tiresome one; we need a new narrative for the next edition of this textbook, full of creativity, possibilities, and progressive models of care—and each reader of this chapter can help to write it!

ACKNOWLEDGMENT

The authors would like to acknowledge the following people who were very generous with their time, resources, and expertise in advising on the new edition of this chapter. Specifically, we thank Megan Kirshbaum PhD, Judi Rogers OTR/L, and Christi Tuleja MS, OTR/L from Through the Looking Glass for their expertise on parenting and child care for women with disabilities. Rosemary Hughes PhD provided valuable advice on domestic violence and women with disabilities. We would like to thank Chris Wisniewski, MSW, Katrina Bullock and Rebecca Thompson from the RIC Women with Disabilities Center who provided technical assistance. And finally, we would be remiss if did not acknowledge the invaluable contributions of the authors involved in the earlier versions: Carol J Gill PhD, and the late Sandra Welner MD. Their influence, passion, and wisdom are still palpable in this updated chapter.

REFERENCES


118. Disabled Women’s Network (DAWN). I Want to be a Mother I have a Disability: What are My Choices? Toronto, ON: DAWN; 1993.


124. Morrell MJ. Guidelines for the care of women with epilepsy.


231. Hassounah-Phillips D, McNeff E. “I thought I was less worthy”: low sexual and body esteem and increased vulnerability to intimate partner abuse in women with physical disabilities. Sex Disabil. 2003;26:15–28.


Aging and Rehabilitation

Aging, an integral part of living, is typically accompanied by gradual but progressive physiologic changes and an increased prevalence of acute and chronic illnesses. Although neither a disease nor a disability per se, aging nonetheless is associated with a higher incidence of physical impairment and functional disability. Many of these functional difficulties occur from the interactions of decreased physiologic reserve with chronic illness. Ongoing research suggests effective interventions to prevent, delay, minimize, or reverse such physiologic declines (1,2). Appropriate roles for geriatric rehabilitation accordingly include not only intervening to reverse disability caused by specific disease or injury (e.g., stroke, hip fracture) but also contributing to preventive gerontology by virtue of promoting structured physical fitness (i.e., wellness) programs and early rehabilitation for common musculoskeletal disorders to avoid progression to disability (1–4).

Significant contributions of rehabilitation to care of older adults include functional assessment (including evaluation of underlying impairments contributing to functional loss and disability) with realistic goal setting, interdisciplinary team care, and efficacious adjustment of therapy interventions (e.g., timing, setting, intensity) to prevent, reverse, or minimize disability (2,5,6). Given the burgeoning number of older persons living longer, Rusk’s observation, as modified by Kottke, becomes ever more relevant: “As modern medicine adds years to life, rehabilitation becomes increasingly necessary to add life to these years” (7). The goals of this chapter are to:

- Discuss relevant aspects of epidemiology and physiology of aging
- Examine the impact of the environment on function in older adults
- Review data on patterns of functional decline in older adults, with particular attention to the effects of hospitalization
- Identify general principles of assessing functional loss as a basis for formulation of rehabilitation plans of care
- Detail appropriate rehabilitation interventions for a number of disabling clinical syndromes and conditions that occur frequently in older adults.

AGING—EPIDEMIOLOGY AND PHYSIOLOGY

Demography and Epidemiology of Aging

Demographic Imperative

The context for the increasing interest, and concern, about health care needs of older adults is found in demographic projections of an expanding elderly population in the United States and other developed countries. At the turn of the 20th century, one of every 25 Americans (4%) was 65 years of age or older. By 1994, this population had increased to one of every eight Americans (12.6%), or 33.2 million (8,9). While the elderly population grew 11-fold during this interval, the under-65 population only increased by a factor of three. Current projections indicate that 86.7 million, or one of every five Americans (20.7%), will be 65 years of age or older by the year 2050 (10). By 2029, all “baby boomers” (Americans born post-WW II, between 1946 and 1964) will be 65 years or over (10). Between now and 2030, the fastest growing segment of the population will be the 65 to 74 year olds, increasing from the current 6.3% to 10.4%. Between 2030 and 2050, however, the 75+ population will outpace the 65 to 74 cohort, reaching 11.6% of the population (every ninth American), while the 65 to 74 year olds will account for 9.0% (10).

Distribution of the elderly is uneven across the United States, with 50% living in just nine states. While California has the greatest number of citizens over the age of 65, Florida has the highest proportion of elderly (18.6%) (8). There are also ethnic aging trends, with projections that by 2050 the proportion of older white individuals will have decreased to 67% (from 87% in 1990), with corresponding increases of older Hispanic and black American citizens of 16% and 10%, respectively.

This demographic phenomenon is not limited to the United States, as a number of developed countries (including Italy, Japan, Germany, Sweden, and Great Britain) show 20% or more of their populations over age 65 (8). As of 1994, there were 357 million older people worldwide, representing 6% of the total population.

There is increasing recognition of differences in health care needs and issues among subgroups of older people. Of particular significance from a health care standpoint is the rapidly expanding relative proportions of the population age 65 years and older who are 75 to 85 years of age (old old) and 85 years of age or older (oldest old). These groups include many of the so-called “frail elderly,” with a disproportionately high prevalence of disabilities and consumption of health services (11). Another related dynamic is recognition of increasing racial and ethnic health disparities among older Americans, as the relative proportions of racial and ethnic minorities increase: from 2003 to 2050 the proportion of Non-Hispanic whites will decrease from 83% to 61%, while Hispanics will increase from 6% to 18%, blacks from 8% to 12%, and Asians from 3% to 8% (12).
Compression of Morbidity

Older people also are living longer: 1990 statistics estimated a longevity of 65 years of age of 15.0 years for men and 19.5 years for women; this is projected to increase further to 17.1 and 22.6 years, respectively, by the year 2040 (13). Contributing factors to these increases in longevity include improved access to health care, advances in medical care, overall healthier lifestyles, as well as better health prior to age 65 (14). The increasingly delayed occurrence of death at all ages appears in part to be due to reduced lethality of such diseases as stroke, cancer, and myocardial infarction, resulting from risk factor reduction as well as improved health care interventions (9). Increasingly, people are surviving their initial encounter with these previously fatal diseases, resulting instead in chronic illness. This trend has been termed the fourth stage of epidemiologic transition (i.e., the postponement of death from degenerative diseases) (15). These significant reductions in mortality are associated with an increasing risk for development of various chronic diseases. Certainly, the incidence and prevalence of many potentially disabling chronic illnesses increase substantially among older adults, including arthritis, osteoporosis with associated fractures, stroke, amputation, and various neurodegenerative disorders (e.g., Alzheimer’s disease, Parkinson’s disease) (2,10,11,13).

This demographic imperative has far-reaching implications for increasingly limited U.S. health care resources and dollars: current national direct cost of medical services for older individuals with chronic conditions is in excess of $470 billion (in 1990 dollars), with a projected near doubling by the year 2050 (13). In 2004, the average annual health care expenses were $8,900 for the 65+ population, compared to $3,028 for the under 65 population (10). The old-old and oldest-old groups consume the greatest proportion of resources, and a disproportionate amount of these health care costs represent nursing home and other institutional care (11). While the overall proportion of elderly individuals residing in nursing homes decreased from 6.8% in 1982 to 4.2% in 1999, these rates vary dramatically by age (16). Only 1% of young old (65 to 74 year olds) reside in nursing homes, contrasted with 20% of the oldest old (85 years of age or older) (9). In fact, the latter group comprises 45% of all elderly nursing home residents.

From another perspective, however, the vast majority of the 85-or-older population is not in nursing homes, and half of those in nursing homes do not necessarily need to be there because they have potentially preventable (or reversible) disabilities related to their chronic disorders (17). In fact, it appears that much of the increased health care cost associated with aging (across all health care settings) is significantly related to activity limitation, rather than chronic disease (18). Fortunately, there is increasing evidence that disability among older individuals may be decreasing (16,19). Reports of limitation of activity due to a chronic condition in the 65+ population have continued to decrease, from 38.7% in 1997 to 32.6% in 2006 (12).

These findings continue to lend credence to the concept advanced by Fries in 1980 of “compression of morbidity,” in which he predicted that if the age of onset of disability could be significantly delayed (e.g., with regular exercise, healthier diets, elimination of smoking, improved health care interventions) in the context of a relatively “fixed” life span, then terminal predeath disability could be compressed into a shorter interval (20). He postulated that health care needs for older people would decrease because they would be relatively healthy and functional until shortly before their death. A related prediction was that this anticipated short duration of predeath morbidity, and accompanying disability, would be expected and accepted with acknowledged futility of medical intervention. There is consistent evidence of increased health care costs associated with aging, with one study documenting nearly one half of lifetime health care costs occurring after age 65 (21). Interestingly, although a number of reports have documented dramatically increasing health care costs near the end of life, it should be noted that these are costs of dying, not of aging per se (22).

Further, there is evidence suggesting that the incremental costs associated with extending life may actually plateau or even decrease (23–25). These findings are of critical significance in the context of increasing focus on cost containment and debate over the feasibility and appropriateness of rationing of health care (26). Twenty years after his initial predictions, Fries (and others) cites increasing evidence in support of the trend of compression of morbidity, even though the mechanisms are not clear (27,28). He reiterates the importance of a research agenda focused on delineation of the epidemiology of disability, determination of the fundamental basis of age-associated chronic conditions, and identification of effective interventions for preventing or delaying resulting disability (27). This call to action has been echoed by others as well (1,2,21). On the other hand, Kane raises disturbing questions regarding potentially adverse economic, cultural, and individual consequences of successfully overcoming the aging (and dying) process and urges ongoing dialogue to further explore these ethical questions (29).

Active Life Expectancy

A derivative of research into longevity and epidemiology of aging relates to issues of quality of life, given the increased incidence of frequently disabling chronic disorders such as degenerative neurologic diseases (e.g., Alzheimer’s disease, Parkinson’s disease), degenerative musculoskeletal conditions (e.g., osteoporosis, osteoarthritis), and multisensory losses (e.g., cataracts, presbycusis). One concept that attempts to delineate quality of life for older individuals has been termed “active life expectancy,” referring to the proportion of remaining life span characterized by functional independence (30). This concept, also referred to as “disability-free life expectancy,” has been expanded to consider both physical and cognitive impairments, as well as their interrelationships (28,31). A significant gender difference in active life expectancy with aging has been identified. As can be seen in Table 59–1, older men have a greater proportionate active life expectancy at all ages. However, due to greater longevity, older women enjoy longer actual durations of active life expectancy than older men, until age 85 (31,32). A recent longitudinal British study confirmed a gender difference, with men aged 65 showing 79% active life expectancy (12.1 of 15.3 years) and women aged 65 showing 57% active life expectancy (11.0 of 19.4 years) (28). Further investigation into
dynamics impacting active life expectancy reveals the deleterious effects of diabetes (decreased total life expectancy and active life expectancy at all ages, with a 25% reduction in active life expectancy in 85 year olds) and depression (reduction of active life expectancy by 6.5 years in 70-year-old males and by 4.2 years in 70-year-old women) (33,34).

Comorbidity, Frailty, and Disability

Although the increasing incidence and prevalence of (often multiple) chronic diseases with aging is well documented, there is no one-to-one correlation between either disease and illness (35) or disease and disability (19). A significant proportion of older people are limited in the amount or kind of their usual activity or mobility secondary to chronic impairments: over 60% of adults with functional impairments due to chronic health problems are 65 years of age or older (13). However, disability in this older population can also be very dynamic, with frequent transitions between periods of independence and disability and levels of disability (36,37). Also, the overall health of progressive cohorts of older persons has been changing. Although there have been predictions that future generations may well be healthier than current generations, due in part to higher levels of education and health awareness (36), there are disturbing trends of increasing obesity and diabetes in older adults, with significant associated morbidity and mortality (10).

Fried et al. have helped to clarify the dynamics of interrelationships among comorbidity, frailty, and disability—terms frequently used interchangeably—by providing discrete definitions of each entity and describing the synergistic impact of each on the other(s) (38). As seen in Figure 59–1, Disability accordingly is defined as “difficulty or dependency in carrying out activities essential to independent living” (such as activities of daily living [ADLs], mobility, instrumental activities of daily living [IADLs]), while Frailty is “a physiologic state of increased vulnerability to stressors that results from decreased physiologic reserves, and even dysregulation, of multiple physiologic systems.” Frailty appears to represent an aggregate expression of risk resulting from age- or disease-associated physiologic accumulation of subthreshold decrements affecting multiple physiologic systems (38). Fried cites evidence in support of a phenotype of the clinically frail older adult, characterized by the presence of a critical mass of three or more “core elements” of frailty (weakness, poor endurance, weight loss, low physical activity, and slow gait speed) (39). Comorbidity is commonly defined as the concurrent presence of two or more disease processes in the same individual. Fried suggests that comorbidity is the aggregation of clinically manifested diseases present in an individual, while frailty is the aggregate of subclinical losses of reserve across multiple physiologic systems (38).

### TABLE 59.1 Active Life Expectancy (Remaining Years of Functional Independence, Compared to Projected Longevity) by Gender and Age Cohort

<table>
<thead>
<tr>
<th>Age</th>
<th>Males</th>
<th>Females</th>
</tr>
</thead>
<tbody>
<tr>
<td>65</td>
<td>82% (11.9 of 14.4 y)</td>
<td>73% (13.6 of 18.6 y)</td>
</tr>
<tr>
<td>85</td>
<td>50% (2.6 of 5.2 y)</td>
<td>35% (2.3 of 6.4 y)</td>
</tr>
<tr>
<td>95</td>
<td>20% (0.6 of 3.2 y)</td>
<td>10% (0.4 of 3.7 y)</td>
</tr>
</tbody>
</table>


### FIGURE 59-1. Definitions of frailty, comorbidity, disability.

Prevalences—and overlaps—of comorbidity, disability, and frailty among community-dwelling men and women 65 years and older participating in the Cardiovascular Health Study. Percents listed indicate the proportion among those who were frailty-dwelling men and women 65 years and older participating in the Cardiovascular Health Study. From Fried LP, Ferrucci L, Darer J, et al. Untangling the concepts of disability, frailty, and comorbidity: implications for improved targeting and care. J Gerontol Med Sci. 2004;59(3):255–263.

In summary, an increasing proportion of older people are frailty, comorbidity, and disability (Fig. 59–2), noting that frailty and comorbidity each predict disability, while disability appears to exacerbate frailty and comorbidity (38). While medical care for each of these entities is typically complex, particularly when they coexist, Fried specifically notes that rehabilitation interventions for disabled older adults to regain function and/or prevent further functional decline need to factor in recognition and treatment of frailty and comorbidity to maximize likelihood of success. She cites growing evidence to suggest that frailty, comorbidity, and disability may be preventable, but with different intervention strategies (38). The importance of prevention is reinforced by the clear association of each entity, especially when concurrent, with higher health care costs and poorer survival (38,40). Recent research is further differentiating between various subtypes of disability in older individuals (transient, short-term, long-term, recurrent, and unstable) and their relative clinical significance (41).

In summary, an increasing proportion of older people are living longer and are at increased risk of developing varying (and changing) degrees of comorbidity, frailty, and functional losses with disability. The challenge for health care providers, accordingly, is to try to prevent the onset/progression of these entities with early and effective medical and rehabilitation interventions, to reverse or at least minimize their deleterious effects on health and function.

Successful Aging
Distinctions have been made between aging processes representing "primary aging" (i.e., apparently universal changes that occur with aging, independent of disease and environmental effects) and "secondary aging," which includes lifestyle and environmental consequences and disease associated with the aging (42,43). A number of tenets associated with aging research are being reexamined, particularly with the observation that a pathologic process may exaggerate an aging process believed to be normal, even before the disease is detected clinically (42). There is increasing evidence that the nonpathologic processes of aging are distinct from, but not necessarily independent of, the pathologic processes of disease (38,44).

Most studies of normal aging have focused on the physiological and biochemical changes occurring with aging, with explicit exclusion of disease. However, it is increasingly apparent that such factors as personal habits (e.g., diet, exercise, nutrition), environmental exposures, and body composition may have significant impact on observed aging changes (43). Rowe has proposed a conceptual distinction between "successful aging" and "usual aging" (45). He suggests that "successful aging" could be characterized by minimal or no physiologic losses in a particular organ system and would comprise a relatively small subset of the total "normal" (i.e., nonpathologic) aging population. The remaining majority of "normal" older adults demonstrate "usual aging," with gradually progressive but significant declines in various physiologic functions.

The significance of this concept lies in the implications for modifiability of usual aging by virtue of addressing such variables as level of physical activity, diet and nutrition, and environmental exposures (27,43,44). This principle is demonstrated in studies documenting the effects of exercise, diet, and drugs on the usual aging observations of carbohydrate intolerance. Rowe proposes that geriatric research into health promotion initiatives concentrates on increasing the proportion of older adults who "successfully age" by identifying and modifying extrinsic risk factors contributing to "usual aging" and decreasing the manifestations of "pathologic aging" by preventing or minimizing adverse effects of acquired disease processes (45). This would reinforce the previously described concept of compression of morbidity, with greater active life expectancy. Indeed, studies are helping to determine which factors distinguish high-functioning older adults from other populations of older adults (28,33,43,46).

Theories of Aging
With continuing research, it appears likely that there is no single cause of aging (44,47). The current concepts of aging characterize the process as extremely complex and multifactorial and suggest that various theories of aging should be viewed not as mutually exclusive, but rather complementary (47). From this perspective, hypotheses based on passive (i.e., random) and/or active processes of genetic programming could be considered in conjunction with superimposed nongenetic
mechanisms (e.g., environment, lifestyle), producing varying individual vulnerability (42,44,45,48). Certainly this would help explain the well-documented phenomenon of differential aging, whereby individuals of the same species appear to age at different rates (44,46). Multiple levels of research suggest that rates of aging are affected to varying extents by heredity, lifestyle, environment, occurrence of disease, and psychological coping abilities (17,32,42,46).

Active investigation continues in the areas of neuroendocrine pacemakers, telomere shortening, and attenuation of inducible stress responses (48–51). A number of studies have also focused on the phenomena of apoptosis and autophagy, related to cell death of proliferating and postmitotic cells respectively, as well as mitochondrial degradation in long-living cells (52,53). There is some evidence suggesting that pathologic stimulation of apoptosis may result in a number of degenerative disorders commonly associated with aging, whereas inhibition appears to be associated with a variety of forms of cancer (52).

**Physiology of Normal Aging**

The normal aging process involves gradual decreases in organ system capabilities and homeostatic controls that are relatively benign (i.e., asymptomatic or subclinical) in the absence of disease or stress (35). Although the older person progressively adapts to these changes without need (or desire) for outside intervention, the steady decreases of physiologic reserves make older adults potentially vulnerable to functional decline as a result of acute and/or chronic illnesses (42,54).

**Characteristics of aging include:**

- Decreased reserve capacity of organ systems, which is typically apparent only during periods of exertion or stress.
- Decreased internal homeostatic control (e.g., blunting of the thermoregulatory system, decline in baroreceptor sensitivity).
- Decreased ability to adapt in response to different environments (e.g., vulnerability to hypothermia and hyperthermia with changing temperatures, orthostatic hypotension with change in position).
- Decreased capacity to respond to stress (e.g., exertion, fever, anemia) (35).
- The end result of these age-related declines is an increased vulnerability to disease and injury, or frailty (38).

**Problems in Study Design**

Inherent with studying physiologic changes in aging adults are several potentially confounding variables, which are unique to this population. Awareness of these dynamics will facilitate more accurate interpretations of clinical studies of older adults, with particular reference to generalizability.

**Definition of Normal**

A significant concern, in view of the heterogeneity of the aging population, is what is truly normal. As noted, there is great variability in rates of aging among healthy elderly and wide variations in individual performance. Further complicating any analysis is a superimposed dispersion of skills due to frequency of significantly impaired function from disease, environment, and lifestyle (32,35). More than 80% of the population over 65 years has at least one chronic disease and 50% two or more disorders (31). Of concern is whether the relative minority of older people who have escaped serious illness should be considered “normal” for the purpose of studies of aging and whether the results of such studies can be generalized to the rest (majority) of the older population.

On the other hand, it is important clinically to be able to differentiate the physiologic consequences of aging (i.e., normal aging) from those of accompanying disease (i.e., pathologic aging) (3). Because detection of disease depends on determination that a patient is other than normal, it is critical to define appropriate age-adjusted criteria for clinically relevant variables in the elderly (55). Although many laboratory values do change gradually with aging, abnormalities should not be a priori attributed to old age. In fact, a number of age-related changes may resemble the changes associated with a specific disease (44). For example, an age-related decline in glucose tolerance is well documented. So dramatic is this change that most people over 60 years of age would be diagnosed as diabetic if traditional criteria, based on studies of primarily younger patients, were applied (56).

**Methodology Limitations**

A number of methodological problems are associated with the study of aging. Well recognized are frequent discrepancies in age reporting, with a tendency to distort upwardly (57). This is coupled with difficulties in verifying reported ages, due in part to lost or nonexistent birth records.

Another major problem in the design and evaluation of aging studies is the relative validity of both cross-sectional and longitudinal studies. Cross-sectional studies, although easier and less costly (in time and money) to perform, often overemphasize (but may also underestimate) age-related changes (58,59). This can result from a cohort bias, due to significant differences in educational, nutritional, health, and social experiences of people born in different decades. Contributing further to this distortion is the high proportion of elderly in the United States who were foreign born, with relatively less schooling. This has implications in particular for studies of psychological and cognitive changes with aging (19).

On the other hand, longitudinal studies tend to underestimate changes due to aging, primarily as a result of withdrawal and survivor biases with high drop-out rates (46,54,58). Some studies have experienced as much as a 50% drop-out rate over just a 10-year period, leading to questions of self-selection for relative preservation of function (and again, the issue of “super-normals”). Subtle changes in methodology over time may introduce laboratory drifts that are difficult to differentiate from true age-related changes (8). A further concern with serial measurements is the potential for distortion due to learning effects.

**Mean Versus Maximal Performance**

Another issue in the characterization of aging is that a focus on “average” or “mean” changes in various parameters can
hide remarkable individual variation, particularly regarding peak performance (20). Consider marathon running, which tends to attract a very select (“supernormal”) population, and, thus, individuals with a higher maximal aerobic capacity. For example, a 50-year-old male with a marathon time of 3.5 hours is in the 99th percentile for his age group, yet this same time would not be an age group record until over 80 years of age. Although there also is a slow linear decline in maximal aerobic performance with aging based on world age-group capacity, this is only on the order of about 1% per year between the ages of 30 and 70 years (20).

Effects of Age on Organ System Performance
There are several general principles regarding aging effects on the performance of various organ systems (60).

Wide Individual Differences in Rate of Aging
Variation between healthy people of the same age is far greater than the variation due to aging alone, and the range of variability increases with aging (20). Linear regressions show average changes with aging, but variation between subjects is so great that it is not always possible to determine accurately if age decrements are linear over the entire age span or whether the rate of decline accelerates in later years (61). However, Fleg et al. recently reported that VO$_{2\text{max}}$ declines more precipitously after the age of 70, especially in men, irrespective of habitual physical activity (62).

Different Organ Systems Age at Different Rates
There is great individual variation in the rate of decline for various organ system functions (35). For instance, there is up to a 60% decline in maximal breathing capacity with aging but only a 15% decline in nerve conduction velocity and basal metabolic rate during the same time interval. Another demonstration of this principle is the localized cellular growth, aging, and death that occur continuously in some tissues and organs (e.g., hematopoietic system, skin, mucosa). Furthermore, a significant decline in function of one organ system (e.g., kidney) does not necessarily entail a similar decline in other organ systems (60).

Age Changes with Complex Performances
Complex performances (e.g., running) will show greater changes with aging because of the need to coordinate and integrate multiple organ system functions (e.g., rate, degree and sequence of muscle contraction, balance, proprioception, vision, cardiovascular response), as opposed to simple performances involving a single system (e.g., renal glomerular filtration) (60).

Age Changes in Adaptive Responses
Adaptive responses (e.g., to temperature change or change in position) are most affected by aging due to a decline in the effectiveness of physiologic control mechanisms (e.g., sensory feedback), which is magnified during stressful situations (e.g., disease, sudden changes in environment) (35,60).

Prevention and Reversibility of Physiologic Decline
There is little question that biologic systems, regardless of the direct effects of aging, can be profoundly influenced by environment and lifestyle (32,35,43). Obvious examples include the deleterious effects of smoking and a sedentary versus active lifestyle (63,64). The modifiability or plasticity of aging is demonstrated by studies in which performance can be improved despite age, within relatively broad ranges (20,65). Physical training can improve or even reverse age-related declines in aerobic power and muscle strength (66–68). These gains have been demonstrated to translate to improvements in functional skills (69,70).

Functional Implications of Organ System Aging
The clinician must be aware of specific age-related physiologic changes to properly understand disease in the elderly because these changes significantly influence not only the presentation of disease but also the response to treatment and potential complications that may ensue. Similarly, such knowledge is essential to understand the underlying mechanisms of functional deterioration secondary to disease and to formulate effective rehabilitation approaches (3). The following is a summary of clinically significant physiologic changes that occur with aging.

Hematologic System
Although anemia (i.e., hemoglobin <13 g/dL in men and <12 g/dL in women) (71) occurs with increasing prevalence with aging, there is convincing evidence that it is not a normal consequence of aging and should be investigated, especially if hemoglobin is less than 10.5 g/dL (72–74). Anemia in older people appears to be due most commonly to iron deficiency (typically from GI blood loss) or chronic disease (such as infection, polymyalgia rheumatica, or cancer) (73). Other potential causes include hemolysis (e.g., secondary to lymphoma, leukemia, or medication effect), B$_{12}$ deficiency (pernicious anemia, diet), or folate deficiency (diet). Of note, D-dimer levels have been shown to double with aging, with even more dramatic increases among blacks and functionally impaired older individuals (75). Increases in the erythrocyte sedimentation rate and C-reactive protein levels also have been noted with aging (76,77).

The functional consequences of anemia can be significant because of further reduction of reserve capacity, such that previously subclinical disease states may become symptomatic (e.g., orthostatic blood pressure changes, change in anginal pattern with lower exercise tolerance) (74,78). This has obvious implications with regard to tolerance of relatively intensive and sustained rehabilitation exercise programs. There is also evidence of correlation of even relatively mild anemia with impaired mobility (79,80). A very anemic older patient may present with nonspecific fatigue and confusion, with the potential for misdiagnosis and mistreatment (72).

There are several related hematologic changes with aging that can affect pharmacokinetics, particularly drug distribution. Decreased drug binding for highly protein-bound drugs (e.g., warfarin, meperidine, tolbutamide) may result in a higher
unbound, or free, drug concentration with correspondingly magnified actions (81). This effect is even more significant for patients taking multiple drugs because of competition for fewer binding sites.

The volume of distribution is also altered in older adults due to a reduction in total body water and lean body mass, with a relative increase in body fat (82). As a result, water-soluble drugs (e.g., digoxin, cimetidine) tend to have a smaller volume of distribution, with higher plasma concentrations and greater pharmacological effect (83). Conversely, fat-soluble drugs (e.g., diazepam, phenobarbital) usually have a larger volume of distribution because of relatively greater storage in fatty tissue. This may result in delayed therapeutic effects, with the potential for unexpected late toxicity. By the same token, prolonged drug effects are seen after dosage change or discontinuation because of the amount of drug stored in adipose tissue (81).

**Gastrointestinal System**

The term “presbyesophagus” has been used to describe multiple changes in the esophageal function commonly observed with aging, such as delayed esophageal emptying, incomplete sphincter relaxation, and decreased amplitude of peristaltic contractions. Only the latter appears to be a direct result of aging, but it is not clinically significant; the other changes are related to associated disease processes and may have significant clinical ramifications (84). Most importantly, there is an increased risk of aspiration with aging due to less coordinated swallowing.

Age-related changes occur throughout the gastrointestinal (GI) system although the more distal portion is most affected (85). Alterations in colon function include slightly decreased force and coordination of smooth muscle contraction resulting in slower transit time, as well as impaired rectal perception of feces (86). The high incidence of constipation in older people accordingly is thought to be related to multiple additional factors, such as low dietary fiber and fluid intake, sedentary habits, and various associated diseases interfering with intrinsic bowel function (e.g., parkinsonism, stroke) (87). A variety of medications are potentially constipating as well, including minerals (e.g., aluminum antacids, iron, calcium), opiates, nonsteroidal anti-inflammatory drugs (NSAIDs), antihypertensives (e.g., calcium channel blockers, clonidine), anticholinergics (e.g., tricyclic antidepressants, neuroleptics, antispasmodics), and sympathomimetics (e.g., pseudoephedrine, isoproterenol, terbutaline) (88). Prolonged use of stimulant laxatives or enemas can also impair bowel contractility and result in constipation or obstruction (87). Older adults often report straining and hard bowel movements along with their constipation (89). Straining may indicate rectal dyschezia (i.e., impaired rectal sensation and contractility).

Fecal incontinence in older people is due most commonly to overflow incontinence secondary to fecal impaction but can also occur as a result of decreased sphincter tone, cognitive impairment (e.g., from drugs, dementia), diarrhea, or dyschezia (87,89). Diarrhea among elderly patients is most frequently caused by fecal impaction, intestinal infection, or drugs (e.g., broad-spectrum antibiotics, digoxin toxicity) but also can be due to chronic laxative abuse (90). More appropriate interventions for bowel regulation include increasing diet fiber, using bulk agents or stool softeners, and avoiding frequent use of enemas or laxatives.

Despite these physiologic changes with aging, little effect is seen on absorption of most orally administered drugs (83). Drug absorption in general is more significantly affected by concomitant administration of multiple drugs; in particular, antacids and laxatives bind to or reduce dissolution of other medications (81).

**Hepatic System**

The primary changes in the hepatic system with aging involve a gradual progressive decline in liver size (5% to 15%) and hepatic blood flow, as well as slowing of hepatic biotransformation, specifically and most consistently microsomal oxidation and hydrolysis (83,86). This can have major implications on the circulating concentration of certain drugs and their metabolites, depending on the mode of metabolism and clearance. Drugs with high first-pass clearance (e.g., propranolol, propoxyphene, major tranquilizers, tricylic antidepressants, antiarrhythmic drugs) are cleared less effectively owing to reduced hepatic blood flow, resulting in greater bioavailability (81). Comorbid processes such as congestive heart failure can exacerbate these effects.

Drugs metabolized by means of phase I biotransformation (i.e., oxidation, reduction, hydrolysis) tend to have prolonged elimination in older people (e.g., diazepam, chlordiazepoxide, prazepam), whereas those undergoing phase II metabolism (i.e., glucuronidation, acetylation, sulfation) generally are not affected by aging changes (e.g., oxazepam, lorazepam, triazolam) (81,88).

It is important to note that studies of drug elimination with aging demonstrate significant interindividual variability that is likely due to genetic variation as well as the effects of such factors as smoking, alcohol, caffeine intake, diet, and concurrent use of other medications (83). As a result, caution should be exercised when using age-based guidelines for dosage determination (61).

**Renal System**

There are a number of age-related anatomic and physiologic changes in the kidney, including decreases in renal mass, number and functioning of glomeruli and tubules, renal blood flow, and glomerular filtration rate (91–93). These reductions in renal function have major implications for drug excretion, with prolonged half-lives for those drugs cleared primarily by glomerular filtration (e.g., cimetidine, aminoglycosides, digoxin, lithium, procainamide, penicillin, chlorpropamide) (81).

Studies show a mean age-related decrease in renal function of about 1% per year, with a decrease in creatinine clearance of 7.5 to 10 mL per decade; however, there is wide variability, with up to a third of older individuals showing no significant decline (92). Because of a corresponding decline in daily urinary creatinine excretion (reflecting decreases in muscle mass), there is no significant change in serum creatinine level with
aging. As a result, neither serum blood urea nitrogen (BUN) (which is dependent on dietary intake and metabolic function) nor creatinine is valid for accurately gauging renal function in older people (81).

Other common physiologic changes with aging include impaired ability to concentrate and dilute urine, impaired sodium conservation, reduction of urine acidification, and decreased ability to excrete an acid load (91). This erosion of reserve capacity allows maintenance of fluid and electrolyte homeostasis under normal conditions, but not with sudden changes in volume, acid load, or electrolyte balance. As a result, older people are more vulnerable to hyponatremia, hyperkalemia, dehydration, and perhaps most seriously, water intoxication (56, 94).

Because of difficulty in concentrating urine in conjunction with a blunted thirst mechanism, a hypernatriemic state with attendant mental confusion can result if an elderly person is stressed by higher than usual insensible losses (e.g., high or prolonged fever, heat exposure, exercise) with poor fluid intake (94). This is pertinent in a rehabilitation setting because patients are often engaged in vigorous activities and may become dehydrated relatively easily.

Just as older patients are prone to volume depletion when deprived of salt, acute volume expansion from an elevated sodium load caused by inappropriate intravenous fluids, dietary indiscretion, or intravenous radiographic contrast dye can result in congestive heart failure, even in elderly patients without preexisting myocardial disease (81, 95). A further potential complication of the use of radiocontrast materials in the elderly is the risk of acute renal failure, which is exacerbated by the presence of preprocedure dehydration (91). Because renin and aldosterone plasma concentrations are decreased by 30% to 50% in the elderly, with increased susceptibility to hyperkalemia, potassium-sparing diuretics (e.g., spironolactone, triamterene) should be used with great caution (88).

Hyponatremia due to water intoxication may be the most serious electrolyte disorder of older adults (56, 94). Most frequently complicating an acute illness, the clinical picture includes nonspecific signs of depression, confusion, lethargy, anorexia, and weakness. Serum sodium concentrations below 110 mEq/L may result in seizures and stupor. The syndrome of inappropriate antidiuretic hormone secretion (SIADH), with water retention and hyponatremia, can occur with infections (e.g., pneumonia, meningitis), strokes, various drugs (e.g., diuretics), or the stress of anesthesia and surgery (92).

**Pulmonary System**

Although progressive declines in pulmonary function are observed with aging, in the absence of significant pulmonary, cardiovascular, or neuromuscular disease, these declines are reflected primarily as a loss of reserve capacity without major functional limitations at rest (96). However, impaired pulmonary function on spirometric testing does indicate increased risk for subsequent disability and several common causes of death in older people, including cardiovascular disease and chronic obstructive pulmonary disease (COPD) (97, 98).

Changes in pulmonary function observed with aging reflect effects of aging per se (in the pulmonary as well as cardiovascular and neuromuscular systems) together with the cumulative effects of inhaled noxious agents (especially cigarette smoke and air pollutants) and infectious processes (96). The latter typically have a far greater impact on pulmonary function.

Progressive decline in a number of pulmonary function tests has been documented with aging, including vital capacity, maximum voluntary ventilation, expiratory flow rate, and forced expiratory ventilation (96). These declines reflect aging changes in the pulmonary system combined with those in related organ systems, which are collectively stressed by the maximum volitional inspiration and expiration required to complete the tests. Examples include stiffening of the rib cage from degenerative calcification of costochondral cartilage (i.e., decreased compliance), weakening of intercostal and abdominal muscles, and increased airflow resistance from small airway narrowing due to decreased elasticity (97). Residual volume and functional residual capacity increase, related to the loss of elastic recoil (increased compliance), although total lung capacity remains unchanged.

Normal gas exchange requires both uniform ventilation of alveoli and adequate blood flow through the pulmonary capillary bed. With increasing age, there is a progressive ventilation-perfusion imbalance due to collapse of small peripheral airways with decreased ventilation of alveoli, resulting in a linear decline in $pO_2$ with aging ($pO_2 = 110 - (0.4 \times \text{age})$) (96). Due to altered thoracic mechanics, $pO_2$ in older individuals is lower in the supine position than with sitting or standing. No changes occur in $pCO_2$ or pH, and oxygen saturation is typically normal or only slightly reduced.

This reduction in arterial oxygen tension is clinically relevant because it represents an additional loss of reserve. Elderly patients are more vulnerable to significant hypoxia from a relatively minor insult (e.g., anemia, congestive heart failure, respiratory infection) or the stress of physical inactivity because they are closer to the steep slope of the oxygen-hemoglobin dissociation curve (96). Blunting of central and peripheral chemoreceptor responsiveness exacerbates this vulnerability further: both hypercapneic and hypoxic ventilatory responses markedly diminish with aging, independent of lung mechanics. There is a significant increase in sleep-related breathing disorders with aging, and this appears to be related to this phenomenon (56).

Maximal oxygen consumption ($VO_{2\text{max}}$), an overall measure of exercise capacity and cardiopulmonary fitness, depends on pulmonary ventilation, cardiac output, peripheral circulatory control (i.e., ability to shunt blood to exercising muscles), and muscle oxidative capacity (i.e., oxygen extraction from the blood). Although a progressive decline in $VO_{2\text{max}}$ is observed with aging, this does not appear to be on a pulmonary basis (97, 99). In fact, it appears that decreases in $VO_{2\text{max}}$ in older adults with mild to moderate COPD are due primarily to cardiac and peripheral muscle deconditioning resulting from limited activity levels (97). Regular exercise to maintain or improve fitness is critical with aging because it is possible to
improve fitness with training at any age, and this is associated with a reduced vulnerability to stress or disease (and thereby increased active life expectancy) (65,69,100–102). The tendency of physicians and society to tolerate (or even encourage) decreased activity among older people, in conjunction with trends toward obesity and increased recumbency, probably contributes more to poor pulmonary function than aging alone (97,103).

Although most attention regarding the high incidence of pneumonia in the elderly is focused on immunologic declines, there appear to be contributing factors, direct or indirect, relating to the pulmonary system itself. Because many pneumonias result from aspiration of the infecting organism, impaired mucociliary function and decreased chest wall compliance with weaker cough (resulting in impaired ability to clear aspirated material or secretions) likely play a role (96,97). Other nonimmunologic contributing factors may include dysphagia, disruption of lower esophageal sphincter integrity, various esophageal disorders, and reduced levels of consciousness.

**Cardiovascular System**

A number of established tenets about the aging cardiovascular system have been revised, based on continuing research using more rigorous methodologies to exclude occult disease and controlling for degree of habitual physical activity. As a result, it now appears that cardiac output at rest and during graded exercise is relatively unaffected by age directly (56,99,104). Although resting heart rate does not change with aging, maximal heart rate with exercise does decrease progressively, related to decreased chronotropic responsiveness to adrenergic stimuli. The clinical formula reflecting this decline in maximal heart rate involves subtracting the age from 220 for men and (0.8 × age) from 190 for women (100,104). Decreased inotropic responsiveness to adrenergic stimulus results in decreased myocardial contractility, with decreased ejection fraction and increased risk of congestive heart failure (105). Maintenance of cardiac output at rest and with modest exercise is accomplished by early involvement of the Frank-Starling mechanism, with increased stroke volume via higher left ventricular end-diastolic volumes (56,104).

Another age-associated change is a decrease in the rate of early diastolic filling, with a much greater dependency on late filling through atrial contraction (104). As a result, older people are more vulnerable to deleterious effects of atrial tachycardia or fibrillation, including congestive heart failure (99,105).

Both cross-sectional and longitudinal studies demonstrate decreases in maximal oxygen consumption with aging, regardless of habitual activity levels (63,99,106). However, physically active people retain significantly greater maximal aerobic capacity with aging compared to their sedentary counterparts (65). In fact, trained elderly subjects may have greater maximal oxygen consumption than sedentary subjects who are much younger (66). Furthermore, endurance training, even when begun in old age, can significantly improve exercise capacity (65,100). Of clinical relevance is that the energy of walking represents an increasing percentage of the total aerobic capacity with advancing age, such that walking becomes a very effective physical conditioning activity (107).

A final age-related physiologic change in the cardiovascular system with important clinical applications is decreased baroreceptor sensitivity (107,108). This results in a diminished reflex tachycardia on rising from a recumbent position and accounts in part (possibly along with blunted plasma renin activity and reduced angiotensin II and vasopressin levels) for the increased incidence of symptomatic orthostatic hypotension in the elderly (Table 59–2), as well as cough and micturition syncope syndromes (104,109).

### TABLE 59.2 Office Management of Orthostatic Hypotension

<table>
<thead>
<tr>
<th><strong>Office Management of Orthostatic Hypotension</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>• Orthostatic hypotension is defined as a reduction in systolic blood pressure of at least 20 mm Hg or diastolic blood pressure of at least 10 mm Hg within 3 minutes of assuming an erect posture. However, the significance of any decrease in blood pressure upon standing should be evaluated in context with associated symptoms.</td>
</tr>
<tr>
<td>• Regardless of whether orthostatic hypotension is symptomatic or asymptomatic, the elderly patient remains at significant risk for future falls, fractures, transient ischemic attacks, and myocardial infarction.</td>
</tr>
<tr>
<td>• Orthostatic hypotension can be acute or chronic. Acute causes include hypotensive medications, dehydration, and adrenal insufficiency. Chronic causes can be further subdivided into those related to aging or age-related blood pressure elevation (physiologic causes) and those due to central or peripheral autonomic nervous system diseases (pathologic causes).</td>
</tr>
<tr>
<td>• The diagnostic evaluation of orthostatic hypotension should include a comprehensive history and physical examination, careful blood pressure measurements, and laboratory studies.</td>
</tr>
<tr>
<td>• Goals of treatment in the elderly patient include ameliorating symptoms, correcting any underlying cause, improving the patient’s functional status, and reducing the risk of complications, rather than trying to attain an arbitrary blood pressure goal.</td>
</tr>
<tr>
<td>• In the most cases, treatment of orthostatic hypotension begins with nonpharmacological interventions, including withdrawal of offending medications (when feasible), physical maneuvers, compression stockings, increased intake of salt and water, and regular exercise.</td>
</tr>
<tr>
<td>• If nonpharmacological measures fail to improve symptoms, pharmacologic agents should be initiated. Fludrocortisone, midodrine, nonsteroidal anti-inflammatory drugs, caffeine, and erythropoietin have all been used to treat orthostatic hypotension due to autonomic failure.</td>
</tr>
</tbody>
</table>

**Immunologic System**

Significant alterations in immunocompetence occur with aging, involving both cellular and humoral immune functions (110,111). Although the total number of lymphocytes decreases by about 15% in older adults, this does not appear to contribute significantly to the marked decline in immunocompetence (112). There is a decline in lymphocyte proliferation in response to antigen stimulation in older adults, as well as a higher incidence of energy (110). Age-related shifts have been observed in the regulatory activities of T cells (i.e., fewer T cells with suppressor or helper activity) and monocytes or macrophages.

Changes in humoral immunity with aging include increases in circulating autoantibodies and immune complexes, with decreased antibody production (110). The latter is characterized by an attenuated response to immunization, with difficulty maintaining specific serum antibody levels.

The increased susceptibility of the elderly to infection is a function of both these age-related changes in immune function and the frequency of concomitant factors that further impair host defenses (e.g., diabetes, malignancy, vascular disease, malnutrition, and stress) (110). Altered local barriers to infection, such as skin breakdown or an indwelling urinary catheter, often compromise resistance to infection further. Common infectious processes in the elderly include influenza, pneumonia, urinary tract infection, sepsis, herpes zoster, and postoperative wound infections.

Of particular clinical relevance is the fact that older people react differently to infections than do their younger counterparts. There is a less active leukocytosis in response to inflammation, and the total white blood cell count often is not increased (although usually there is still a shift of the differential count to the left) (113). The older patient may have less pain or other symptomatology, and frequently absent, or only low-grade, fever.

**Endocrine System**

The endocrine system also undergoes significant changes as we grow older. There is a gradual decrease in glucose tolerance with aging, although the fasting blood sugar level remains relatively unchanged (56). Accordingly, age-adjusted criteria for diabetes mellitus have been developed. This age-related decline in glucose tolerance is due to reduced sensitivity of tissues to the metabolic effects of insulin or insulin resistance (114,115). Compounding these aging changes are secondary conditions that further reduce tissue sensitivity to insulin, including lifestyle changes (e.g., obesity, diet changes, stress, sedentary lifestyle), other diseases (e.g., chronic infections, prolonged immobilization), and effects of medications (45,116). Older adults with diabetes are also at increased risk for common geriatric syndromes (e.g., depression, falls) (117).

Of clinical importance is the risk for untreated hyperglycemia, osmotic diuresis, and dehydration, potentially leading to hyperosmolar nonketotic coma or ketoacidosis (116). Certain drugs can cause or potentiate hyperglycemia (e.g., thiazide diuretics, glucocorticoids, tricyclic antidepressants, phenothiazines, phenytoin) (118). Control of serum glucose in older diabetics with oral sulfonylureas or insulin can be fragile, with significant risk for hypoglycemia. However, even borderline hyperglycemia appears to result in accelerated atherosclerosis and multiple end-organ involvement (56). On the other hand, recent data indicate that extremely tight glycemic control in type 2 diabetics may also be harmful (119). Of interest are the contributions of obesity and physical inactivity to the increased incidence of diabetes in older adults and the benefits of weight loss and regular exercise in improving control (45).

There are multiple other endocrine changes associated with aging. The primary clinical impact of altered thyroid physiology with aging is the need to maintain a high index of suspicion for the unusual presentation of thyroid disease (120). Presenting signs and symptoms of the older thyrotoxic patient may include palpitations, congestive heart failure, angina, atrial fibrillation, major weight loss associated with anorexia, and either diarrhea or constipation (121). Goiter and serious ophthalmopathy frequently are absent. Apathetic hyperthyroidism may not be recognized until late in the course of illness: patients appear depressed and withdrawn, with clinical clues of muscle weakness, dramatic weight loss, and cardiac dysfunction (116). Signs and symptoms of hypothyroidism essentially are unchanged with aging, but the diagnosis still may be delayed because of the many similarities between the stereotype of senescence and the hypothyroid state (e.g., psychomotor retardation, depression, constipation, cold intolerance). In view of the higher incidence of hypothyroidism in older adults, routine periodic screening of thyroid function is warranted (121).

The relationships among the hypothalamus, pituitary, and adrenal cortex remain unchanged with age, with preserved diurnal rhythm and stress response (116). However, in older women, serum cortisol concentrations vary more over the course of a day, and mean daily cortisol and ACTH (adrenocorticotropic hormone)-stimulated serum cortisol levels are higher (122,123). Primary adrenocortical disease is uncommon in the elderly. Significant hyponatremia or hyperkalemia, suggestive of adrenocortical insufficiency, is not uncommon in the elderly but more often is secondary to drugs (e.g., thiazide diuretics, chlorpropamide, carbamazepine) (88).

Age-related changes in gonadal function are well documented. There are variable and gradual declines in serum testosterone levels in healthy men with aging, likely due to partial testicular failure; however, there is no indication for routine androgen replacement (56,124). Postmenopausal declines in estrogen levels are well documented, with clinical expression variably including vasomotor instability syndrome (i.e., hot flashes), atrophic vaginitis, and osteoporosis (43,125). Controversy continues over prophylaxis and treatment of the latter, particularly with regard to potential benefits of dietary supplements and exercise (43,45,126). The reader is referred to Chapter 39 for further details.

**Thermoregulatory System**

Older people have only a mildly impaired temperature regulation system due to a combination of diminished sensitivity to temperature change and abnormal autonomic vasomotor...
As a result, they have a reduced ability to maintain body temperature with changes in environmental temperature and are vulnerable to both hypothermia and hyperthermia. The risk of hypothermia is compounded further by impaired thermogenesis (i.e., inefficient shivering), with potential aggravation by a variety of conditions (e.g., hypothyroidism, hypoglycemia, malnutrition) or medications (e.g., ethanol, barbiturates, phenothiazines, benzodiazepines, narcotics).

Conversely, diminished sweating (due to higher body temperature to initiate sweating and decreased sweat production) is a major contributing factor in heat exhaustion and heat stroke in hot conditions. Hypohidrosis is aggravated by anticholinergics, phenothiazines, and antidepressants.

Two thirds of deaths from heat stroke occur in people over 60 years of age, reflecting this impairment in regulatory systems. This has major implications for rehabilitation exercise programs, particularly when combined with a tendency for dehydration.

Sensory System
Deterioration of vision is one of the most recognized sensory changes occurring with aging. The most common visual change with increasing age is a gradual loss of the ability to increase thickness and curvature of the lens to focus on near objects (i.e., presbyopia) and physiologic miosis. Cataract formation, with opacification of the lens, occurs to some degree in 95% of the 65-or-older population. The elderly are also at significantly higher risk for further disease-related visual decrements (e.g., glaucoma, macular degeneration, diabetic retinopathy). The result of these various changes is a loss of visual acuity, decrease in lateral fields of vision, decline in both dark adaptation ability and speed of adaptation, and higher minimal threshold for light perception. These changes have obvious implications in relation to the higher incidence of falls in the elderly, particularly at night.

Gradual decline in hearing acuity (i.e., presbycusis) also is characteristic of aging, although again a number of treatable disorders can cause superimposed hearing loss (e.g., wax occluding the outer canal, cholesteatomas, acoustic neuromas). Older people most commonly manifest a conductive hearing loss, possibly due to increased stiffness of the basilar membrane or distortion of perceived sound with increase in threshold sensitivity, narrower range of audibility, abnormal loudness, and difficulty discriminating complex sounds. Continuing advances in hearing aid technology make remediation of such hearing deficits increasingly feasible. Early recognition and treatment of hearing impairments are particularly critical in the presence of cognitive deficits to avoid adverse sequelae of social isolation and development of paranoid ideations or frank psychiatric reactions.

Neurologic System
Numerous changes in the functioning of the neurologic system have been noted with aging. Three important areas of dysfunction accompanying normal aging include declines in short-term memory, loss of speed of motor activities (with slowing in the rate of central information processing), and changes in posture, proprioception, and gait.

The major controversy over neurologic changes with aging concerns cognitive functioning. A significant proportion of the observed decline in fluid intelligence with aging appears to be related to a decrease in the rate of central information processing. There is progressive deterioration in performance after age 20 on timed motor or cognitive tasks, including abstraction tests (e.g., digit symbol substitution test), reaction time tasks, and other tests requiring speed in processing of new information. Although there are declines with aging in motor and sensory nerve conduction velocities and rate of muscle contraction, they account for only a fraction of these slowed responses.
Many aspects of learning and memory remain relatively intact during normal aging, including immediate or primary memory as measured by digit span recall, retrieval from long-term storage, storage and retrieval of overlearned material, and semantic memory (137). However, age-related impairments have been documented consistently in tasks involving episodic short-term memory and incidental learning (138). Examples include difficulties with free recall of long (i.e., supraspan) lists of digits or words and paired-associate and serial recall learning, for both visually and verbally presented material. What these investigations indicate is that older adults are capable of new learning, but at a slower rate (137).

Because much of rehabilitation involves learning, these findings have major implications for rehabilitation programming for elderly people with disabilities. This is particularly true in the context of superimposed cognitive deficits, given that intellectual ability is an important determinant of the effectiveness of a standard geriatric rehabilitation program (4).

A final area of neurological age-related physiologic changes involves posture, proprioception, and gait (139). Older people in general are noted to demonstrate progressive declines in coordination and balance, related in part to impaired proprioception (140). This may have significant implications for degree of mobility and stability, although there are a number of common, potentially concomitant, pathologic changes that may contribute further to gait problems in the elderly (e.g., vertebral compression fractures with kyphosis, arthritis, degenerative cerebral changes, cerebral infarcts) (130,141).

### Musculoskeletal System

There is a well-documented progressive loss of muscle strength with aging, on the order of 14% to 16% per decade (men and women) for lower-extremity muscles and 2% (women) to 12% (men) per decade for upper-extremity muscles (59). A major contributing factor to this observed decline in strength appears to be an overall decrease in muscle cross-sectional area and mass with age (142). However, there may be significant contributions of cellular, neural, or metabolic factors to changes in strength, as loss of strength was observed even without loss of muscle mass (59). Further, significant gains in muscle strength, as well as functional mobility, have been demonstrated in older individuals with a structured, high-intensity resistance exercise program, even in frail nursing home residents up to 96 years of age (143).

The high prevalence of both osteoporosis and degenerative joint disease (i.e., osteoarthritis) in the elderly again raises the question about normal physiologic changes versus ubiquitous pathologic processes (45,144). The physiologic changes and sequelae associated with osteoporosis are discussed further in Chapter 39.

Distinction of the “disease” of osteoarthritis from the normal or usual aging changes that occur in weight-bearing joints can be made on a biochemical basis: with osteoarthritis, there are increases in the water content of cartilage and the ratio of chondroitin-4-sulfate to chondroitin-6-sulfate, with decreases in keratin sulfate and hyaluronic acid content (the opposite of what occurs in aging) (144). There is a strong relationship between aging and osteoarthritis: Degenerative joint changes in weight-bearing joints are essentially a universal occurrence in both genders by 60 years of age (145). These changes include biochemical alteration of cartilage, especially the proteoglycan component, with reduced ability to bear weight without fissuring; focal fibrillation and ulceration of cartilage, and eventual exposure of the subchondral bone (144). The tear hypothesis of osteoarthritis suggests that this process is the result of the cumulative stresses of a lifetime of joint use. Accordingly, “primary” osteoarthritis results from the stress of repetitive weight loading (e.g., spine, knees) or strain (e.g., distal interphalangeal joints), whereas “secondary” osteoarthritis may be related to occupational factors or congenital factors with unusual patterns of stress (e.g., congenital hip dysplasia). There appear to be other factors operating, however, because there are specific differences in distribution and prevalence between genders and races, and other explanatory models have been proposed (145–147). Obesity appears to be a risk factor for knee osteoarthritis in particular, although it is not clear whether this is due to a mechanical or a metabolic etiology. Further details regarding arthritis can be found in Chapter 31.

### Genitourinary System

Benign prostatic hyperplasia is an almost universal occurrence in men older than 60 years of age and develops under hormonal rather than neoplastic influence (148). Of note is that the median lobe of the prostate, which is not palpable rectally, can cause a ball-valve obstruction during micturition. Accordingly, after ruling out other etiologies (e.g., anticholinergic or diuretic medication side effects), cystoscopy should be considered in patients with persisting obstructive symptomatology but minimal prostatic tissue on rectal examination to detect median lobe hypertrophy (149). Usual indications for surgical intervention (e.g., prostatectomy) include increasing obstructive symptoms, recurrent/persistent gross hematuria, bladder calculi, recurrent infections, and postvoid residual volumes greater than 100 mL (148,150).

Incontinence in the elderly, although increasingly prevalent with advancing age, should be regarded as a symptom of underlying disease; it does not result from the natural aging process (151). Normal aging typically results in decreases in bladder capacity, ability to postpone voiding, detrusor contractility, and urinary flow rate (152). Postvoid residual volumes are typically increased, with a tendency for increased urine output later in the day, as well as propensity for uninhibited detrusor contractions. Each of these changes predisposes older adults to incontinence, but none alone precipitates it. Common medical conditions associated with incontinence are listed in Table 59–3, while precipitating (and reversible) causes of transient incontinence are detailed in Table 59–4.

The primary clinical significance of these aging changes is that the new onset, or exacerbation, of incontinence in an older person is likely due to a precipitating factor outside the urinary tract (151). Usually, remedial intervention can restore continence.
### TABLE 59.3 Common Medical Conditions Associated with Incontinence in the Elderly

<table>
<thead>
<tr>
<th>Condition</th>
<th>Effect on Continence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neurologic disease</td>
<td></td>
</tr>
<tr>
<td>Cerebrovascular disease; stroke</td>
<td>DO from damage to upper motor neurons; impaired sensation to void from interruption of</td>
</tr>
<tr>
<td></td>
<td>subcortical pathways; impaired function and cognition</td>
</tr>
<tr>
<td>Delirium</td>
<td>Impaired function and cognition</td>
</tr>
<tr>
<td>Dementia</td>
<td>DO from damage to upper motor neurons; impaired function and cognition</td>
</tr>
<tr>
<td>Multiple sclerosis</td>
<td>DO, areflexia, or sphincter dyssynergia (dependent on level of synergy)</td>
</tr>
<tr>
<td>Multisystem atrophy</td>
<td>Detrusor and sphincter areflexia from damage to spinal intermediolateral tracts</td>
</tr>
<tr>
<td>Normal-pressure hydrocephalus</td>
<td>DO from compression of frontal inhibitory centers; impaired function and cognition</td>
</tr>
<tr>
<td>Parkinson’s disease</td>
<td>DO from loss of inhibitory centers; impaired function and cognition; retention and</td>
</tr>
<tr>
<td></td>
<td>overflow from constipation</td>
</tr>
<tr>
<td>Spinal cord injury</td>
<td>DO, areflexia, or sphincter dyssynergia (dependent on level of injury)</td>
</tr>
<tr>
<td>Spinal stenosis</td>
<td>DO from damage to detrusor upper motor neurons (cervical stenosis); DO or areflexia</td>
</tr>
<tr>
<td></td>
<td>(lumbar stenosis)</td>
</tr>
<tr>
<td>Metabolic disease</td>
<td></td>
</tr>
<tr>
<td>Diabetes mellitus</td>
<td>Detrusor underactivity due to neuropathy, DO, osmotic diuresis; altered mental status</td>
</tr>
<tr>
<td></td>
<td>from hyper- or hypoglycemia; retention and overflow from constipation</td>
</tr>
<tr>
<td>Hypercalcemia</td>
<td>Diuresis; altered mental status</td>
</tr>
<tr>
<td>Vitamin B&lt;sub&gt;12&lt;/sub&gt; deficiency</td>
<td>Impaired bladder sensation and detrusor underactivity from peripheral neuropathy</td>
</tr>
<tr>
<td>Infectious disease</td>
<td></td>
</tr>
<tr>
<td>Herpes zoster</td>
<td>Urinary retention if sacral dermatomes involved; outlet obstruction from viral</td>
</tr>
<tr>
<td></td>
<td>prostatitis in men; retention and overflow UI from constipation</td>
</tr>
<tr>
<td>Human immunodeficiency virus</td>
<td>DO, areflexia, or sphincter dyssynergia</td>
</tr>
<tr>
<td>Neurosyphilis</td>
<td>DO, areflexia, or sphincter dyssynergia</td>
</tr>
<tr>
<td>Tuberculosis</td>
<td>Inanition and functional impairments (sterile pyuria found in ≤50% of genitourinary</td>
</tr>
<tr>
<td></td>
<td>TB cases)</td>
</tr>
<tr>
<td>Psychiatric disease</td>
<td></td>
</tr>
<tr>
<td>Affective and anxiety disorders</td>
<td>Decreased motivation</td>
</tr>
<tr>
<td>Alcoholism</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Functional and cognitive impairments; rapid diuresis and retention in acute</td>
</tr>
<tr>
<td></td>
<td>intoxication</td>
</tr>
<tr>
<td>Psychosis</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Functional and cognitive impairments; decreased motivation</td>
</tr>
<tr>
<td>Cardiovascular disease</td>
<td></td>
</tr>
<tr>
<td>Arteriovascular disease</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Detrusor underactivity or areflexia from ischemic myopathy or neuropathy</td>
</tr>
<tr>
<td>Congestive heart failure</td>
<td>Nocturnal diuresis</td>
</tr>
<tr>
<td>Other organ system diseases</td>
<td></td>
</tr>
<tr>
<td>GI disease</td>
<td>Retention and overflow UI from constipation</td>
</tr>
<tr>
<td>Musculoskeletal disease</td>
<td>Mobility impairment; DO from cervical myelopathy in rheumatoid arthritis and</td>
</tr>
<tr>
<td></td>
<td>osteoarthritis</td>
</tr>
<tr>
<td>Peripheral venous insufficiency</td>
<td>Nocturnal diuresis</td>
</tr>
<tr>
<td>Pulmonary disease</td>
<td>Exacerbation of stress UI by chronic cough</td>
</tr>
</tbody>
</table>

DO, detrusor overactivity; TB, tuberculosis; UI, urinary incontinence.


Contrary to stereotypes, although there is a decrease in sexual functioning with aging, most older people retain sexual interest and desire and to a variable extent, capability (153–155). Older men experience a decrease in ability to have psychogenic erections and require more intense physical stimulation for erection; erections may be partial, and orgasm with ejaculation may occur without full engagement (156). The force of ejaculation is less, along with a less intense sensation of orgasm. Impotence may be caused by a variety of diseases (e.g., atherosclerosis, diabetes, hypothyroidism) and medications (e.g., antihypertensives, phenytoin, cimetidine). Treatment of erectile dysfunction in older men has been revolutionized with development of the vacuum tumescent device, advances in penile prostheses, and availability of such medications as sildenafil and alprostadil (157,158).

Older women experience postmenopausal changes, including increased fragility of the vaginal wall and attenuation of the excitement phase (e.g., decreased vaginal lubrication) (153). Common sexual difficulties identified included partner’s impotence, anorgasmia, decreased libido, and insufficient opportunities for sexual encounters. Despite these changes, most women maintain the ability to engage in sexual intercourse throughout the life cycle (159).
IMPACT OF THE ENVIRONMENT ON THE HEALTH OF OLDER ADULTS

Psychological and Social Issues in Aging

Ageism and Myths of Aging

Butler coined the term “ageism” (or agism) to describe negatively biased perceptions of older people by the younger population in today’s youth-oriented culture, as well as perceptions of old age by elderly individuals themselves (160). There are many adverse sequelae of ageism, including devaluation of older people (by themselves, as well as others both younger and older), diversion of health care professional focus from the real health problems of older patients, the dearth of physicians interested and trained in geriatric medicine, and lack of curriculum time in medical schools regarding geriatrics (161–163). There is even new evidence of a negative impact of perceived discrimination on mortality in older adults (164). According to Rowe, it is time to “discard the many derogatory myths about older people, who are often seen as sick, senile, silly, sexless, and sedentary, as well as inflexible, irritable, noncontributing, and too old for preventive interventions” (165). Many of today’s older adults are survivors of the Depression and the World Wars—they built much of the life and standard of living we now enjoy. The evidence is clear: the majority of elderly are cognitively intact, live independently in the community, and are fully independent in ADL (17,19).

Cumulative Changes

There is increasing awareness of the critical interrelationships, particularly for older people, of physical health, mental health, and life circumstances. The emotional and life stress associated with major losses are well documented, and older people may be exposed progressively to multiple significant losses: job, income, health, functional ability and independence, parents, spouse, siblings, children, friends, social roles and status, and self-esteem (166). There are in fact few norms or defined role expectations regarding appropriate behavior or activities in old age (167,168). Bereavement, isolation, poverty, illness, and physical disability all are associated with a higher incidence of depression in older adults (169), which in turn is associated with decreased physical and cognitive functioning, disability, and increased mortality (34,170–172).

Social Support Networks

Social support networks include a wide variety of sources that can be categorized as informal (family), semiformal (church, clubs, family doctor, local pharmacist), and formal (health care system, social service agencies, insurance companies, etc.) (173,174). Older persons often use supports from a combination of these networks (13). There is mounting evidence of the positive impact of social support networks on the cognitive, health, and functional statuses of older individuals (174–176).

Elderly people with children usually live near them and visit frequently or at least maintain regular telephone contact (177). Elderly people without children tend to maintain closer ties with young relatives or with siblings (178). It is important to consider the extended family, including cousins, in-laws, and others, with regard to support networks, rather than just immediate household members (167,174,179).

Institutionalization of an impaired older person usually is the last resort for families, used only when all other efforts fail; in fact, 64% of individuals over the age of 85 who are dependent in self-care or homemaking still live in the community (180). Families, rather than the formal system of government and agencies, provide the bulk (up to 90%) of personalized long-term care for their disabled older relatives (13). This includes home health and nursing care, personal care, household maintenance, transportation, cooking, and shopping. In 1990, 73% of elderly disabled individuals relied exclusively on such informal support and care networks (13).

---

**TABLE 59.4 Precipitating Causes of Transient Incontinence in the Elderly**

<table>
<thead>
<tr>
<th>Cause</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delirium, confusion state</td>
<td>UI resolves once underlying cause(s) treated</td>
</tr>
<tr>
<td>Urinary infection</td>
<td>UI may be only symptom of infection; antibiotic trial warranted in asymptomatic persons only on initial evaluation and with new onset/exacerbation of UI</td>
</tr>
<tr>
<td>Atrophic urethritis, vaginitis</td>
<td>Aggravates stress or urge UI; agitation can be presenting symptom in demented patients</td>
</tr>
<tr>
<td>Medications</td>
<td>Any agent that impairs cognition, mobility, fluid balance, bladder contractility, or sphincter function; many agents impair several functions</td>
</tr>
<tr>
<td>Psychiatric disorders</td>
<td>Severe depression or psychosis</td>
</tr>
<tr>
<td>Increased urine output</td>
<td>Frequency or nocturia; causes: excessive fluid intake, diuretics, hyperglycemia, hypercalcemia, volume overload (congestive heart failure, venous insufficiency, hypothyroidism, hypoalbuminemia, and drug-induced peripheral edema)</td>
</tr>
<tr>
<td>Restricted mobility</td>
<td>Treat underlying cause; provide a urinal or bedside commode</td>
</tr>
<tr>
<td>Stool impaction</td>
<td>Urge or overflow UI; fecal incontinence common</td>
</tr>
</tbody>
</table>

UI, urinary incontinence.
With advancing age, however, older adults tend to have increasingly limited and relatively fragile support systems. Dependency in aging parents can result in significant physical, emotional, and financial stresses on their family network (177). An alternative support system may evolve gradually over a period of time as the older person loses family support (e.g., death of spouse and siblings, children moving away and unable to actively assist). Such a system might include friends and neighbors in an extended network to assist with shopping, cooking, cleaning, and self-care (174,181).

With whatever combination of support systems, a significant additional insult (e.g., onset of a new disease or complication) may overtax an already marginal arrangement. It is commonly observed that as the patient’s dependence on the formal network of the health care system increases, the informal or family network support decreases (176). Furthermore, if the elderly person is hospitalized for a prolonged period, the network(s) may dissipate and may be difficult or impossible to reassemble (182). The critical importance of maintaining the integrity of support networks is illustrated by the observation that for every aged impaired person in a nursing home there are two equally impaired older people living in the community (3). The difference is the role played by the latter’s informal support systems, providing most of their long-term care.

Increasingly, issues concerning family functioning with aging are being studied. Even when family members are seemingly available to assist older relatives, their support cannot always be counted on unless they too receive help. Fortunately, there is evidence that patients’ families can benefit from educational interventions to help prevent weakening in this crucial source of patient support. Caregivers of older patients with cancer and chronic pain are often frustrated, fearful, and anxious. Patient care improved when caregivers were provided with guidelines on what they could do within the home to help the patient (176). Similarly, caregivers of stroke survivors had less depression and better outcomes with more formal training (183). Nursing home placement has been delayed by specific family interventions for patients with Alzheimer’s disease (184).

Caregiver burden is another dynamic receiving increasing attention (185). Increased caregiver burden can be associated with increased mortality after controlling for known risk factors (186). Adding to the physical stress of providing personal care aid may be the unpleasantness of incontinence or exhaustion due to a relative’s sleep disorder. Behavioral problems, such as agitation or impulsivity with poor safety awareness, create proportionately greater caregiver burden than the demands of providing physical assistance (185). Physical and emotional health problems among caregivers have been documented, including depression and immunosuppression. Physical aggression on a caregiver by a patient is not uncommon and may lead to reciprocal abuse (187). Potential intervention strategies include encouraging use of other support systems to augment care provided by family members, as well as the use of respite programs (188,189).

**Functional Impact**

Physically impaired older people tend to become socially isolated, which can result in exacerbation of medical problems, functional deficits, and mental health problems (particularly depression) (34,171). Other factors contributing to a vicious cycle of depression, withdrawal, and functional decline may include the stress of multiple losses, malnutrition, chronic illness, pain, and adverse drug effects that aggravate depression (172). Unfortunately, the environment too often fosters dependency. A classic illustration is the acute hospital setting, where the focus is on routinely providing care and assistance, rather than encouraging self-care (190,191).

Additional psychosocial barriers can interfere with maintaining or improving functional ability in the elderly. Handicapping sequelae of ageism include devaluation of elderly disabled (by themselves as well as others), lack of interest (actual and/or perceived) among health care professionals in their problems, and limited opportunity for access to appropriate rehabilitation services (161,165). Further attitudinal obstacles encountered among disabled elderly include the “right of dependency,” perceived as earned by virtue of longevity, and the “apathy of fatigue,” both physical and emotional, associated with multiple illnesses and hospitalizations (192). Increasingly, physical environments are being recognized as either preventing or contributing to disability (193). While full discussion is beyond the scope of this chapter, the development of a requirement for some housing to meet visitability criteria (one entrance without stairs, bathroom on first floor, and wide bathroom doorway to accommodate a wheelchair) is encouraging. Housing adapted in this manner may more easily accommodate older adults as they age and facilitate social visits to or from friends by minimizing physical barriers.

The obvious conclusion, and why rehabilitation plays a key role in restoring function in disabled older people, is the importance of awareness and intervention regarding significant psychoemotional and social factors affecting their health. Many of these factors can be anticipated and prevented, or at least minimized in terms of their adverse effects. As with any complication, prevention is the best treatment.

**Cumulative Functional Sequelae of Disease**

Older adults can experience acute onset of disability just as in younger people from such conditions as stroke, amputation, spinal cord injury (SCI), and traumatic brain injury. However, many experience a gradual progression of difficulties in function. The effects of multiple and chronic illnesses usually are gradual over time with cumulative erosion of organ reserves, leaving the elderly person reasonably functional with various adaptations, such as walking more slowly or taking more frequent rests (27,35,45). As functional problems develop, it can be hard to determine if the disability can be treated generically (i.e., regardless of the contributing diseases and other factors) or it is due to a reversible underlying disease process that needs to be treated. Also, an elderly individual may be only marginally functional with little or no reserve capacity, so that even a relatively minor superimposed acute
complication or disease process (e.g., influenza) may result in functional decompensation (3,194). Of even greater concern is that this significant functional decompensation may be difficult to reverse even though the intercurrent acute illness is appropriately treated and resolves (194–197).

Underreporting of Illness
Older people as a group are more vulnerable to functional sequelae of diseases for a variety of reasons, including ageism. The latter commonly results in underreporting symptomatology related to illness (35,160). Health care providers may not be trained adequately to evaluate and treat symptoms and signs of functional disability in older patients and may as a result not recognize the significance of vague and inconsistent symptoms. Older people themselves may think that such vague symptoms are a natural result of aging (197). As a result, the underlying disease process may become quite advanced before care is sought, making treatment that much more difficult.

From the older person's viewpoint, the available system of care may seem unresponsive (35). Physician offices can be perceived as inconveniently located, with inadequate parking and limited access for the physically impaired. A typically brief physician encounter may not allow for development of rapport and full elaboration of symptoms. Busy office staff may appear to be uninterested or discourteous.

Other issues may contribute further to underreporting of illness. There may be denial of disease coupled with fear of consequences, especially financial (160). Depression is common among older people and may result in the attitude, “What have I got to gain?” (35,171). Increasing isolation, with fewer opportunities for others to observe and react to changes in appearance or behavior, is an additional barrier. Finally, older people may not recognize significant symptoms or seek medical attention because of cognitive impairments, which not infrequently may be secondary to or aggravated by an underlying and potentially reversible disease process (135).

Altered Response to Illness
There is often an altered response to illness in the elderly, which contributes to a delayed, or incorrect, initial diagnosis (35). Many specific diseases present with atypical signs and symptoms. For example, the presentation of a myocardial infarction in an elderly person is less likely to include classic retrosternal chest pain; more often it will involve nausea, dizziness, syncope, or congestive heart failure with decreased activity tolerance (78). Furthermore, a wide variety of diseases may present with similar nonspecific symptoms, including confusion, weakness, weight loss, and general “failure to thrive” (35,198). Accordingly, the differential diagnosis of possible disease processes is much broader in elderly patients.

Variable Patterns of Disease
Further confounding accurate elucidation of the underlying illness are the frequent changes in disease patterns and distribution (35). Abnormalities in one organ system may be accompanied by secondary abnormalities in other organ systems. Traditional medical training focuses on disease recognition and treatment in a relatively young population, with emphasis on synthesizing multiple signs and symptoms into a single unifying diagnosis (35,199). Older people more typically have concurrent symptomatology, relating to multiple diseases. Although accurate diagnosis is important, the functional impact of each disease, particularly the cumulative and additive impact of multiple diseases, must be determined (3).

Atypical Responses to Treatment
There is an increased frequency of many chronic diseases in this population, including anemia, osteoarthritis, osteoporosis, cardiovascular disease, malignancy, and malnutrition. Table 59-5 lists types and patterns of disabilities associated with various chronic conditions. Palliation and prevention of secondary complications frequently are more appropriate and realistic goals than is cure of the primary condition (35,199). There often are atypical (and potentially confusing) behaviors and responses to treatment, however, due to coexisting diseases and decreased functional reserves of multiple organ systems (e.g., affecting drug metabolism and distribution) (35).

Older people also are more prone to a wide variety of comitant and complicating diseases, which may further cloud diagnosis and treatment decisions. Examples include thrombophlebitis, dehydration, fluid and electrolyte disturbances, adverse drug interactions or toxicity, pressure ulcers, pneumonia, and general deleterious effects of deconditioning due to inactivity, which occurs earlier and with greater severity in older adults (200,201).

Frailty as a concept refers to more than just older individuals who experience functional loss; it represents a state of vulnerability resulting from the balance and interplay of medical and social factors (38,39,202).Characteristics of frail older institutionalized persons included female gender, being unmarried, absence of a caregiver, presence of cognitive deficit, functional impairment, and medical condition (e.g., diabetes mellitus, stroke, Parkinson's disease) (203).

“Frailty” is becoming better characterized as a biologic syndrome of older adults associated with decreased reserve and resistance to stressors (38,39,204). Contributing are the multiple decrements in physiologic systems, as previously discussed. Clinical characteristics have been defined as unintentional weight loss (10 lb in prior year), self-reported exhaustion, weakness (decreased grip strength), slow walking speed, and decreased physical activity. Using these criteria, if three or more of these characteristics are present, the individual may be classified as frail (38,204).

Models of Disease Presentation
Given these multiple dynamics, alternative models of disease presentation in older persons may prove useful (199). In one series of proposed models (Table 59–6), the most basic is the Medical Model, in which the patient’s symptoms and signs are fully explained by one disease. The Synergistic Morbidity Model explains the functional loss that may have suddenly occurred as the result of the additive effects of several diseases. In the
**Attribution Model**, a patient might attribute a deterioration in health to an already diagnosed chronic illness when in fact another undiagnosed problem is present (e.g., hypothyroidism superimposed on stroke). In the more complex **Causal Chain Model**, a medical-psychiatric interaction occurs in which one illness causes another in addition to causing functional decrements, with the presenting complaint representing the proverbial last straw of decompensation. Finally, in the **Unmasking Event Model**, a stressful external event unmasks an underlying, stable, or slowly progressive chronic condition.

<table>
<thead>
<tr>
<th>Characteristics Entered into Each Model</th>
<th>Association with Disability in:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A Mobility/Exercise Tolerance Demanding Tasks</td>
</tr>
<tr>
<td>Age</td>
<td>•</td>
</tr>
<tr>
<td>Sex (male)</td>
<td></td>
</tr>
<tr>
<td>Angina (nitroglycerine use)</td>
<td>•</td>
</tr>
<tr>
<td>Myocardial infarction</td>
<td>•</td>
</tr>
<tr>
<td>Congestive heart failure</td>
<td>•</td>
</tr>
<tr>
<td>Stroke</td>
<td>•</td>
</tr>
<tr>
<td>Claudication</td>
<td>•</td>
</tr>
<tr>
<td>Arthritis</td>
<td>•</td>
</tr>
<tr>
<td>Lung disease</td>
<td>•</td>
</tr>
<tr>
<td>Depressive symptomatology</td>
<td>•</td>
</tr>
<tr>
<td>Hearing impairment</td>
<td>•</td>
</tr>
<tr>
<td>Cognitive impairment: digit symbol substitution</td>
<td>•</td>
</tr>
<tr>
<td>Cancer</td>
<td>•</td>
</tr>
<tr>
<td>Weakness</td>
<td>•</td>
</tr>
<tr>
<td>Balance problems, last year</td>
<td>•</td>
</tr>
<tr>
<td>Dizziness, last 2 wk</td>
<td>•</td>
</tr>
<tr>
<td>Body mass indexa</td>
<td>•</td>
</tr>
<tr>
<td>Weightb</td>
<td>•</td>
</tr>
</tbody>
</table>

• Indicates significant association (P < 0.01), adjusting for: clinical sizes, mini-mental state exam score, hypertension, visual impairment, diabetes, left ventricular systolic dysfunction by echocardiography, carotid stenosis, and grip strength as well as variables in the table.

*a Increase of 10 lb.


**TABLE 59.6 Alternative Models of Illness Presentation**

<table>
<thead>
<tr>
<th>Model of Illness</th>
<th>Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Medical Model</td>
<td>Symptoms and signs correlate directly (and solely) with a specific disease (traditional model of medical diagnostic thinking).</td>
</tr>
<tr>
<td>2. Synergistic Morbidity Model</td>
<td>Multiple concurrent chronic diseases, subclinical (or at least not severe enough to seek medical attention) until reaching a threshold of cumulative morbidity and functional decline.</td>
</tr>
<tr>
<td>3. Attribution Model</td>
<td>Underlying stable chronic disease(s), with new onset of symptoms attributed by the patient (and possibly, the clinician) to the known chronic disease(s), but in actuality caused by a new, unrecognized condition.</td>
</tr>
<tr>
<td>4. Causal Chain Model</td>
<td>Medical-psychiatric interaction involving an underlying disease with multiple secondary conditions/impairments (potentially including mental health problems, such as depression), which aggravate the underlying disease, forming a “vicious cycle.”</td>
</tr>
<tr>
<td>5. Unmasking Event Model</td>
<td>Underlying stable or slowly progressive chronic disease(s), effectively compensated to minimize functional impact, until the occurrence of a new stressful event (e.g., death of a loved one, relocation), which results in decompensation (and the risk of being misinterpreted as a new condition).</td>
</tr>
</tbody>
</table>

that had previously been well compensated and unrecognized (199). Alternative classification models such as these help organize the significant and varying complexity of signs and symptoms, to more rapidly and accurately diagnose and treat older persons with multiple interacting problems.

**Effects of Acute Hospitalization**

There is increasing recognition of the multiple deleterious effects of acute hospitalization on older people, separate and distinct from sequelae due primarily to their presenting illness (195,205). Disorientation due to the foreign hospital environment and relatively infrequent and brief interactions with unfamiliar health care personnel may contribute to bizarre and inappropriate behavior, including delirium and agitation (171,190,206). Contributing to this may be relative sensory and social isolation with few familiar environmental cues or social interactions, especially if the patient is confined to a private room or intensive care setting. Moreover, there are atypical routines and schedules (e.g., blood drawing, vital sign checks at odd hours), which, coupled with unusual noises (e.g., overhead paging, machines, other patients), may contribute to insomnia. A patient with insomnia typically is treated with a sedative medication, which may begin a cycle (or cascade) of other side effects and interactions that may adversely affect the patient’s health (195,205).

Increased incidence of medical and iatrogenic complications in older adults is well documented (35,199). Drug side effects, complications, and toxicity, together with adverse interactions related to polypharmacy, make up a large proportion of such morbidity (35,81). There also is a greater frequency of diagnostic and therapeutic misadventures in this age group, related in part to decreased organ reserve with resultant increased vulnerability (205).

There also are a variety of emotional sequelae of hospitalization that may affect health and functional statuses. Anxiety and confusion relating to the underlying illness and prognosis, or just to hospitalization itself, may interfere with cooperation with medical treatment or therapy programs (207–209). Depression from similar origins may result in dependency and poor motivation to cooperate or improve function (170,171). Functional dependency frequently is reinforced during acute hospitalizations, both by the older patients who expect hospital staff to assist and by hospital staff who tend routinely to perform self-care tasks without taking the extra time to supervise the patient in performing his or her own self-care (192,195). Documented functional decline after hospitalization for acute medical illnesses also may result from other, as yet unexplained, factors (205). Deconditioning from inadequate activity during the hospital stay may contribute to poorer functional outcomes, and a recent study indicates that older adults are more sensitive to the deleterious effects of bed rest as compared to younger individuals (210). Preliminary results are mixed as to whether exercise during hospitalization can help improve functional outcomes (211,212).

In addition to significant implications for health care in the hospital setting, these sequelae related to acute hospitalization often affect social support systems and discharge disposition (182,195). The elderly patient may experience loss of confidence or motivation as a result of multiple insults and complications, coupled with erosion of functional abilities from deconditioning (200,201). This by itself will put greater stress on often relatively fragile social support systems, making it more difficult for elderly patients to return home to their prior living situation (182).

To try to address these problems, several randomized trials have investigated alternative models attempting to change the organization of care (and outcomes) for older persons during hospitalization. Use of a geriatric consultation team approach did not yield improved outcomes (213). In another randomized clinical trial of 651 patients 70 years of age and older, the experimental patients received care on a special unit, which had the additional components of daily team conferences, active discharge planning, and use of therapy staff for functional training—the core components of an acute rehabilitation unit. Patients in the experimental group were discharged at higher functional levels, and fewer were discharged to skilled nursing facilities (14% vs. 22%). Neither cost of hospitalization nor length of stay was increased (214). In a Veterans Administration study, older patients, after stabilization, were randomized to usual care or to a geriatric rehabilitation unit (215). At hospital discharge, patients receiving the geriatric unit care had greater improvements in scores for four of the eight SF-36 subscales, ADL, and physical performance. These studies suggest potential strategies to improve acute hospital care for older persons, with less functional sequelae. Unfortunately, similar programs have not been routinely implemented despite their demonstrated benefit (216).

**Effects of Deconditioning**

Deconditioning can be defined as the multiple changes in physiology and anatomy induced by physical inactivity and reversed through physical activity (201). This topic is covered in detail in Chapter 48, but will be discussed here relative to unique aspects relating to aging.

Older adults’ (75 to 120 years old) physical functioning and associated physical activities have been classified by Spirduso into five categories: physically elite, physically fit, physically independent, physically frail, and physically dependent (217) (Fig. 59–4). Deconditioning can be one component, in addition to disease, that may contribute to a patient’s lower level of function. In addition, low fitness is an important predictor of cardiovascular and all-cause mortality in older adults (218).

As noted before, older adults have a more substantial loss of muscle mass than younger adults with bed rest inactivity (210). However, the associated declines in lower-extremity strength and maximal aerobic capacity in these older subjects were similar to that reported in younger adults exposed to similar periods of bed rest. Reversal of the deleterious effects of deconditioning has been amply demonstrated through focused muscle-strengthening programs as well as comprehensive exercise programs in both nursing homes and the community.
These programs variably included exercises for flexibility, muscle strength, and aerobic endurance.

The functional consequences of deconditioning in older people may be of major clinical significance and may be confused with changes intrinsic to aging or changes from diseases (201). In addition to deconditioning, sarcopenia is another process that can affect overall function in older individuals (222). Deconditioning per se may result in functional losses when certain threshold values for physical performance are crossed (223). Quadriceps weakness may progress to the point of dependency in getting in and out of a car solely from progressive deconditioning, not related to intrinsic aging or new onset of disease (200). Multiple factors associated with falls may originate from deconditioning, or be exacerbated by deconditioning (224,225). For people living in the community, factors associated with falls include impairments in static balance, leg strength, and hip and ankle flexibility (226,227). In nursing home patients, falls are associated with decreased muscle strength at the knees and ankles (228,229). Weakened muscles also may contribute to other injuries and pain syndromes by allowing abnormal forces to act on bone, joints, ligaments, and tendons. In addition, lack of exercise is being viewed increasingly as a risk factor not only for functional loss but also for onset of various disease processes, including cardiovascular disease and diabetes, among others (230).

Deconditioning affecting older people can be differentiated into acute inactivity secondary to bed rest (such as during acute illness) and chronic inactivity from sedentary lifestyles (often more difficult to reverse) (201). A variety of types and combinations of exercises are available to treat deconditioning in older individuals; a precise prescription of a therapeutic exercise and activity program, including appropriate precautions and instruction, is essential (65,107,231,232).

Psychological issues in maintaining exercise habits are being studied increasingly (219). Currently, health professionals can help older persons increase their physical activity levels by discussing the issue openly and then helping the patient decide what approach may be most appropriate (103). One option may be group classes, which offer the benefits of social interaction and support, and perhaps even friendly competition. Exercising on their own at home may be preferable for some, especially if they are self-conscious about their own abilities (233).

Disability Prevention in Older Adults

The future impact of increased disability as the population ages is of major concern to clinicians, health care administrators, insurers, and policy makers. Increasingly, prevention strategies are being proposed both to improve the number of disability-free years of life as well as to contain health care costs (16,19,28,31,34,36,38). Some of the themes around prevention can be organized according to the concepts of primary, secondary, and tertiary preventions. Primary prevention involves preventing the onset of a disease (e.g., annual influenza vaccine), whereas secondary prevention involves the diagnosis and treatment of asymptomatic diseases to prevent the development of symptoms (e.g., treatment of hypertension to prevent stroke or myocardial infarction). Tertiary prevention involves treatment once a disease becomes symptomatic to avoid complications (e.g., deep venous thrombosis prophylaxis and appropriate mobilization to prevent skin breakdown in poststroke patients). However, as previously discussed, what becomes increasingly important in older persons with chronic disease (comorbidity) is focused attention to preventing or minimizing frailty and, therefore, decreasing risk of development of disability (38).

Another model that may help in the development of prevention strategies has been proposed by Lawrence and Jette (234). They have studied the application of a model for the disablenent process in 1,048 community-residing adults (mean age 74 years) without functional limitations or disabilities over

![TABLE: Fitness Levels](image-url)

<table>
<thead>
<tr>
<th>Fitness Level</th>
<th>Physically Elite</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>*High risk &amp; power sports</td>
</tr>
<tr>
<td></td>
<td>*Senior Olympics</td>
</tr>
<tr>
<td>Physically Fit</td>
<td>*Moderate physical work</td>
</tr>
<tr>
<td></td>
<td>*Endurance sports</td>
</tr>
<tr>
<td>Physically Independent</td>
<td>*Very light physical work</td>
</tr>
<tr>
<td></td>
<td>*Low physical demand activities (e.g., golf, driving)</td>
</tr>
<tr>
<td></td>
<td>*All IADLs</td>
</tr>
<tr>
<td>Physically Frail</td>
<td>*Light housekeeping</td>
</tr>
<tr>
<td></td>
<td>*Some IADLs</td>
</tr>
<tr>
<td></td>
<td>*All ADLs</td>
</tr>
<tr>
<td></td>
<td>*May be homebound</td>
</tr>
<tr>
<td>Physically Dependent</td>
<td>*No or only some ADLs</td>
</tr>
<tr>
<td></td>
<td>*Needs home-based or institutional care</td>
</tr>
</tbody>
</table>

![FIGURE 59-4](image-url)
a 6-year period. The model hypothesized a process in which risk factors (age, gender, education, body mass, and physical activity measured as frequency of walking one mile) would lead to functional limitations with or without the presence of pathology or impairments. Over time, the functional limitations (e.g., inability to walk well) would lead to subsequent disability such as inability to go shopping. Guralnik et al. likewise found that among nondisabled persons living in the community, impairments in the lower extremity were highly predictive of later disability (235). Fried et al. characterized a functional level of “preclinical disability” (task modification but no difficulty by self-report) in a study of women 70 to 80 years old (236). Measurement of their physical function was intermediate between women with high function (no modifications) and disability (difficulty with tasks). The potential role of interventions involving increased physical activity and exercise as types of primary prevention of disability for older persons is supported by studies documenting the dynamic nature of functional loss and impact on both morbidity and mortality (37,38,40).

There may in fact be significant overlap of risk factors for multiple problems in older persons. For instance, the risk factors associated with falls, incontinence, and functional dependence are similar—slowed chair stand, decreased arm strength, decreased vision and hearing, and either a high anxiety or depression score (237).

From these types of models, successful prevention strategies are being tested, often entailing multiple interventions covering multiple domains. For instance, incidence of falls can be reduced by a combination of medication adjustments, exercise, safety training, and environmental modifications. A comprehensive nurse-practitioner evaluation program for community-residing seniors resulted in a delay in onset of disability and decreased nursing home admission (238). Another study demonstrated that higher self-efficacy was associated with lesser functional decline in persons with diminished physical capacity (239). Research also suggests that self-efficacy (i.e., a person’s confidence or belief that he or she can achieve a specific behavior or cognitive state) may be modifiable and therefore help guide preventive strategies (240). Research continues to identify which targeted intervention for which specific risk factor in which specific patient at what specific point in time will be most efficacious in preventing or minimizing disability (38).

PRINCIPLES OF ASSESSMENT AND REHABILITATIVE MANAGEMENT OF OLDER ADULTS WITH DISABILITY

Assessment of the Older Patient
Rehabilitation medicine and geriatric medicine share common principals and have complementary approaches (241). This is seen in patient assessments. Rehabilitation of older adults requires attention to multiple areas. Benefits of organized, comprehensive assessments and care have been documented for acute onset disabilities like stroke and are described elsewhere in this textbook. Assessments of older patients with more insidious onset of multiple disabilities are evolving. They are collectively referred to as comprehensive geriatric assessment (CGA). Pioneering work by Dr. Marjorie Warren in Britain laid the foundation for CGA while evaluating many institutionalized older adults (242). The process was formally defined in the United States in a National Institute Consensus Development Conference in 1988 (243). CGA is “a multidisciplinary evaluation in which the multiple problems of older persons are uncovered, described, and explained, if possible, and in which the resources and strengths of the person are catalogued, need for services assessed, and a coordinated care plan developed to focus interventions of the person’s problems” (243).

Since that time, numerous studies have evaluated CGA programs in different sites of care. While the majority focuses on the general medical geriatric population, a growing literature examines benefits in specific areas like oncology. CGA programs vary depending on the clinical setting and program design. A general conclusion from this work is that there can be benefits depending on patients targeted and the nature of the program. Targeting aims to identify older adults who are not terminal and who have multiple conditions. These patients also have associated disabilities amenable to known multifactorial interventions. Usually, beneficial programs include actual control of care provided rather than consultation alone.

Benefits for some, but not all, inpatient programs have included shorter lengths of stay, better management of specific syndromes, delays in functional decline, decreased bodily pain at 1 year, decreased need for home health care services, and higher likelihood of living at home in 6 months (242,244,245). In some recent studies, mortality at 1 year has not been changed. Some, but not all, outpatient CGA programs have shown delays in functional decline, improved mental health at 1 year, and earlier diagnosis of common health problems like cognitive impairment, depression and anxiety, and incontinence (245–248).

CGA topics are similar to what needs to be assessed during comprehensive rehabilitation. Issues include sensory, cognitive, emotional, functional, social, financial, legal, and home issues. Other commonalities are screening and evaluation of syndromes like incontinence and falls. These syndromes are common entities that result from multiple diseases and multiple risk factors (237,249).

A standardized framework for assessing older patients, and managing care, can facilitate both rehabilitation and follow-up care by other providers. With so many types of issues requiring assessment, and follow-up, a practical standard framework would help avoid oversight of relevant factors. Communication among team members would be easier. This is especially important because older adults receive care from multiple providers in multiple settings. One estimate suggests Medicare patients receive annually, on average, care from five specialists and two primary care providers in four different practice sites (250). Another estimate for stroke survivors is that they can receive care from up to 89 providers distributed among eight different health care teams in seven different locations (251).

The Siebens Domain Management Model (SDMM) is one practical clinical framework for organizing care for older
The Siebens Domain Management Model

<table>
<thead>
<tr>
<th>Clinical Problems</th>
<th>Time/Setting 1</th>
<th>Time/Setting 2</th>
<th>Time/Setting 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. Medical/Surgical Issues</td>
<td>A. Symptoms</td>
<td>B. Diseases</td>
<td>C. Prevention</td>
</tr>
<tr>
<td>II. Mental Status/Emotions/Coping</td>
<td>A. Cognition</td>
<td>B. Emotions</td>
<td>C. Coping/Behavioral Symptoms</td>
</tr>
<tr>
<td>III. Physical Function</td>
<td>A. Basic Activities of Daily Living (BADLs)</td>
<td>B. Intermediate Activities of Daily Living (IADLs)</td>
<td>C. Advanced Activities of Daily Living (AADLs)</td>
</tr>
<tr>
<td></td>
<td>Home mobility</td>
<td>Community mobility</td>
<td>Avocational</td>
</tr>
<tr>
<td></td>
<td>Self-care</td>
<td>Other</td>
<td>Vocational</td>
</tr>
<tr>
<td>IV. Living Environment</td>
<td>A. Physical</td>
<td>B. Social</td>
<td>C. Financial/Community Resources</td>
</tr>
</tbody>
</table>

**FIGURE 59-5.** The two organizing constructs of the Siebens Domain Management Model are the domain classification of patients’ health-related issues (Vertical, or Y axis) and time, including identification of process factors (assessments and plans), outcomes, and care setting (Horizontal, or X axis). Copyright Hilary C. Siebens MD 2001, 2010. Used with permission.
resources. The patients’ physical environment includes man-made structures as well as the natural world (257). Specific details include the type of housing and job site, building access, and proximity of basic necessities like food and banking services. For older persons with progressive disabilities, transitioning to more supportive home environments may be a critical goal in their care.

Patients’ social environment is their interface with other people at many levels of association, ranging from social networks to social institutions (257). These interactions determine whether or not personal physical help or assistance with instrumental or advanced ADL is available. Caregiving by family and friends can be a rewarding experience. On the other hand, sometimes caregivers face significant burdens. Assessing caregivers’ experiences, and assisting through education and supports when necessary, is an integral part of rehabilitation management (258).

Finally, financial and community resources are critical to determining what medical and rehabilitation services are available as well as what personal financial resources patients and their families may have. This can facilitate determination of what medications, services, and additional equipment are feasible as part of the rehabilitation program.

Practical applications of the SDMM include structuring of physicians’ decision making in their documentation of their assessments/plans (252). Other applications include using it as a structure for team and family conferences, for teaching rehabilitation in residency programs, for medical student teaching about the care of older adults, and for efficient geriatric assessments in sites like emergency departments (259–263).

Rehabilitation Goal Setting
Increasingly, there is recognition that health care and rehabilitation goals may not be uniformly shared by patients and their caregivers (264,265). For example, patients may want to return home after rehabilitation, while rehabilitation staff may strongly disagree, at times, out of undue fear of what may occur (266). Our focus as health care providers is typically on general health and well-being, as well as functional independence. Our tacit assumption is that these goals are necessarily shared by our patients and their caregivers and perhaps that these are the only goals they value. However, there is evidence of differing goals, and priorities of goals, between health care providers, case managers, patients, and their caregivers (264,265,267). Alternative goals also valued by patients and their caregivers included such areas as education and referrals, social/family relations, emotional issues, and caregiver burden. Of significance is that patients and family caregivers often focus on process goals and shorter time frames, as opposed to typical health care providers’ emphasis on specific outcomes and longer range time frames (265).

Prescription of Rehabilitation Programs
An appropriate therapeutic prescription is critical to the success of a rehabilitation program. It must be based on a careful analysis of the patient’s current functional limitations, with realistic goal setting in the context of premorbid functioning and anticipated improvement in medical status (268). Specific therapy techniques will be used based on physical status (e.g., neurologic, musculoskeletal, cardiovascular systems) and medical stability. Social and cultural barriers to certain exercises or activities also must be taken into account. The patient should participate in goal setting, such that the rehabilitation plan is relevant to his or her own goals. Without a strong therapeutic alliance between all health professional team members and the patient (and family), progress will be slow and/or limited.

Functional training approaches usually are well accepted by most older individuals because they can clearly see the relevance and importance. Many therapeutic goals can be achieved by incorporating formal therapy techniques into the context of functional tasks. Examples of this approach include the following techniques:

- Remediate perceptual problems during eating or meal preparation
- Increase range of motion with dressing training
- Strengthen through inclined sanding or woodworking projects
- Aerobic conditioning through adapted competitive sports.

Prescription of the rehabilitation program must be tailored to the person to accommodate limitations imposed by comorbid medical problems (65,100,107). Cardiovascular (e.g., blood pressure/pulse response, cardiac symptomatology) and pulmonary (e.g., use of oxygen) restrictions should be established as appropriate. Weight-bearing limits can be accommodated by means of assistive devices or aquatic therapy. In patients with limited exercise tolerance, it may be necessary to use flexibility with therapy scheduling and duration of treatment, with frequent rest periods. A number of authors have suggested practical rehabilitation guidelines that summarize these recommendations (107,269–271). Resnick proposes a seven-step model to help motivate older individuals to exercise, including education, goal setting, role modeling, and verbal encouragement/rewards (272).

Realistic goal setting is complex in any rehabilitation setting but may be complicated further in older patients in two unique respects. First, older adults frequently have potential caregivers (e.g., spouse, siblings, children) who also are aging and have medical problems of their own. Thus, the fitness and capability of the proposed caregiver after discharge must be considered in the discharge planning process (178,179,182). The second problem relates to the limited remaining life expectancy of the older disabled patient. For example, diabetic amputees over 65 years of age have average survival of approximately a year (273). In view of their potentially limited longevity, an expedited rehabilitation program to facilitate return home with family would be most appropriate. This would also apply to people with conditions that are progressively disabling and require increasingly frequent episodes of hospitalization and/or skilled nursing care.
A number of potentially negative or counterproductive team dynamics can develop in rehabilitation settings that may interfere with or limit a patient’s progress (274,275). Patients and families tend to trust their health care providers as the experts who will know and do what is best for them. It is critical that the rehabilitation team maintain vigilance for and strive to counter such negative attitudes as paternalism (e.g., overriding patient goals judged by the team to be unrealistic or inappropriate), arrogance (e.g., presumptive familiarity by addressing patients by their first name without requesting permission or preference), and self-fulfilling prophecies (e.g., patients judged not to have potential for improvement not receiving as much attention or effort as patients felt to have a good prognosis). Other team issues that can impact effectiveness include relative team member roles (e.g., lack of clarity or excess rigidity), communication barriers, or decision-making conflicts (274). The ongoing challenge for rehabilitation teams is to foster an individual and collective philosophy of respect for and empowerment of the individuals they treat, facilitating functional independence. The latter includes (appropriate) risk taking (referred to as “dignity of risk”).

Significance of Functional Status in Placement Issues
Reference already has been made to the critical nature of functional status with regard to the ability to live independently in the community. Older people often live alone and must perform their own self-care and other daily activities, including homemaking. Issues of safety in this home environment are raised, particularly after an acute adverse event or illness (e.g., a fall with hip fracture). A patient who achieves a level of mobility (i.e., ambulation or transfers) requiring only close supervision or contact guard assistance because of an occasional loss of balance may not be able to safely return home alone. Home-based supervision by an aide, unless paid privately, typically is available only a few hours a day, 5 days a week, for relatively limited intervals. In this instance, return to a community setting may only be possible if a relative is available to live with this individual. If not, he or she may not even qualify for a boarding home or intermediate care/assisted living facility; most intermediate care facility admission criteria include independent safe locomotion. This patient accordingly may require a skilled nursing facility (SNF, or nursing home) admission as the only source of 24-hour/day supervision available. Unfortunately, alternatives are limited for more homelike settings with other residents of similar functional level.

Health Care Policy Issues
There is increasing evidence that health care financing can have a significant (if unintended) impact on the structure, delivery, and outcomes of health care. Chan et al. have analyzed the effect of changes in the Medicare payment system for rehabilitation hospitals on length of stay and charges, as well as discharge destinations (276–278). Their analyses reveal that lengths of stay and charges dropped significantly the year after implementation of the new Medicare prospective payment system for rehabilitation, compared to the base year, for all diagnostic categories (276). A further dynamic involved an increased likelihood of discharge from a rehabilitation hospital to an SNF, particularly for (often older) patients with longer lengths of stay, under the new prospective payment system with its fixed reimbursement cap (277).

Chan et al. have also documented associations between type of insurance (Medicaid, health maintenance organizations (HMOs), commercial fee-for-service plans) and relative frequency of postacute hospital discharge settings (SNF vs. rehabilitation hospital) for patients with traumatic brain injury (279). Discharge from acute care hospitals to SNFs was significantly more likely for patients with Medicaid, and to a lesser extent with HMO coverage, than for patients with fee-for-service commercial insurance.

On the other hand, there is also evidence that programs specifically designed to improve quality of care for older patients can effect changes in health care delivery and outcomes. Coleman, on behalf of the HMO Care Management Workgroup, describes an evidence-based approach to improving quality of geriatric care within managed care organizations (280). Their premise is that by expanding the managed care organization focus to include older members at risk for functional decline and subsequent frail health (with higher health care costs), “upstream” approaches to maintain or improve functional reserve will result not only in improved health but also lower health care costs. There is clear evidence to support the premise of the correlation between activity limitation and increased health care costs (281).

Jencks et al. cite improvement in care rendered to Medicare patients over a 2-year interval, based on tracking of (and possibly related to feedback for) a number of quality of care indicators by the Quality Improvement Organization (QIO) program of the Centers for Medicare and Medicaid Services (CMS) (282). The quality indicators included inpatient management of acute myocardial infarction, heart failure, stroke, and pneumonia, as well as ambulatory preventive measures for breast cancer, diabetes, and pneumonia. Despite the historical shortcomings of quality improvement efforts on a national level as documented in Healthy People 2010 (283), Jencks is optimistic that improvement in health care quality is not only possible but is occurring. He notes this trend should be reinforced by alignment of quality indicator tracking between such national agencies as CMS, the Joint Commission, and the National Quality Forum.

Along similar lines, Calkins et al. propose changes in components of the health care system to improve the health of older people, such as more detailed emergency room evaluations and more focused (and sustained) home health care services (284). What is clear in any case is that health care policy continues to be in flux, trying to address the challenges of financing and delivering appropriate care to older persons (278).
REHABILITATION APPROACHES TO COMORBIDITY AND GERIATRIC SYNDROMES

Medical Issues
Reassessment of Medical Status
Transfer of a patient from an acute medical or surgical service to a rehabilitation unit has been identified as an opportunity for a fresh and objective reassessment of medical status (285). Such an appraisal is even more critical in geriatric rehabilitation and should include confirming the accuracy of referral diagnoses, evaluating for previously unrecognized conditions, and reviewing medications for continuing appropriateness (199). Often an elderly patient’s physiologic status has markedly improved or stabilized by the time of transfer, warranting consideration of altering dosage or discontinuing certain medications. Creating and using tools like a patient care notebook with medication lists is one process that can help clarify medication use (286). Patients can be encouraged to take these notebooks to all health care appointments.

Avoiding Adverse Drug Effects
The incidence of adverse drug reactions approaches 25% in people older than 80 years of age (35). Adverse drug reactions account for 11% to 20% of hospitalizations in people older than 65 years of age and are a frequent cause of mental deterioration (287–289). The reasons for drug problems in older adults revolve around five key interrelated areas:

1. Polypharmacy
2. Medications not taken as prescribed
3. Increased susceptibility to adverse reactions
4. Altered pharmacokinetics
5. Altered receptor sensitivity

Polypharmacy
Polypharmacy is frequent in older people and often compounded by the frequent use of nonprescription drugs, resulting in preventable adverse drug reactions and unnecessary financial costs (290,291). Physicians all too often contribute to this problem; one study found that one in five older patients received a prescription for a potentially inappropriate medication (292). Frequently, known risk factors are not taken into account when prescribing a particular medication and dosage (293,294). Medication histories obtained by physicians often are inaccurate in both inpatient and outpatient settings (295,296), related in part to lack of questioning and underreporting of over-the-counter (OTC) medications, which are used almost as frequently as prescription medications by older adults (297,298). Over 50% of OTC medications are oral analgesics, with the remainder consisting of cough-and-cold preparations, vitamins, antacids, and laxatives (296,298).

The hospital setting is an excellent opportunity to discontinue drugs of questionable value because careful monitoring is possible. Accordingly, all medications should be reviewed carefully on admission to a rehabilitation unit and regularly thereafter. In addition to OTC medications, special attention should be directed to digitalis preparations, NSAIDs, and psychotropic medications (299).

In addition to unawareness of concurrent medications, other factors have been identified as potentially influencing prescribing habits and contributing to polypharmacy in older people. These include pharmaceutical advertising (particularly involving new drugs inadequately tested in older people) and patient and family expectations (or even demands) for treatment, usually in the form of a prescription (300). In one nursing home study, treatments (i.e., medication prescriptions) occurred more often if a patient’s problem produced ongoing discomfort for the staff (e.g., agitation); conditions tended to be undertreated if symptoms were intermittent or had less impact on staff (e.g., arthritis) (301).

Medications Not Taken as Prescribed
Patients do not take medications as prescribed in one third to one half of all cases (302). Patients over 75 years of age who live alone are especially likely to not take medications as prescribed (296,298). Reasons given by patients in one study for not taking medications included “feeling the prescribed dosage was too high” and experiencing problematic side effects (303). These patient choices can cause postdischarge deterioration; patient education prior to discharge is critical to avoid such problems. Allowing patients to self-medicate in the hospital with flexible administration times may be a useful way to monitor their understanding and an opportunity to reinforce the need for medications (304). This strategy also may help to decrease the frequency of incorrect drug frequency or dosage, omitted medications, and use of expired medications (305). However, challenges remain in determining the best methods for enhancing patient adherence to medication prescriptions (306).

Drug toxicity can result when a patient is admitted to the hospital and given all the drugs (in the “correct” dosages and frequency) that have been previously prescribed but were not being taken at home. Such a problem should be suspected when a patient shows a decline in cognitive or functional status 5 to 10 days after admission and other medical workup is unrevealing (307).

Increased Susceptibility to Adverse Reactions
Adverse drug reactions appear to be more common in older patients, even when medications are given in the proper dosages (308). This may be related to a relative lack of resiliency in their homeostatic mechanisms (35,299). Although not all adverse drug reactions are avoidable, some investigators suggest that 70% to 90% could be anticipated and prevented (291,301,302). That some patients choose not to adhere to prescriptions given by physicians may in fact help decrease the frequency of adverse drug reactions (303). Serious side effects can occur secondary to OTC medications, especially from antihistamines with anticholinergic side effects, leading to fatigue and confusion even in middle-aged adults (297).
Altered Pharmacokinetics

There are a number of age-related changes in pharmacokinetics (see Chapter 65) with significant implications for drug dosing, timing of dosage changes, and potential for unexpected toxicity or interactions. The reader is referred to the previous discussion of organ system aging for hematologic, GI, hepatic, and renal systems.

Altered Receptor Sensitivity

Age-related changes in receptor sensitivity to drug effects are an additional reason for untoward drug effects in older people. There is evidence, for example, that benzodiazepines and warfarin have a greater effect at similar concentrations in the young compared with the elderly (300,308). Such changes are difficult to evaluate separately from the pharmacokinetic changes related to aging.

The common philosophy of “go low, go slow” is sound advice when prescribing drugs for older people. Evaluation of response to drug therapy is critical, and elimination of unnecessary medication is essential for improving function. Patient and family education as to the indications, contraindications, and adverse effects of drugs are even more important in older individuals to try to improve adherence and avoid adverse reactions (304,305).

Hypotension

Symptomatic orthostatic hypotension can occur in many older patients after even relatively short periods of bed rest (201,309). Orthostatic hypotension is defined as a decline of 20 mm Hg or more in systolic blood pressure when rising from supine to standing, usually accompanied by symptoms of dizziness or light-headedness (310). It is accordingly a frequent problem during early remobilization of elderly patients in rehabilitation settings. Symptoms can persist if there are underlying problems with blood pressure maintenance related to drug therapy, salt restriction, or autonomic dysfunction (78).

Evaluation of the orthostatic patient should include a review of medications (particularly nitrates, antihypertensives, levodopa, diuretics, phenothiazines, and tricyclic antidepressants), examination for autonomic dysfunction (e.g., pupillary response, abnormal sweating) or recent fluid loss, and laboratory tests to rule out abnormalities in aldosterone and cortisol levels (199,310,311). Metanephrines should be evaluated when hypotension occurs with episodic hypertension, a hallmark of pheochromocytoma (311).

Treatment for symptomatic orthostasis includes discontinuing any prescribed or OTC medication that could be contributing to the hypotension. The patient should be instructed to exercise (i.e., ankle dorsiflexion/plantar flexion) before arising, to sit up initially, and then to stand up slowly while holding onto a support. Thigh-high elastic stockings or an abdominal binder may help minimize lower-extremity blood pooling (310) (see Table 59-2). High-sodium diet and fludrocortisone acetate, a synthetic mineralocorticoid, are useful for plasma expansion in the absence of congestive heart failure. Other medications to consider include NSAIDs (inhibit prostaglandin synthesis), clonidine or midodrine (α2-adrenergic agonists), propranolol (blocks β-2 vasodilatory receptors), pindolol (a β-adrenergic antagonist with intrinsic sympathomimetic activity), or phenylpropanolamine (a sympathomimetic) (299,300,310).

Significance of Mental and Emotional Status

Depression

Depressed mood is a significant problem in older persons and is often missed. Given the importance of improving recognition of depression, several recent reviews have covered the topic in detail (312–315). Rates of major depression vary from 16% to 30% for elderly clinical populations, and prevalence rates in community-dwelling older persons range from 2% to 5%. The risk of depression has been estimated to be threefold greater for older persons with disability, compared to their functionally independent counterparts (170,316). The reverse appears to also be true: depression in older adults is associated with significant reductions in active life expectancy (34). Essential is the distinction, sometimes hard to make, between depressed mood that will respond to supportive counseling and more severe depression requiring more aggressive intervention (e.g., psychotherapy, medication, electroconvulsive therapy) (317). Symptoms can include depressed mood, poor motivation, fatigue, and suicidal ideation. The rehabilitation team should maintain a high index of suspicion for the presence of depression that may require aggressive treatment. Vegetative signs suggestive of more severe depression may include the following:

- Sleep disturbance
- Loss of appetite
- Constipation
- Impaired concentration
- Poor memory
- Psychomotor retardation (171).

Other less specific complaints include other somatic symptoms such as pain and ill-characterized dyspnea (169). Depressed patients may appear as if they have a dementia syndrome (312,313).

In many patients with mild reactive depression, the activity and milieu of the rehabilitation unit will alleviate the depression. Progress in therapy and the support of peers and staff often are therapeutic. When the depression is more profound, antidepressants may prove helpful; however, medical contraindications may limit or preclude their use (312). Medications such as trazodone and the selective serotonin reuptake inhibitors (SSRIs) have lower anticholinergic activity and should be considered in addition to those tricyclic antidepressants with the fewest anticholinergic side effects (e.g., doxepin, nortriptyline, desipramine) (313). Doses should be started low, usually at night initially, and increased gradually.

Few studies have investigated the role of psychological and social interventions for older persons (312,314). Depressed outpatient volunteers in generally good health have shown benefits with cognitive behavior, interpersonal, and short-term psychodynamic therapies. Clinically, interventions like
these are potentially beneficial and require more evaluation, especially because medication alone cannot address such associated issues as altered life role (especially with disabilities), chronic medical illness, and losses of a spouse and/or close friends (171,314).

Anxiety
Anxiety syndromes are another frequent problem during rehabilitation (318). Symptoms can manifest in multiple systems, as in depression. Careful differential diagnosis is required to distinguish between primary anxiety disorders and those secondary to medical illness or medication. A detailed interview and chronology of onset are needed, and assistance from psychiatry often is needed. Diagnostic categories include adjustment disorder with anxious mood, generalized anxiety disorder, posttraumatic stress disorder, and panic attacks (319,320). When feasible, nonpharmacologic interventions, such as behavioral management techniques, physical therapy for muscle relaxation, or psychotherapy, should be used first (321). Nonetheless, judicious and appropriate medication use is frequently necessary to control anxiety symptoms and facilitate participation in the rehabilitation process. Commonly used options include lorazepam, buspirone, imipramine, and SSRIs (320,321).

A baseline anxiety disorder may be exacerbated by hospitalization, leading to agitation with nonpurposeful excessive motor activity. Depression, as well as preexisting psychoses, can also present with agitation (318). The former usually responds to the more sedating tricyclic antidepressants (e.g., doxepin). Paraphrenia is an example of a psychosis occurring in the elderly that often presents with agitation (321). This paranoid psychosis, with onset typically late in life, is characterized by bizarre paranoid delusions in a socially isolated person. Antipsychotics are critical to the management of this problem, coupled with therapeutic alliance with a physician and an attempt to redevelop social contacts for the patient. A history of preexisting psychiatric disorders should always be investigated in the agitated patient. Schizophrenia can continue into old age, although exacerbations respond well to antipsychotics (318).

Delirium
Delirium, a syndrome characterized by the acute onset of fluctuating cognitive deficits in conjunction with attention disorder and disorganized thought, can cause sleep disturbances, hallucinations, and agitation (322,323). It occurs more often in patients with prior cognitive impairments and can coexist with a dementia, making accurate diagnosis very difficult (324). Any acute medical illness can present in the older person with delirium without the classic signs of the underlying acute illness (199). Infections, dehydration, stroke, hypothermia, uremia, heart or liver failure, and pulmonary emboli are the most common examples of this phenomenon (325). Drug toxicity is another frequent cause of delirium in the elderly, with common offenders including neuroleptics and narcotics (324,325). There are evolving data on factors associated with postoperative delirium states (326). Delirium is still present in 16% of admissions to postacute facilities, presenting diagnostic and management challenges in the SNF setting (327). Delirium represents a medical emergency, with significant independent morbidity (328); identifying the cause is critical to its resolution (199,323).

Dementia
Severe dementia occurs in about 5% of individuals over 65 years of age (with mild to moderate forms in another 10%) and in about 20% of those over 80 years of age (329,330). It is found in more than half of nursing home residents and is the most common precipitating cause of admission (329). Women appear to be affected more frequently than men. Insidious onset of memory loss, loss of abstract reasoning and problem-solving ability, impairment of judgment and orientation, and personality changes with relatively intact alertness and awareness are all hallmarks of the disease (331). A patient with early dementia, premorbidly not interfering with daily activities, can become severely disoriented during an acute hospitalization (191,199,332). This agitated confusion may resolve without any specific therapy in 1 to 2 weeks. Appreciation of this possibility is important with regard to evaluating and working with this population in the rehabilitation setting.

Fifty to sixty percent of dementias represent dementia of the Alzheimer type, and another 20% are multi-infarct in origin (330). The remaining large number of potentially reversible causes of dementia include the following conditions:

Subdural hematoma
Brain tumor
Occult (normal pressure) hydrocephalus
Syphilis
Hypothyroidism or hyperthyroidism
Hypercalcemia
Vitamin B12 deficiency
Niacin deficiency
Drug toxicity
Depression
Cardiac, renal, or hepatic failure (331)

Diagnostic evaluation should always be performed to rule out these possible causes. Even if one of these potentially treatable etiologies is established, however, reversibility of the dementia may be limited because of permanent damage from the condition (333). Although there are differing guidelines for recommended laboratory tests for evaluating dementia (331), a standard dementia workup, in addition to detailed history and physical examination with cognitive screening, should include at least a complete blood count (CBC), blood chemistry profiles (including electrolytes, creatinine, BUN), erythrocyte sedimentation rate, and thyroid function studies.

Other investigative studies, such as serologic testing for syphilis, serum folate, serum cobalamin, drug screening, collagen vascular profile, urinalysis for heavy metals, or imaging of the brain (e.g., CT [computed tomography], MRI [magnetic resonance imaging]), can be undertaken if there is concern that
they will clarify the etiology, such as multi-infarct dementia versus Alzheimer’s disease (331,334). A trial off all medications, if possible, probably is warranted in all patients with new onset of dementia. Many clinicians also routinely give a trial of antidepressants to newly identified dementia patients because occult depression frequently coexists with mild dementia (198,335). Amelioration of the depression may improve overall functioning in this situation.

Patients with moderate or severe dementia can be limited in their ability to learn in the rehabilitation setting because their ability to form new memory is poor. Day-to-day carryover may be limited and makes certain types of therapeutic gains difficult to achieve (336). A rehabilitation trial may still be justified in such situations to clarify learning abilities and to train the family in appropriate care of a patient with a new disability. For instance, the patient may show the ability for procedural learning (learning by performing the activity) even if declarative learning (learning from verbal instruction) is impaired (184,207).

When evaluating the elderly patient for admission to a rehabilitation program, it is critical to determine the mental status before onset of the new disability by talking to family or others who have observed the patient. Too often the mental status as seen in the acute hospital setting underestimates the patient’s cognitive function when healthier and in a more supportive and stimulating environment (337).

Discharge planning for patients with dementia needs to include family education as to the nature of the patient’s cognitive strengths and weaknesses and how to handle potential behavioral problems (184). Community resources for adult day care and respite care programs may be very helpful for families, as well as educational materials like the Alzheimer’s Disease Education and Referral Center (ADEAR) website for family and healthcare professional information.

Pharmacological treatment of older patients with severe dementia, particularly with behavioral disturbances, is controversial and confusing. In 1995, the Food and Drug Administration (FDA) issued “black box” warnings for atypical (second generation) antipsychotic medications in elderly patients, due to multiple clinical trials documenting greater mortality risk (339,340). This left clinicians with unclear choices for treatment of elderly dementia patients with psychosis and/or agitation. However, other investigators conclude no increased risk from use of either atypical antipsychotics or conventional neuroleptics (341). Current treatment guidelines emphasize shared decision making with caregivers, thorough documentation, and frequent monitoring (339,340).

**Common Geriatric Syndromes**

**Incontinence**
An all too common complication, devastating to patient self-esteem and family commitment to patient care, is urinary incontinence. For diagnostic classification and evaluation procedures, the reader is referred to Chapter 51. Several recent reviews cover this topic thoroughly (342–345). Treatment for incontinence in the older patient hinges on proper diagnosis (See Tables 59-3, 59-4), which usually is possible with a complete history of the problem combined with careful neurologic, pelvic, rectal, and mental status examinations. Laboratory studies should include urinalysis, culture and sensitivity, serum creatinine and BUN, and a postvoid residual urine volume (342,346). A voiding diary often is helpful in determining the nature of the problem, and cystometrics/urodynamics may also be indicated (347).

Treatment is directed at the cause of the incontinence. Unfortunately, many of the etiologies have no uniformly successful therapy, and there may even be multiple causes. A timed voiding program is useful in many patients, offering toileting opportunities at regular intervals to try to maintain continence (342). Initially, the intervals are very short (e.g., every 15 to 20 minutes), with a progressive increase as indicated. Modifications of this technique include patterned urge-response toileting (PURt) (348) and functional incidental training (FIT) (349), with reports of excellent success in nursing home settings.

Surgical procedures may be useful in the treatment of prostatic hypertrophy and sphincteric incompetence (343,347). Anticholinergics (e.g., tolterodine, trospium) frequently are useful in the management of detrusor instability, but with the potential risk of retention (257). Other pharmacologic approaches include direct smooth muscle relaxants (e.g., oxybutynin) and imipramine (342,343). Overflow incontinence due to detrusor decompensation (from overstretching) may require long-term indwelling catheterization, although frequent intermittent catheterization and cholinergic drugs to stimulate detrusor contraction may be helpful (350). Excellent patient and health care professional educational materials are available (351,352).

Bowel incontinence may imply severe bilateral brain disease or loss of sensory input from the rectal ampulla (353). Biofeedback has been shown to be helpful in managing sensory bowel incontinence (354,355), but the management of incontinence secondary to diffuse brain disease usually requires a behavioral approach with bowel movements induced by suppositories at regular intervals (353).

**Sleep Disorders**
Sleep disorders and daytime fatigue are related problems common in the hospitalized elderly person as well as those living in the community (200,205,356). The hospital environment alone can disrupt the sleep cycle, an effect further compounded by foreign routines (vital sign checks and medication administration at odd hours), unfamiliar noises (from machines, overhead paging, and neighboring patients), and the depression often associated with the onset of new major chronic illness (205). Sleep deprivation at night leads to fatigue during the day. Napping during the day further disrupts nocturnal sleep patterns, and a vicious cycle can ensue (356).

It is important to document whether insufficient sleep is actually occurring because patients can complain of sleep difficulties when no problem is documented and they
remain alert throughout the day. Simple reassurance in such cases is warranted. In cases of documented sleep disorder, contributing factors such as delirium, medication toxicities, depression, anxiety, restless leg syndrome, chronic pain syndrome, or nighttime medical problems (e.g., congestive heart failure, angina) should be considered (199,205). It is also important to differentiate acute insomnia from chronic insomnia. Acute insomnia (present for <1 month) is often related to a stressor (e.g., bereavement) and is treated with support and short-term, intermittent medication. Chronic insomnia (persisting for >1 month) should be viewed more as a symptom of another illness (356). Hypnotics should be used judiciously and only if other interventions, such as improved sleep hygiene and treatment of the underlying illness, are unsuccessful.

After addressing these issues, good sleep hygiene practices may help. These include a regular sleep schedule, keeping the patient out of the bed and bedroom until bedtime, a snack before bedtime, daily exercise, relaxing activities in the evening before bedtime, and instruction in mental imagery or deep breathing relaxation techniques to be used as needed in bed at night (357). In addition, patients should not watch clocks during the night. Naps during the day should be avoided unless absolutely needed, briefly after lunch.

Only if these interventions fail should a sleep medication be considered. Selected antidepressants can be used in low doses at night to take advantage of sedative side effects, while minimizing anticholinergic activity (e.g., trazodone) (358). If a benzodiazepine-type hypnotic is used, the choice should be one with a very short half-life (e.g., zaleplon, zolpidem) to avoid accumulation with hangover effects (359,360). In general, diphenhydramine should be avoided because of anticholinergic effects (299,300). In the patient who remains persistently fatigued without clear organic cause, occult depression should be suspected (198,311). Additional nonmedication treatments that appear promising in older adults include behavioral treatments (e.g., stimulus control to induce good sleep hygiene behaviors) and increased exposure to light (361,362). Recent research indicates melatonin can be a useful adjunct in the management of sleep disorders in the elderly (363,364).

Pain

Pain is very common in older people, and studies are increasingly focused on this subject (365–367); prevalence estimates range from 25% to 50% of community-dwelling elderly people to 45% to 80% of nursing home residents (366,368). The consequences of pain are significant and include depression, decreased socialization, sleep disturbance, impaired ambulation, and increased health care use and costs (365). The pain experienced and reported by older people is no less threatening than that experienced by younger people and similarly, must be addressed promptly (366).

Special considerations in managing pain in older persons include difficulty in assessment secondary to patient fears, the higher incidence of comorbid illnesses compared with younger persons, complications in reporting pain in patients with memory and other cognitive impairments, validity difficulties with proxy reporting, and the importance of assessing functional implications of the pain (366,369). Furthermore, physicians understandably tend to attribute new pain to prior conditions (e.g., osteoarthritis). Cognitive impairment does not mask pain at the time of patient questioning, but accurate reporting of past pain is not necessarily reliable (370). Patients may be able to respond appropriately to pain intensity scales concerning current pain, with visual cueing as needed and taking short attention spans into account (371). Special functional considerations include the recognition that advanced and elective ADLs may be more sensitive to changes in pain.

Common etiologies of pain in older people include osteoarthritis, cancer, herpes zoster, temporal arteritis, polymyalgia rheumatica, and atherosclerotic peripheral vascular disease (368). Approaches to pain management are similar across age groups and include use of physical modalities (e.g., heat, cold, massage), transcutaneous electrical nerve stimulation (TENS), biofeedback, hypnosis, and distractive techniques (365). Concurrent depression with pain may occur as in younger persons, requiring direct assessment for depression and intervention if required (313). Of note, older patients with depression are more likely to report pain as a somatic expression of their mood disturbance (370).

Medications for pain should be prescribed judiciously and in conjunction with nonpharmacologic approaches (365). Acetaminophen remains one of the best initial medications to be used routinely in patients with pain (372,373). Nonsteroidal anti-inflammatory medications are problematic in this population, given the limited study of patients over age 65 (374) and the known fourfold higher risk of peptic ulcer disease (375). Long-term use of opiate analgesics is appropriate for malignant pain and probably in some cases of nonmalignant chronic disabling pain unresponsive to other medications. Tricyclic antidepressants or anticonvulsants may be useful in treating neuropathic pain (365). Physical mobility and activities should be encouraged as much as possible. All these treatments are best administered as part of a multidisciplinary team approach, regardless of the setting (e.g., home, SNF, or outpatient department) (366).

A particularly challenging clinical population includes older adults who experience chronic pain, often with repeated failures to respond to traditional medical or surgical treatments. The cognitive-behavioral model of therapy, developed with younger persons, may prove useful in these patients (366,367). This therapy is safer, more effective, and probably lower in cost than a long-term analgesic regimen, especially if applied early in the course of an evolving pain syndrome. This model divides contributory factors to the pain experience into biomedical variables, psychological variables (e.g., pain coping strategies, depression, personality), and socioeconomic variables (e.g., social support, spousal criticism). Behavior therapy encourages wellness behaviors, and cognitive therapy helps patients reassess how they view themselves and their pain experience. This model also
highlights the importance of assessing family behaviors in the presence of pain. Specific interventions for family coping can be beneficial (376,377). Family and caregiver training, as well as semiformal social supports in the community, may be especially important in the setting of chronic pain in older persons (366). For example, chronic back pain sufferers (typically with associated depression) tend to exhaust their social support (378). Preventing the resulting social isolation would likely improve efficacy of treatment intervention and avoid a cascade of complications (196).

**Falls**

Many of the age-related physiologic declines in multiple organ systems combine to increase dramatically the incidence of falls in the elderly, including visuoperceptual difficulty, postural instability, impaired mobility, orthostatic hypotension, lower-extremity weakness, and vertigo due to degenerative or vascular changes in the vestibular apparatus (224,225,379). Other factors contribute to increase the risk of falling, including environmental hazards, adverse effects of medications, concomitant acute or chronic disease states, depression, apathy, or confusion (330,380–382). A model attempting to identify the degree of risk for recurrent falls stratified patients into high and low risk depending on sitting and standing balance, walking ability, and stair climbing (383). In addition, attitudes toward risk were measured, as were social supports and environmental status. Recurrent falls were associated with impaired mobility, risk-taking behavior, and environmental score.

Prevention of these injurious falls is more problematic. A recent prospective study of 9,516 community-residing white women (average follow-up 4.1 years) found that the likelihood of hip fracture increased in the presence of multiple risk factors and low bone density (384). Suggested possible interventions to decrease risk include maintaining body weight, walking for exercise, avoiding long-acting benzodiazepines, minimizing caffeine intake, and treating impaired visual function. Tai chi reduced risk of multiple falls by 48% in a randomized control trial in community-residing persons 70 years of age and older without chronic illness, many of whom had fallen in the prior year (385). Whether this intervention would work for older persons with chronic illnesses needs to be assessed. In another study of community-dwelling elderly with at least one risk factor for falls, a multifactorial intervention (medication adjustments, behavioral instruction, and exercise) reduced falls from 47% in the control group to 35% in the intervention group (386). What multiple studies appear to consistently substantiate is that exercise is an important component of any fall prevention strategy (387–389).

Fear of falling resulting in decreased mobility is also a clinical problem in many older adults (390,391). Patient training in fall recovery techniques and education regarding adaptive and preventive strategies are being evaluated as methods to help prevent activity restriction from the fear of falling (392).

**REHABILITATION INTERVENTIONS FOR COMMON DISABLING CONDITIONS IN OLDER ADULTS**

**Hip Fracture**

Although 95% of falls in older persons fortunately do not result in serious injury (393), hip fractures continue to be one of the most serious sequela (394). Strategies considered for intervention to prevent hip fracture have included both public health initiatives (e.g., emphasizing weight-bearing exercise) and individualized approaches focused on high-risk patients (395–397). The most effective approaches have yet to be worked out. In situations involving rehabilitation of an elderly patient after repair of hip fracture secondary to fall, it is critical to evaluate, and treat, the cause of the incident fall to prevent future recurrence.

A number of controversies involve proper care of the elderly patient after hip fracture. The literature is increasingly evaluating factors affecting outcomes from hip fractures and potential cost-effective changes in practice (398–402).

Issues relating to preoperative decisions include how long to wait for medical stabilization. One guideline suggests that hip fracture patients who have two or fewer comorbidities should have the operation within 2 days of admission but that a longer delay is beneficial for patients with three or more comorbidities (403). Another study found that a delay of greater than 4 days, or delay due to the management of acute medical problems, resulted in increased mortality (404).

Several factors affect the decision to operate and selection of the most appropriate type of surgery. A tendency to treat hip fractures conservatively (i.e., nonsurgically) in elderly patients with dementia is countered by findings of better function with less morbidity and mortality with surgical management (405). For patients with severe cardiovascular disease that contraindicates general anesthesia, percutaneous pinning with Ender rods under local anesthesia can be performed. Femoral neck fractures can be treated either by resection of the femoral head with endoprosthetic metal prosthesis with or without simultaneous osteotomy, or by internal fixation with multiple pins with delayed weight-bearing. Although intertrochanteric fractures traditionally are managed by internal fixation with nail or compression screw with delayed weight bearing, some studies suggest patients can be mobilized much earlier without complication with and improved morbidity and function (398,401,405).

The postoperative period can be divided into the acute hospital period and posthospital (or postacute) care. The urgency of early mobilization after repair of hip fracture is twofold: the vulnerability to many postoperative complications (e.g., pulmonary problems, thromboembolism, genitourinary sepsis, delirium) and the risk of secondary complications from bed rest or relative inactivity (398,399). In one recent study of the acute hospital period after surgical repair of hip fracture, factors associated with discharge directly to home (which occurred in only 17% of 162 hip fracture survivors) included prior community residential status, age under 85 years, absence
of postoperative complications, achieving independence in bed mobility and ambulation with a walker, and a greater number of physical therapy sessions during hospitalization (401). Also, early initiation of physical therapy has been found to be beneficial for improving mobility in the first 2 months after hip fracture (406).

By far, the majority of hip fracture patients in the United States receive postacute hospital care in other facilities—either acute rehabilitation hospitals/units or skilled nursing facilities. These settings are increasingly necessary as hospital length of stay continues to dramatically decline. Which setting is best for which patient is still not clear, although a recent prospective evaluation found that hip fracture patients had greater functional recovery at 12 weeks with treatment in an acute rehabilitation setting as compared to an SNF (407). What is necessary during the recuperative phase from hip fracture, regardless of setting, is close attention to the multiple medical problems that can arise (399,400). Optimal length of stay in these settings likewise is not yet clear and also continues to decline.

Arthritis and Joint Replacements
Management of arthritic conditions in older people, just as in a younger population, must be individualized with close monitoring of benefits (405,408). Treatment principles are comparable, although the balance between rest and activity is much more delicate because of the adverse sequelae of inactivity in the elderly. There is evidence that older people with arthritis may respond better to therapeutic programs and often are more patient and compliant with long-term exercise and activity programs (409,410). Treatment goals include relief from fear, fatigue, stiffness, and pain; suppression of the inflammatory process; prevention or correction of deformity; and maximizing function (373,409). This is accomplished via a combination of psychological, pharmacologic, physical, and surgical measures.

Important psychological approaches have been developed through the Arthritis Self-Help Course. The multifaceted interventions include education and exercise (411). Part of the benefit may derive from facilitating the patient’s ability to manage his or her own chronic condition. Successful exercise interventions include programs of focused muscle strengthening (e.g., quadriceps strengthening for osteoarthritis of the knee), general conditioning, and aerobic activities (412–414).

Patients can be educated about the beneficial results of using various assistive devices to maintain independent community living, such as a firm chair of appropriate height with armrests, utensils with built-up handles, elevated toilet seat with grab bars, or ambulation aids (e.g., cane, walker). As with any patient care equipment, having patients try to use various devices before purchase will help ensure actual functional use (373).

Pharmacologic interventions for pain control should start with acetaminophen (up to 2 to 4 g/day) (372,373). Opiates (codeine or hydrocodone) can provide additional analgesia for breakthrough pain. Topical capsaicin cream may be helpful for persistent knee or finger pain, although usage may be limited by cost, need for frequent application, and initial burning sensation with application. NSAIDs may be needed for control of pain and inflammation but with great care and close monitoring given the increased risk of acute renal impairments and GI bleeding (372). In older patients with a history of gastritis or ulcers who require use of NSAIDs, concurrent administration of misoprostol or other cytoprotective agents should be considered (373,415). The use of the nutraceuticals glucosamine and chondroitin sulfate has been controversial in terms of efficacy but is currently the focus of a National Institutes of Health multi-center trial (416).

A limited number of intra-articular steroid injections can be considered (generally no more than two to three per year for any joint) but with anticipation of only short-term benefit (373). The FDA has approved another type of intra-articular injection (hyaluronic, a glycosaminoglycan) for patients who have failed other forms of therapy (373).

Age should not be a primary factor in considering potential benefits of surgical intervention in the elderly arthritic patient (372,416). Significant functional gains may be realized with an appropriately timed procedure (e.g., ligament or tendon repair, osteotomy, arthroplasty, prosthetic joint replacement) to improve stability and range or to decrease pain (409). Attention to preoperative and postoperative therapy programs and early mobilization is critical to maximize functional gains and minimize secondary complications from inactivity. Further details of rehabilitation management, including principles for prescription of medication and therapeutic modalities, can be found in Chapters 31, 40, and 61–66.

Stroke
Appropriateness of intensive rehabilitation for older stroke patients sometimes is questioned, given the limited and sometimes conflicting research data. Studies to date suggest that age may have a negative or no effect on functional outcomes, that elderly stroke patients may require longer lengths of stay to achieve the same functional gains as younger patients, and that functional outcomes may be similar in differing rehabilitation settings (417–427). The most significant aspects of rehabilitating elderly stroke patients relate to the severity of their neurologic and functional deficits, medical stability and impact of their frequent multiple comorbid conditions on endurance, and their ability to understand, cooperate, and learn in therapy sessions. Severe language or cognitive deficits, significant neglect or apraxia, poor balance or endurance, or recurrent medical complications/instability may negatively impact the feasibility and goals of a rehabilitation program.

Clinical practice guidelines for stroke rehabilitation that incorporate these and other variables have been published (428). It is clear that depression, a common complication after stroke at any age, is particularly problematic in older patients due to the deleterious effect on cognitive functioning (429). There is evidence of potential to enhance neural plasticity after stroke, even in older individuals, with improved functional recovery (430). Research continues in an effort to clarify the most appropriate
(and cost-effective) role, timing, methods, setting, intensity, and duration of rehabilitation services for older people after stroke (427,431). Chapter 23 reviews concepts of stroke rehabilitation in detail, with reference to older patients.

Amputation
Although a detailed review of rehabilitation of patients with dysvascular amputation can be found in Chapters 45 and 74, several aspects require emphasis here. An ageist bias may result in the belief that a patient’s age should be a factor in determining whether or not to prescribe a prosthesis. Other comorbidities rather than age per se are the relevant determinants for prosthetic fitting (427,432,433). A number of studies have documented the successful outcomes of rehabilitation programs for older amputees, including bilateral amputees and amputees with concurrent hemiplegia (433–437). Even in the face of severe medical comorbidity (e.g., cardiovascular disease), a prosthesis still may be both therapeutic and functional, even if only from the standpoint of standing, transfers, or cosmesis (432). For bilateral amputees, although energy costs are significantly higher and ambulation training more difficult, prosthetic fitting still may be useful to allow periodic standing during the day and for walking short distances in the home, which are therapeutic from both an aerobic exercise and psychological standpoint (434). Wheelchair locomotion will usually be a preferable alternative for longer distance travel in view of significantly lower energy costs and ability to stop and rest. The former criterion of successful crutch ambulation to justify prosthetic prescription is not appropriate (432,433).

Spinal Cord Injury
Although SCI usually is considered a disability occurring primarily in the younger population, there is increasing recognition of its significance for older people. Not only is there a significant incidence of SCI in a growing older population (5.4% in the 61- to 90-year-old age group) (438), but there is also increased survival rate in an aging population injured earlier in their lives (439). The result is a much higher prevalence of older SCI patients, subject to the usual age-related morbidity and mortality. Similarities have been observed between aging morbidity and that of chronic SCI patients who are not old (440).

Epidemiology of SCI with older age at onset differs from that of younger populations. The etiology of injury is much more likely to be falls (60% in the 75-and-older age group), followed by motor vehicle accidents (32% in the 75-and-older group) (438,441,442). SCI from metastatic disease and cervical myelopathy occurs primarily in older adults. There is a marked increase in proportionate incidence of quadriplegia (QC) and quadriparesis (QI) in the elderly (67% in the 61- to 75-year-old age group, 88% in the 75-and-older group), as opposed to the more nearly equal distribution between paraplegia (PC) and paraparesis (PI) and QC and QI in younger age groups (438). Elderly SCI patients are much more likely to be QI as opposed to QC (441).

There is a progressive disparity in 10-year survival rates between SCI and non-SCI populations with advancing age at injury (438). For those 70 to 98 years of age, the grouped SCI 10-year survival rate is 32%, compared with 48% for their non-SCI counterparts. Life expectancies reported for SCI patients differ depending on whether patients who die before discharge from rehabilitation programs—usually within the first year post injury—are included in the analysis. If such first-year fatalities are included, life expectancy for SCI patients injured at 60 years of age is 6.5 years for PI, 5.9 years for PC, 4.2 years for QI, and 1.9 years for QC, compared with 20.0 years for the non-SCI population (438). Two-year life expectancy is 59% for SCI patients between 61 and 86 years of age, compared to 95% for their younger counterparts (439,443). Older patients with SCI were more likely than their younger counterparts to develop various medical complications, such as pneumonia, GI hemorrhage, pulmonary emboli, or renal stones (444,445). Although overall survival post-SCI is reduced for older adults, with increased morbidity, there does not appear to be a direct relationship between age and functional outcome (446,447).

These significant life expectancies, and potential for functional gains, make rehabilitation efforts appropriate for all patients following SCI, regardless of age (445,446). Rehabilitation goals should be comparable with those for a younger SCI population (see Chapter 27), except as impacted by comorbidity (e.g., arthritis with limitation of hand function, deconditioning, etc.). Personal care assistance becomes more critical with regard to ability to return to the community to live.

Traumatic Brain Injury
Although the concerns regarding falls in older persons are focused primarily on the risk of hip or other skeletal fractures, there is evidence of a significant incidence of traumatic brain injury as well (448,449). Similarly to older individuals after stroke, the rehabilitation interventions after brain injury must factor in premorbid and current cognitive status, severity of neurological and functional deficits, and comorbidity issues. Experience with the federally designated Traumatic Brain Injury Model Systems Project reveals that elderly brain-injured patients are capable of significant functional improvement, but more often at a slower pace (with longer lengths of stay and higher costs) (448,449). A further consideration in the differential diagnosis of an older individual with new onset of cognitive deficits in the context of multiple falls may be the possibility of a postconcussion syndrome, rather than a traditional etiology of dementia (450). Chapter 24 provides detailed rehabilitation approaches to traumatic brain injury.

GERIATRIC REHABILITATION PERSPECTIVES
Geriatric Rehabilitation Settings
Older people with disabling medical conditions may have difficulty tolerating and participating in an intensive inpatient rehabilitation program, owing to such factors as severity of deficits, medical comorbidity, and deconditioning (451,452).
Combined with the quest for least costly health care alternatives, the ideal system of inpatient rehabilitative care would provide for varying levels of intensity and settings (268). Indeed, there is evolving interest in the role and effectiveness of various levels of postacute rehabilitative care, encompassing hospital-based comprehensive or “acute rehabilitation,” SNF, or nursing home-based “subacute rehabilitation,” long-term acute care hospitals (LTACHs), or outpatient clinic-based therapies. Each of these settings offers varying degrees and frequency of physician involvement in medical care and supervision and varying frequency and intensity of rehabilitation therapy treatments. Outcome studies continue to address the question of which older patients, with what underlying disability, will benefit most cost-effectively from which level(s) of care, and at what time postonset of disability (425,426,451,452).

Another nontraditional setting for rehabilitation is the day hospital, which provides comprehensive, relatively intensive and structured rehabilitation therapies designed to reverse disability and train family members to facilitate maintenance of the patient at home (453,454). This provides a greater intensity of therapy, with a wider array of equipment, under closer medical supervision, than usually is feasible in a home-based treatment program. Day hospitals may allow earlier transition from inpatient rehabilitation hospitals/units to the more familiar and comfortable home setting, with lower health care costs (453).

There is also an increasing interest and program development in augmented home care services, including rehabilitative care (455,456). New and innovative programs to provide intensive, rehabilitation services in the home are being developed and tested (457). A randomized controlled trial of an occupational therapy preventive assessment and treatment program for older people living in the community showed improvements across various health, function, and quality of life domains (458). Such community-based programs may prove cost-effective and feasible and help resolve accessibility problems in both urban and rural settings.

These alternative levels and settings of rehabilitation services for older adults provide the potential for a continuum of care, facilitating individually tailored rehabilitative care that can be modified to meet an individual patient’s changing needs over time. Further research is required to document the cost-effectiveness and benefits of these varied rehabilitation programs, particularly among subsets of different disability and age groups.

Critical Issues Relating to Outcomes

The rehabilitation approach to care of older individuals with disability, like geriatric care in general, must be longitudinal in perspective and coordinated with other aspects of the individual’s health care, not episodic and in isolation. The team work background and training of the physiatrist are an ideal base to accomplish this critical goal of geriatric care.

There are a number of critical issues impacting the quality, cost-effectiveness, and outcomes of rehabilitation interventions. Determining the appropriate treatment setting, timing, and duration of care is of obvious importance. Coupled with this is the need for individualized, realistic, functional, and relevant goal setting, which includes engagement by the older patient during formulation. Facilitating access to needed services is critical, whether referencing insurance coverage or physical transportation to the care setting. In keeping with the longitudinal perspective, periodic reevaluation is necessary, with review and revision of therapeutic goals as warranted.

Even though “maintenance” treatment or activity is typically unfunded (as opposed to “restorative” care), the concept nonetheless is appropriate. By “maintaining” an older individual’s functional status, he or she can stay in the community at far lower costs than if institutionalized. Waiting for a patient to deteriorate from lack of “maintenance” care risks initiating a cascade of complications with concomitant decreased prospects of regaining premorbid function. It is also important to remember the potential benefits of group therapy/activities (e.g., peer support and encouragement, even friendly competition), as opposed to individual treatment.

The need for patient and family education and training, perhaps on multiple occasions or even continually, should be recognized with respect to compliance and follow-through with health care recommendations and treatment. Finally, the longitudinal perspective mandates long-term follow-up to monitor for complications, recidivism, or underlying disease progression (with prompt intervention, if necessary), as well as to assess counseling or respite needs.

Role of the Physiatrist in Geriatrics

Physicians from various specialties traditionally have had differing perspectives on their respective roles in geriatrics (459). Physiatrists may serve a variety of roles relating to geriatrics, depending on the practice setting. These contributions range from providing primary care in a rehabilitation inpatient hospital setting or subacute (SNF-based) setting (336,460) to consulting in various health care settings (e.g., acute care hospital, SNF, day hospital, or home health care) to outpatient care (336,402,453). In the latter settings, the physiatrist assesses functional and medical statuses, works closely with therapy staff, helps formulate appropriate rehabilitation goals, helps coordinate interdisciplinary team care if needed, and monitors the efficacy of therapy.

Hoenig provides an interesting analysis of rehabilitation providers, noting that physicians of whatever specialty typically act at the disease or impairment level (e.g., prescription of medication, surgical procedure, etc.), may function as rehabilitation team leaders, and are often designated as gatekeepers in facilitating access to rehabilitation services (via prescription of therapies, insurance authorization, etc.) (461). She points out that there is a great deal of overlap and variability in the roles played by various rehabilitation providers, with a need to network, communicate, and coordinate to best serve the needs of older patients.

Further progress in this regard is demonstrated by an interdisciplinary interaction and consensus process among ten medical and surgical specialties, spearheaded and funded by the American Geriatrics Society and John A. Hartford Foundation,


Primary Care for Persons with Disability

The issue of providing quality medical care to all persons has been brought to the forefront with the current changes in medicine and health care reform. Terms such as "cost containment," "appropriate utilization of resources," and "quality management" are heard by practicing physicians on a daily basis in reference to their patient care interactions. The impetus to make changes in our current health care system has also been influenced by economic issues (1). Health care costs have increased exponentially in recent years, and even though these costs have risen significantly, patients' health outcome and satisfaction have not risen proportionately. When the United States was compared with numerous other countries with respect to health outcomes and satisfaction in relation to cost, the United States ranked lowest of the nations studied, compared with countries such as The Netherlands, which was among the best (1). When the same countries were compared with respect to their percentage of primary care physicians, the results were similar, with the United States having the smallest percentage of primary care physicians and the countries with greatest satisfaction having the greatest percentage of primary care physicians. When the cost factor is eliminated, however, health care quality in the United States is among the finest in the world. Unfortunately, it is not accessible and available to all. Because of lack of income, lack of insurance, isolation, language, culture, or physical disability, millions still face barriers to quality health care.

The U.S. Department of Health and Human Services' Health Resources Service Administration (HRSA) has seven strategic goals that support its mission. They are as follows:

- Improve Access to Health Care
- Improve Health Outcomes
- Improve the Quality of Health Care
- Eliminate Health Disparities
- Improve the Public Health and Health Care Systems
- Enhance the Ability of the Health Care System to Respond to Public Health
- Emergencies
- Achieve Excellence in Management

The 2008 budget for this organization approximates $5.8 billion, some of which is allotted to those with disabilities. The Bureau of Primary Health Care within the Health Resources Service Administration is allotted $1.98 billion for 2008 (2).

Although there is general consensus in the United States and abroad that primary care is a critical component of any health care system, there is considerable imbalance between primary and specialty care in the United States (3). The proportion of specialists in the United States is more than 70% of all patient care physicians, whereas in other industrialized countries, 25% to 50% of physicians are specialists (3). There are multiple factors and forces that have led to the dissatisfaction with primary care fields. Medical students are choosing non–primary care specialities at a greater rate each year. The percentage of U.S. medical graduates choosing family medicine decreased from 14% in 2000 to 8% in 2005 (4). Issues such as non-competitive income, excessive/unpredictable work hours, and concerns about patient outcomes due to shorter and more rushed office visits have also led to 75% of internal medicine residents eventually choosing to be subspecialists or hospitalists, rather than general internists (4).

Further work has shown that based on staffing patterns in classic health maintenance organizations (HMOs), there are about 3.1 times more pathologists, 2.5 times more neurosurgeons, 2.4 times more general surgeons, 2.0 times more cardiologists and neurologists, 1.9 times more gastroenterologists, 1.8 times more ophthalmologists, and 1.5 times more radiologists in the nation than would be needed (3). Although it is clear that there is a plethora of specialists and a need for greater primary care services, considerable room for debate remains on its true impact on health care costs and quality of patient care.

Beginning in the late 1980s and continuing to the present, recognition of the difficulties that persons with disabilities face in accessing quality health care has become apparent. Persons with disabilities represent approximately 10% of the world’s population, yet they are among the most underserved groups (5). Although the medical literature is relatively sparse on this issue, multiple conferences and publications have addressed this topic. The Association of Academic Physiatrists and the American Academy of Physical Medicine and Rehabilitation have both previously developed position statements on the provision of primary care services to persons with disabilities, lending their support to physiatrists who choose to provide these services; however, there are a variety of opinions among practitioners as to how these services are best provided (6,7).

This chapter provides an overview of the primary care issue, with special emphasis on persons with disabilities, as well as a discussion of issues on health promotion in this
population. A practical approach to primary medical care in a general population that can easily be adapted to persons with disabilities follows. The chapter concludes with a review of several models of primary care and describes management issues more specific to those with disabilities.

DEFINITIONS OF PRIMARY CARE

The HRSA has defined primary care based on the following three anchoring principles: (a) the routine medical care and services people receive on first contact with the health care system for a particular health incident, that is, prevention, maintenance, diagnosis, limited treatment, management of chronic problems, and referral; (b) assumption of longitudinal responsibility for the patient regardless of the presence or absence of disease (i.e., all of a person’s health care needs—physical, psychological, and social—are met); and (c) integration of other health resources when necessary (gatekeeper function) (8).

The Institute of Medicine (9) has provided a definition as well, which states that primary care is the provision of integrated, accessible health care services by clinicians who are accountable for addressing a large majority of personal health care needs, developing a sustained partnership with patients, and practicing in the context of family and community. “Integration” includes comprehensive, coordinated, and continuous services. “Accessibility” refers to eliminating geographic, financial, and cultural barriers to seeing the caregiver. “Health care services” includes hospitals, nursing homes, office, school, home, and intermediate care facilities. “Clinicians” can be physicians, nurse practitioners, physician assistants, or similar health care practitioners. “Accountable” refers to the clinician being responsible for quality of care, patient satisfaction, efficient use of resources, and ethical behavior. “Majority of personal health care needs” describes the full spectrum of physical, mental, emotional, and social concerns. “Sustained partnership” is a long-term relationship that includes health promotion, disease prevention, and the management of disease itself. “Context of family and community” includes an understanding of the patient’s social background and support systems.

Another approach at defining primary care is from the patient’s perspective (10). A primary care physician is a trusted physician who (a) performs all preventive care necessary to safeguard health, (b) diagnoses and treats self-limiting conditions, (c) diagnoses serious conditions and either treats those for which he or she has expertise or refers the patient to the best available expert for treatment. A strength of this functional patient-oriented definition is that it delineates the three tiers of health care provided by all primary care physicians.

Primary care may be distinguished from specialty care by the time, focus, and scope of services provided to the patients (3). Primary care as noted above is first-contact care on entry into the health care system. Specialty care generally follows primary care upon referral from the primary care provider. Whereas primary care addresses the person as a whole, specialty care usually focuses on specific diseases or organ systems. Because primary care providers see patients at their initial interface with the health care system, they are presented with a variety of symptoms and concerns that may represent early stages of disease that are not yet easily classified into specific diseases or organ systems. Through the various roles of the primary care provider, but especially the gatekeeper function, referral to specialty care occurs when organ- or diagnostic-specific disease is identified that is beyond the scope of services provided by the primary care provider. Although primary care is comprehensive in scope and is present throughout the continuum of care, specialty care tends to be limited to specific illness episodes, the organ system involved, or the disease process identified.

PRIMARY CARE ISSUES IN GRADUATE MEDICAL EDUCATION

Three specialties typically are thought of as primary care fields. The largest is general internal medicine, followed by general pediatrics and family practice. Obstetrics and gynecology, although regarded by the American College of Obstetrics and Gynecology and the American Medical Association as a primary care provider for women, does not traditionally fulfill this role because it does not meet the “whole body” medicine criterion often cited as the standard for judging whether a specialty offers primary care (11). Nevertheless, it is clear that many women consider their gynecologist to be their primary care provider (12).

In its third report (13), the Council on Graduate Medical Education (COGME) stated that generalist physicians are trained, practice, and receive continuing education in a broad set of competencies to care for the entire population in office, hospital, and residential settings; provide comprehensive age- and sex-specific preventive care; evaluate and diagnose common symptoms; treat common acute conditions; provide ongoing care for chronic illnesses and behavior problems; and seek appropriate consultation for other needed specialized services. Given these required competencies, COGME concluded that family physicians, general internists, and general pediatricians are properly trained to function as generalist physicians. Although other physicians provide elements of primary care, COGME also noted that physicians who are broadly educated as generalist physicians provide more comprehensive and cost-effective care than do other specialists and subspecialists (14).

The recent emphasis on health care system reform has sparked the decades-old debate as to who is a generalist physician and has reemerged with important implications for physician workforce policy and medical education (14). In its third report (13), COGME recommended that the nation set a goal that at least 50% of all physicians be practicing generalist physicians. The Association of American Medical Colleges (AAMC) recommended that a majority of graduating medical students be committed to generalist careers (15).
the AAMC define generalist physicians as residents who complete a 3-year training program in family medicine, internal medicine, or pediatrics and who do not subspecialize. Given the enhanced role and growing prestige of the generalist physician and the increased emphasis on primary care, other physician groups, including physical medicine and rehabilitation (PM&R), have suggested that they be included in this category (6,14). In its fourth report (16), COGME stated that the designation of a specialty being included as primary care should be based on an objective analysis of training requirements in disciplines that provide graduates with broad capabilities for primary care practice. In an analysis of the training requirements of numerous specialties, including those typically thought of as primary care specialties (family practice, internal medicine, pediatrics) and several that have proposed inclusion as a primary care specialty (emergency medicine and obstetrics and gynecology), only the previously established primary care fields actually prepared their residents in the broad competencies required for primary care practice (14). Although PM&R was not included in the analysis, based on the current training requirements and a similar analysis, it would not fulfill the necessary training needs.

The ability to provide effective primary care or appropriate treatment for life-threatening illnesses not only depends on adequate training but a continued interest in the area of concern, as well as continued experience based on an appropriate number of cases (10). The growing quality agenda and the move toward mandatory Maintenance of Certification by the American Board of Medical Specialties (and Maintenance of Licensure by the Federation of State Medical Boards) shed additional light on the increasing requirements to maintain expertise in primary care activities. It is not enough to be adequately trained during residency in specific aspects of primary care or the treatment of a serious disease or high-risk procedure because this training quickly becomes outdated in the face of the rapid advances in clinical medicine. The individual choosing to be a primary care provider needs to remain current in the management of any condition that is diagnosed and is needing treatment (10).

In its eighth report (17), COGME evaluated five models that attempted to determine projections for the number of physicians needed for the current century. The differences in the five models lie in the degree to which historic increases in the demand for specialists are assumed to continue in the increasingly competitive managed care setting. COGME hypothesized that market forces will at least balance increasing demand for specialty services resulting from new technology. Consequently, increasing demand for specialists was not as anticipated. The ultimate requirement for generalists and specialists will depend on the configuration of future health care systems. COGME anticipated increased utilization of nurse practitioners and physician assistants, both in specialty care and in primary care. The overall trend, however, was felt to be an increased emphasis on generalists within managed systems of care, thus reducing the demand for specialists.

**GENERALIST VERSUS SPECIALIST**

Although it is well accepted that there is an imbalance between primary and specialty care, numerous arguments have been advanced by advocates of both the generalist and specialist perspectives (3,13,18). The generalist perspective points to the research that supports the efficacy of primary care and indicates that generalists have broader medical knowledge and skills; are better trained in psychosocial, preventive, and community aspects of care; and provide less costly care and are more accessible, factors that make them preferable to specialists as primary care providers (18). In addition, the generalists’ cross-disciplinary skills provide for more efficient referral patterns when using their gatekeeper function.

Specialists, on the other hand, assert that their training before subspecialization is equivalent to generalists, allowing them to deal with primary care issues (similar to generalists) as well as manage problems within their specialty (that generalists might have to refer elsewhere) (18). Numerous studies support the specialist viewpoint that primary care can be provided in an efficient and cost-effective manner by the same practitioner who provides more sophisticated, specialized, and up-to-date medical services (18).

**MEETING THE POSTREHABILITATION HEALTH CARE NEEDS OF PEOPLE WITH DISABILITIES**

The traditional distinctions between primary care and specialty care become much less clear when we consider the ongoing postrehabilitative health care needs of people with disabilities (11). Here, the boundaries between primary and specialty care overlap considerably. It is not clear where one ends and the other begins. The handoff from rehabilitative care in the rehabilitation center to primary care in the community is not straightforward. This is best understood when we consider the nature of the ongoing health care needs of people with disabilities and why many primary care issues have significant rehabilitative or functional content.

It is difficult to generalize about the ongoing health care needs of people with disabilities, in part, because different disabling conditions have widely varying pathophysiologies, comorbidities, and functional consequences. These differences often obscure the fact that people with disabilities experience most of the same health conditions experienced by people without disabilities. However, people with disabilities are at greater risk for certain common health conditions than are those in the general population, often experience these conditions differently, and may require a somewhat different and extended therapeutic regimen that takes into account both their underlying impairment and their functional limitations. However, people with disabilities observe that many health care providers are often unable to look beyond the disabling condition to address the health problem that precipitated the provider-patient encounter in the first place.
SIX CHARACTERIZATIONS

There are many ways one can characterize the ongoing health care needs of people with disabilities relative to those without disabilities. At the risk of overgeneralization, we note six ways in which the ongoing health care needs of people with disabilities are different from those in the general population. These characterizations are limited mainly to people with the types of conditions commonly seen in inpatient rehabilitation settings (19,20).

First, people with disabilities generally have a thinner margin of health that must be carefully guarded if medical problems are to be averted (21). This observation applies to health conditions that people with disabilities share with the nondisabled population (e.g., upper respiratory infection, pneumonia) as well as conditions more likely to appear among people with disabling conditions (e.g., urinary tract infections, renal failure, pressure sores). It should be emphasized that people with disabilities are not “sick” and that most are generally very healthy. However, their impairments and functional limitations often render them more vulnerable to certain health problems.

Second, people with disabilities often do not have the same opportunities for health maintenance and preventive health as those without disabilities. For example, people with mobility limitations usually have fewer opportunities to participate in aerobic activity needed for good cardiovascular health, and people with paralysis may not be able to detect certain health conditions early because they cannot experience pain in certain body regions (21).

Third, people with disabilities who acquired their impairment early in life may experience onset of chronic health conditions earlier than people in the general population. For example, it is believed that people with long-standing mobility limitations are likely to have an earlier onset of coronary artery disease than the general population. Likewise, people with mobility limitations may experience an earlier onset of adult diabetes because of obesity and may experience an earlier onset of renal disease (e.g., pyelonephritis) because of a neurogenic bladder dysfunction (22).

Fourth, people with disabilities who acquire a new health condition, apart from their original impairment, are likely to experience secondary functional losses. Thus, the functional consequences of a new chronic health condition are usually more significant for a person who already has a disabling impairment. The onset of exertional angina or a rotator cuff injury, for example, may require that the person upgrade from a manual to an electric wheelchair and from a conventional automobile to an adapted van.

Fifth, people with disabilities may require more complicated and prolonged treatment for a given health problem than do people without disabilities. For example, using a plaster cast for a broken leg may be complicated by the individual’s vulnerability to a pressure sore when the individual has no sensation in the lower limbs. Likewise, a person with a disability may require a longer recovery period after an acute episode of illness or injury because of preexisting functional limitations that limit a person’s participation in various therapies (e.g., using a treadmill or exercise bicycle after an acute myocardial infarct).

And sixth, people with disabilities may need durable medical equipment and other assistive technologies that require some level of functional assessment. Today, these devices are often prescribed by physicians who have only a rudimentary understanding about the fit between various types of equipment and the needs of the individual consumer. A poor fit between the individual and an assistive device can reduce functional capacity and may induce the individual to abandon the device, the combination of which is wasteful for both the individual and society.

These six characterizations are not exhaustive. A more complete characterization will be important in sorting out the respective roles of traditional primary care disciplines and the various specialty disciplines in managing the health care needs of people with disabilities. The six characterizations point out that traditional distinctions lose their meaning when managing the health care needs of a person with a disabling or chronic health condition.

IMPACT ON HEALTH CARE UTILIZATION AND EXPENDITURES

These six characterizations are borne out in the above-average rates of health care utilization and expenditures among people with disabilities. The 2005 Medical Expenditure Panel Survey (MEPS) provides a broad overview of the health care utilization and expenditure experience of adults with selected functional limitations. Using variables available in the MEPS, we considered persons as having a disability if they meet any one of the following criteria: (a) use mobility aids or equipment, (b) are limited in major activity (work or housework only), or (c) require help or supervision with at least one activity of daily living (ADL) or instrumental activity of daily living (IADL) (i.e., essential tasks for maintaining ones living environment and residing in the community) (23).

Using this definition of disability, individuals with disabilities comprise 15.9% of the adult population (age 18+). Yet, in 2005, they accounted for 37.0% of all physician visits made by adults, 47.6% of all adult prescriptions (including refills), half of all hospital discharges, 64.4% of all nights spent in the hospital by adults, and 47.7% of all adult-related health care expenditures (Fig. 60-1).

The disproportions reported above are also reflected in estimates of the utilization and expenditure experience of individual adults with disabilities. While only 3.4% of adults with disabilities had no health care expenditures in 2005, 23.6% of individuals without disabilities had no health care expenditures. Of those with at least $1 expenditure, the median expenditure for people with disabilities was $5,663, compared to $1,133 for people without disabilities. These figures represent more than a doubling over a 9-year period. Adults with
disabilities also typically pay more out of their own pocket ($780) for health care than the nondisabled expenditure paid from all sources ($326) (Fig. 60-2). (Median total and out-of-pocket expenditures are calculated only for those who had at least a $1 total or $1 out-of-pocket expenditure, respectively. Of note, out-of-pocket expenditures do not include health plan premiums.)

ACCESS TO PRIMARY CARE

Over 80% of disabled persons have at least one secondary medical condition (24). Many of the secondary health conditions experienced by people with disabilities, however, are entirely preventable through scrupulous health maintenance strategies and timely interventions by knowledgeable practitioners before new health conditions become emergent or even life threatening (23,25,26).

People with disabilities often express frustration about the barriers that limit their access to primary care. These barriers can be grouped into (a) process barriers and (b) structural-environmental barriers (27).

**Process barriers** refer mainly to the communication, attitudinal, and knowledge barriers in the provider-patient relationship (27). Persons with disabilities often describe their providers as not being knowledgeable about the management of secondary health conditions and the impact of these conditions on their functional capacities (21,28–30). They remark that they must make considerable effort to educate their primary care providers about their disability and how it needs to be taken into consideration when new health conditions are addressed (31). In other instances, a provider may have difficulty looking past the disabling impairment to also consider the primary and preventive care needs of a patient with a disability. Nearly two decades of studies have made the same observations. In a 1989 survey of 607 respondents with...
disabilities (e.g., spinal cord injury, multiple sclerosis, cerebral palsy, postpolio) in the Washington, DC area—a seemingly dated study, researchers found that many people had difficulty finding a physician who was knowledgeable about their health care needs (32). Respondents who had a previous rehabilitation experience often indicated that they consulted with their original physiatrist when they were not confident about a therapy recommended by their primary care physician. Subsequent smaller scale studies underscore the process and communication barriers cited by individuals with disabilities (27,33–35).

Health plans and health care providers are sometimes said to lack “disability literacy” or “disability competence,” akin to “cultural competence,” when encountering individuals with disabilities (36–40). The lack of “disability literacy” and “competence” occurs despite the fact that health plans and providers have considerable contact with individuals with disabilities owing to the disproportionate use of health care services by individuals with disabilities.

Structural-environment barriers include the “physical, social, and economic” environment in which primary care services are rendered (27). They include, but are not limited to, the physical layout of the physician’s office building, the medical equipment, the patient’s financial resources, and the business interests of the provider (41). Inaccessible entrances, a lack of adjustable diagnostic equipment (e.g., radiologic equipment, examining tables), narrow hallways, and inadequate parking are all reported examples that can make the primary care visit a difficult experience (42,43).

Primary care providers may prefer to serve only a very limited number of people with disabilities because people with disabilities often take longer to process in the course of an ordinary office visit and slow down a busy office practice (23,44). Primary care physicians serving patients participating in capitation-based managed care plans have even less incentive to serve high-use populations, such as individuals with disabilities.

The rise of managed care beginning in the mid-1990s and continuing into the present is commonly thought to be adverse to individuals with disabilities. Researchers report that managed care has made access to downstream services such as rehabilitation and assistive technology more difficult than under traditional fee-for-service plans (45). Yet, they also report that individuals participating in managed health care plans have a more regular source of primary care than those in fee-for-service plans but also had somewhat less incentive to serve high-use populations, such as individuals with disabilities.

The persistence of these process and structural barriers is remarkable given the requirements of the 1990 Americans with Disabilities Act (ADA) and its ensuing regulatory requirements. Some requirements of the ADA may not apply to smaller private practices. Yet, the Act has heightened awareness of society’s obligation to better meet the accessibility needs of individuals with disabilities. Recent ADA-inspired litigation involving inaccessible health care facilities has failed to dramatically increase awareness in this area (47,48).

**MULTIDIMENSIONAL CHARACTER OF THE ISSUES**

This brief review illustrates how the boundary issues between primary and specialty care, in the case of people with disabilities, need to be addressed at several different levels. First, there is the issue of the functional and rehabilitative implications inherent in many primary care encounters. Second, there is the issue of knowledge base and whether traditional primary care providers are adequately equipped to address the needs of people with disabilities. Third, there is the issue of whether primary care providers have the facilities to accommodate people with disabilities. And fourth, there is the issue of whether our system of health care financing discourages providers, primary care or otherwise, from addressing the primary health care needs of people with disabilities.

**PHYSICAL MEDICINE AND REHABILITATION AS A PRIMARY CARE SPECIALTY**

With these issues in mind, the notion of PM&R as a primary care specialty can now be brought into focus. PM&R is the medical specialty that addresses the needs of severely disabled persons. Primary care physicians often refer these patients after a catastrophic illness. When these patients require comprehensive inpatient rehabilitation, the physiatrist usually serves as their primary caregiver during their relatively short hospital stay but subsequently sends patients back to the referring physician for follow-up general medical care. Follow-up visits at the physiatrist’s office typically concentrate on rehabilitation issues (11).

In contrast, patients with spinal cord and brain injuries often present from referring specialists, that is, neurosurgeons and orthopedic surgeons, with no regular primary care physician. After hospitalization, their follow-up care and health maintenance to prevent rehospitalization become major issues. Physicians in spinal cord and brain injury centers often choose to become the primary care physicians for these patients. Because of the specialized and often complex medical needs of the person with severe disability, primary care physicians in the community often welcome the physiatrist’s involvement and encourage his or her assistance in providing for the primary care needs of their mutual patients (11).

Primary care services are often requested, by those persons with severe disability, to be performed at the rehabilitation facility. In a survey (49) of 144 outpatients with spinal cord injury, it was found that 48% considered their rehabilitation physician as their primary care physician. Fifty-one percent of persons also requested that all of their medical care be provided by the physicians at the rehabilitation facility. The primary reasons for these requests included maintaining continuity of care with the physician who the patients felt best understood their specialized needs, as well as easier integration with the rehabilitation hospital.
for other ancillary health needs (e.g., seating clinic, physical and occupational therapies).

Although it appears that many patients would like their primary care needs met by the physiatrists, there is not a clear consensus of whether this is a practical approach to the issue of providing this care. The problems described in developing PM&R as a primary care specialty were outlined partially earlier in this chapter. Additional considerations include manpower or workforce issues, the preferences professed by current practitioners in the field as well as those in training, and concerns dealing with PM&R residency curricula and the ability to provide adequate training in primary care issues based on our present residency requirements.

As previously noted, it is widely believed that there is a shortage of generalists/primary care physicians. There are several ways to reach the goals previously set forth by COGME (13,17) and the AAMC (15). They include having more medical school graduates enter primary care fields, reducing the number of specialty residency positions, and encouraging current practitioners in subspecialties to broaden the scope of their practice to include a more primary care role (50,51). An additional option is to change the current curriculum of some of our residency programs to include more training in primary care issues.

Even with the proposed paucity of generalist physicians, it was widely held in the 1990s that there would be an oversupply of physicians by the year 2000. In fact, in COGME’s fourth report (16) and eighth report (17), it was proposed that the number of federally funded entry-level positions in graduate medical education be restricted to 110% of the number in medical school in 1993 and that 50% of the graduates should be generalist physicians. Since that time, others have made alternative suggestions, including expanding the enrollment of U.S. medical schools to fulfill the shortfall between the supply of graduates of U.S. medical schools and entry-level positions in graduate medical education (52). COGME’s 16th report (53) in 2005 set a prediction of a physician shortage, with the demand for physicians being significantly outweighed by the supply by 2020. They recommended that medical schools expand the number of graduates by 3,000 per year by 2015 in order to help the predicted shortfall of 85,000 needed physicians in 2020. It was also noted that this increase will not likely meet the needs of the predicted increase in demand alone; however, improvements in productivity and changes in the health care delivery system might help achieve balance.

The AAMC quickly followed COGME with their own statement on the physician workforce needs (54). The AAMC noted that between 1980 and 2005, the nation’s population grew by 70 million people (a 31% increase). As baby boomers (those born between 1943–1960) age, the number of Americans over age 65 will grow as well. Baby boomers with a disability and those who become disabled as they age are increasing as well. In general, the elderly/disabled average two to three times as many physician visits, therefore approximately 50% more visits to the doctor in 2020 than 2000. In addition, since 1980, first-year enrollees in U.S. medical schools per 100,000 population has declined annually. Physician retirees will increase from 9,000 in year 2000 to 23,000 in 2025. Also, it is evident that those recently completing residency are among a new generation that are asking for increased quality of life and fewer work hours. Consequently, America will have fewer and fewer doctor/work hours each year relative to our continually growing population (54). The AAMC called for a 30% increase in U.S. medical school enrollment by 2015. This will result in an additional 5,000 new medical physicians annually. They suggested that this could occur through a combination of enrollment increases in existing medical schools and the establishment of new U.S. medical schools. Even with these proposed increases, the AAMC predicts that there will still be a shortfall. With the current trend of medical students moving into more specialized practices (nonprimary care), as noted previously, and the shortfall of physicians being predicted, the need for an increase in primary care providers will be even greater.

In the late 1990s, a workforce study was performed to determine the current and future manpower needs for the PM&R practitioner (55). The model was based on the assumption that current residency capacity as well as utilization of physiatry skills remains constant at its 1994 to 1995 level. The results were that the demand for physiatrists will continue to exceed supply, on average, through the year 2000. Excess supply has emerged and will continue to emerge in selected geographic areas in the future. In order to maintain a level of demand, it was recommended that the field should emphasize the role of physiatrists in providing efficacious and cost-effective health care. An additional option would be to further broaden the scope of practice to include primary care services. An increasing trend of current graduates of PM&R residencies to enter musculoskeletal outpatient practices has also expanded this demand but not in the area of primary care. There is an ongoing effort to reassess the current and future PM&R workforce needs to help us to address our role in the provision of primary care services.

Unfortunately, there are several issues that will likely dampen the success of this effort. The above workforce projection is based on the premise that managed care will continue to grow at a moderate level. PM&R is not typically designated as a primary care provider in our current HMO and managed care systems. In addition, even if PM&R were designated as a primary care provider for persons with severe disability, it is unlikely that this would have a significant impact on the shortage of generalists. The total percentage of currently practicing physiatrists as compared with the total number of U.S. physicians is less than 1% (11). Based on our current training capacity, it is doubtful that there will ever be the number of physiatrists necessary to provide direct primary care for the large majority of persons with disability (11).

Some of the above assumptions are based on the premise that current practitioners and those in training in PM&R are willing to provide primary care services. A previous survey (51) of 106 PM&R physicians (55 physiatrists and 51 PM&R
residents) showed that only 39% agreed that PM&R should be designated as a primary care specialty. The majority also felt that these services should be restricted to those with severe disability (e.g., spinal cord and brain injuries). Overall, 53% felt that physiatrists are competent in providing general medical care, but only 38% were convinced that the current 4-year PM&R residency sufficiently prepares physiatrists to assume the role of a primary care provider.

Current PM&R training and residency curricula do not place a great focus on several areas that are essential to providing primary care services. These include health promotion and education, as well as preventive services (11). Significant changes would need to occur in our current residency requirements to provide the above training. This would likely require extension of our current 4-year training requirement, as well as substantial adjustments to our current residency experiences. A series of recommendations (56) for changes in the current PM&R residency curriculum were proposed to assist PM&R residency programs in providing for education in the provision of primary care. These include applying general preventive care principles and interventions to those with disabilities, expanding the role of “continuity clinics,” publishing a “study guide” on the topic of primary care for the disabled similar to the study guides currently published by the American Academy of Physical Medicine and Rehabilitation, and providing continuing education on common medical problems seen in the disabled and non-disabled population. Additional consideration was given to developing special or added qualifications in primary care for the disabled or expanding the availability of double board certification (current primary care–oriented combined programs involve internal medicine and pediatrics) to include family practice. Proposals such as these would require significant changes in our current training programs that may ultimately result in a negative impact on the continued growth and attractiveness of our field (11). Changes such as increasing the length of residency training and altering "quality of life issues" for residents and practicing physiatrists (the three "Ls" of primary care: low pay, long hours, low prestige) may be viewed as adverse by prospective PM&R residents (11). In fact, this may ultimately lead to further limitation of the accessibility of care to those persons with severe disabilities. Since the publication of the above recommendations, the issue of primary care within PM&R has become less prominent.

Health Promotion in Persons with Disabilities

The overall life course health profile of an individual with a disabling condition is the result of interaction among our disability management strategies, general health care practices, biologic and socioenvironmental factors, and lifestyle behaviors. Therefore, when physiatrists address the longitudinal health care needs of those with chronic disabilities, they must view disability-related health management and general health-promoting strategies as equally important components of care. In order to do this, they must enhance their frames of reference and incorporate the concepts of health promotion and secondary condition risk reduction.

Health Promotion and Related Models

Health promotion has several features that overlap with both primary care and medical rehabilitation. Most notable, all three emphasize education and encouragement of self-responsibility, and all address the potential or actual impact of a given physical or cognitive/emotional condition across several dimensions of health. Finally, all address both health maintenance and disease prevention so as to enhance and protect functional capacity over the life span.

As a general concept, health promotion describes all efforts directed toward helping individuals modify their lifestyles and behavior so as to promote a state of optimal health. Health promotion per se is not disease or health problem specific. However, the most important health-promoting behaviors recommended, as reflected by the Leading Health Indicators (LHI) in Healthy People 2010, the national health promotion agenda for the first decade of the new century, are: proper nutrition, weight control, smoking cessation, stress management, physical fitness, elimination...
of any drug or alcohol misuse, disease and injury prevention, development of social support, and access to health care for regularly scheduled health surveillance to monitor health status (57).

Early 20th century definitions of optimal health emphasized freedom from disease and issues related to hygiene. However, contemporary definitions of health typically reflect its complex multidimensional nature, incorporating biopsychosocial models related to the interaction of physical, emotional, and environmental spheres (58). Indeed, the Department of Health and Human Services defined optimal health as having a “full range of functional capacity at each life stage, allowing one the ability to enter into satisfying relationships with others, to work, and to play” (59). It posits that the pursuit of optimizing health in a given individual is impacted by the interplay between genetics, environment, lifestyle behaviors, personal health care availability, provider performance, and disease/condition management.

The merging of contemporary definitions of health, including the World Health Organization (WHO) International Classification of Functioning, Disability and Health (ICF) (60), physiatry-oriented models of care, and the disabled individual’s own perceptions of well-being, can provide a conceptual framework upon which specific health-promoting and secondary condition risk reduction strategies can be formulated within the context of chronic disability. On many levels, the notion of health promotion for individuals with disabilities is a relatively new and emerging area in research and prevention programming. However, the impetus for closing the gap of disparity for this population of individuals is reflected in the agendas of Healthy People 2010, Healthy People 2020, and the relatively recent Surgeon General’s Call to Action to Improve the Health and Wellness of Persons with Disabilities (61). The guiding principle for this Call To Action is as follows: “Good health is necessary for persons with disabilities to secure the freedom to work, learn and engage in their families and communities.” There are four goals that will help accomplish this vision (61):

GOAL 1: People nationwide understand that persons with disabilities can lead long, healthy, productive lives.

GOAL 2: Health care providers have the knowledge and tools to screen, diagnose, and treat the whole person with a disability with dignity.

GOAL 3: Persons with disabilities can promote their own good health by developing and maintaining healthy lifestyles.

GOAL 4: Accessible health care and support services promote independence for persons with disabilities.

Health Protection and Secondary Risk Reduction

Intimately related to health promotion is health protection. Health-protecting behaviors, although overlapping to some degree with health-promoting behaviors, emphasize preventive measures that guard or defend an individual against specific injuries or illnesses. Public health interventions designed to protect health have traditionally been conceptualized as primary, secondary, and tertiary preventive measures. Primary prevention involves risk reduction and includes those activities undertaken to reduce the circumstances that would result in the subsequent development of a disease process or illness. Attention can be directed to the host (e.g., immunization, counseling on lifestyle behaviors), to the environment (e.g., elimination of physical hazards), or to a specific agent, if one is identified (e.g., contaminated water). Secondary prevention emphasizes the early detection and prompt intervention against asymptomatic disease processes in evolution. Screening efforts characterize this level of prevention. Tertiary prevention attempts to minimize disability from existing disease through medical treatment, education, and rehabilitation. Efforts to prevent the development of secondary conditions known to occur in those with specific disabilities incorporate principles from all three of these traditional models of prevention/health protection (62,63).

Expansion of these public health concepts for application among those with disabilities can provide the conceptual grounding for the development of disability-specific prevention protocols. Specifically, primary prevention for those with an existing disability should include appropriately tailored measures to eliminate risk factors for chronic conditions not necessarily directly related to their primary disability. Interventions may include protocols for health-promoting activities such as smoking cessation, weight control, reduction of substance abuse, increasing physical activity if feasible, and screening for age- and gender-specific carcinoma. Tailoring of these measures includes deliberate attention to the economic, logistic, architectural, and attitudinal obstacles to primary health care often encountered by persons with disabilities. The high prevalence rates of lifestyle risk factors amongst persons with disabilities warrants systematic attention to these areas (Table 60-1).

Secondary prevention measures in those with chronic disability should focus on ongoing anticipatory strategies to minimize the adverse health impact over time of the primary disability, superimposed aging issues, or new injuries. Emphasis is placed on early detection of secondary conditions that, if left unaddressed, can have deleterious effects on organ systems, performance of ADLs, and/or community reintegration over time (58,61). Tertiary prevention measures are then activated when appropriate.

### TABLE 60.1 Prevalence of Risk Factors in Persons with Disabilities and Persons Without Disabilities

<table>
<thead>
<tr>
<th></th>
<th>Persons with Disabilities (%)</th>
<th>Persons without Disabilities (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Current smoker</td>
<td>30.2</td>
<td>21.7</td>
</tr>
<tr>
<td>Physically inactive</td>
<td>25.3</td>
<td>13.4</td>
</tr>
<tr>
<td>Obese</td>
<td>20.5</td>
<td>18.6</td>
</tr>
</tbody>
</table>

Tertiary prevention incorporates ongoing interval efforts to maximize and maintain functional capacity over the life course. Education in new skills and equipment is pursued as functional abilities change. Strategies to combat secondary deterioration in the performance of instrumental ADLs are pursued with attention directed to stabilizing or improving access to comprehensive specialized care and stabilizing access to personal care assistance/support services. Additional attention may be given to ongoing vocational rehabilitation and problem solving and the socioeconomic disincentives and obstacles often experienced by disabled individuals seeking a place in the workforce.

**Health Literacy: Building a Knowledge Base**

Health literacy is the capacity to obtain, process, and understand basic health information and services needed to make appropriate health decisions (58). It has been shown that thoughtful advice and counsel (i.e., patient education) about behaviors, lifestyle, and self-care practices that influence overall health can have far greater impact on health and longevity than specific screening tests or procedures (64,65). It is for this reason that emphasis on awareness and education is at the crux of health promotion activities. However, the success of health promotion educational efforts depends on a variety of issues that can be grouped into three categories: (a) issues related to health care professionals themselves, (b) issues related to the patient, and (c) issues related to clinical/environmental circumstances (66). Comments here are limited to the first category. Specifically, addressing the former, rehabilitation professionals may lack self-efficacy regarding their knowledge base and skills necessary for education and motivating individuals in health-promoting behaviors. Additionally, their own personal health enhancement beliefs and practices, their underestimating of patient interest or motivation, or their overestimating of patient knowledge will also influence the nature of clinical encounters and related patterns of education or referral (66,67). Continuing medical education activities, collaboration with primary care providers in the community, small group discussion, and case studies can be useful to physiatrists interested in building their knowledge base over time (68,69).

In order to facilitate effective patient education and behavioral change, the rehabilitation professional must have a clear understanding of available epidemiologic assessment and intervention information related to commonly encountered disability-related and general health–related issues. Upon establishing this knowledge base, the physiatrist who so chooses can then routinely pose and answer the following questions during routine medical encounters:

1. What are potential disability-related or general health–related problems of which the patient should be aware?
2. What steps are necessary to clarify if the patient is at risk for specific conditions?
3. Is the problem present?
4. If present, what should be done?

Questions 1 and 2 require knowledge of (a) available epidemiologic data and risk factor information and (b) specialty-specific technical assessment skills. Question 3 requires skills in interpretation of the data secured, and Question 4 requires knowledge in appropriate education and therapeutic intervention options (57,70,71). Several factors other than knowledge predict the likelihood that individuals will adopt health-promoting and health-protecting behaviors. While providing knowledge and education are important parts of effective program design, other factors that affect the health decisions of patients should also be considered. Individual perceptions of risks, perceptions of self-efficacy in adopting recommended behaviors, physical and social environmental factors, and the perceived costs and benefits are four factors that inform personal health decisions. Incorporating the message design guidelines presented below can help more effectively promote behavior changes that result in enhanced health status (72).

To this end, investigators (72) have developed five specific guidelines regarding message design:

1. Messages should contain features that relate appropriate levels of risk.
2. Messages should contain features that bolster consumers’ beliefs that they are capable of adopting the recommended behaviors.
3. Messages should contain features that promote efficacy of recommendations.
4. Messages should contain features that encourage consumers to overcome environmental and social impediments.
5. Messages should contain features that promote the benefits and minimize costs.

While “marketable messages” are typically considered the purview of educational materials, the physiatrist can incorporate the specific principles into routine physician-patient encounters.

**The Interdisciplinary Assessment of Health**

A systematic health assessment is necessary to establish and document an individual’s current health status, lifestyle practices, and psychosocial variables that can influence health. Thereafter, goals can be established and pursued in a manner appropriate to the individual’s unique circumstances, resources, and personal desires (70). Particular attention must be directed to the social and environmental aspects of living with a disability that are typically of a magnitude sufficiently significant to greatly influence health-promoting activities (69,73–75).

Health risk appraisal (HRA) instruments are often used in the general population in risk factors assessment, education, and behavioral change programs. However, there are no such guideline-driven instruments for use in the disabled populations. Nonetheless, simple icebreaker-type health status assessments designed for use in the general population can be used to stimulate interest and initiate health promotion discussions between disabled individuals and rehabilitation professionals. Additionally, despite the current lack of valid and reliable HRA instruments for those with disabilities, the adoption of certain
general health risk assessment principles related to smoking, weight control, nutrition, physical activity, cancer screening, mental health, and family history risk factors for disease seems appropriate for those with chronic disabilities (66). While the potential benefits of inquiring and counseling related to lifestyle behaviors are known, studies have indicated that individuals with disabilities are often not asked about these dimensions of health during routine visits. It is postulated that medical attention is often directed to disability-specific issues without attention toward general health-screening inquiries (Table 60-2). Table 60-3 provides a partial list of health assessment issues that should be addressed with individuals with physical disabilities. The following section reviews specific general health-promoting activities that should be addressed and encouraged during physiatric encounters wherein the physiatrist is the functional primary provider.

### Smoking Cessation

Smoking only further compromises a disabled person’s already impaired physiologic reserves for good health. Therefore, it is vitally important for physiatrists to encourage and help every patient who smokes, to quit. Cigarette smoking remains the single most preventable cause of death (78,79,81). Unfortunately, only half of smokers report that their primary care physicians have even asked about their smoking status during the past year (82).

This lack of physician attention to smoking status and smoking cessation recommendations is particularly disturbing in light of the fact that options for office practice-based smoking cessation strategies are readily available from a variety of sources (83,84). The Centers for Disease Control (CDC) and the Agency for Health Care Policy and Research (AHCPR) have produced clinical practice guidelines on smoking cessation (85). Emphasis was placed on assessment and intervention strategies designed to be brief, requiring 3 minutes or less of direct clinician time. Physicians were encouraged to include smoking status (current, former, never) on the vital signs stamp at every clinic/office visit. The practice guidelines recommend that clinicians acknowledge that the majority of smokers who quit smoking will gain weight. Patients should be told that most individuals will gain less than 10 lb, although a small minority may gain substantially more (80,86). Follow-up for positive reinforcement or pre-behavioral change encouragement over time is encouraged. The physiatrist and interdisciplinary team, in particular, are often in a relationship with patients wherein cross-discipline reinforcement and follow-through can be instituted during routinely scheduled inpatient or outpatient clinical encounters.

### Nutrition and Weight Control

For many individuals with chronic disability, the promotion and protection of their nutritional health can be compromised by a variety of physical and socioeconomic difficulties. Bulbar signs/swallowing difficulties, limitations in upper-extremity function, income restrictions, and food procurement and preparation difficulties are several confounding variables that can adversely impact food selection and consumption.

Mounting scientific evidence, however, has underscored the direct relationship between diet and health (87). Poor diet has been implicated or identified as a risk factor in several chronic diseases that have become the leading causes of death in the United States: heart disease, diabetes, stroke, and some forms of cancer. In response to this, numerous health agencies, most notably the U.S. Department of Health and Human Services, the American Heart Association, and the National Cancer Institute, developed proposed dietary guidelines for the American population (83,88,89).

For those with disabilities, nutritional health-related secondary conditions can manifest themselves in the form of compromised skin integrity, increased skin vulnerability from cachexia, suboptimal wound-healing capacity, adult-onset glucose intolerance and dyslipidemias (90), bowel evacuation problems, fluid intake-related genitourinary tract difficulties, and functional compromise associated with weight gain. Identifying and reducing environmental, socioeconomic, and disability-related physiologic or functional skills risk factors for poor nutritional health often require a team approach (91).

### Table 60.2: Percentage of Counseling/Inquiries Made at Last Health Maintenance Examination for Persons with Disabilities

<table>
<thead>
<tr>
<th>Counseling/Inquiry Topic</th>
<th>Inquiry Made (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Changes in functional status</td>
<td>38.3</td>
</tr>
<tr>
<td>Emotions</td>
<td>41.8</td>
</tr>
<tr>
<td>Diet</td>
<td>35.3</td>
</tr>
<tr>
<td>Exercise</td>
<td>38.8</td>
</tr>
<tr>
<td>Smoking</td>
<td>15.4</td>
</tr>
<tr>
<td>Alcohol consumption</td>
<td>13.4</td>
</tr>
<tr>
<td>Sleep</td>
<td>41.3</td>
</tr>
<tr>
<td>Pain</td>
<td>46.3</td>
</tr>
<tr>
<td>Sunscreen use</td>
<td>11.9</td>
</tr>
<tr>
<td>Sexuality(ever discussed)</td>
<td>28.4</td>
</tr>
</tbody>
</table>

### TABLE 60.3 Health Promotion Assessment and Health Maintenance Activities in Those with Physical Disabilities: Secondary Impairment and Secondary Disability Screening and Early Detection Activities

<table>
<thead>
<tr>
<th>Physical Health Functions</th>
<th>Assessment</th>
</tr>
</thead>
</table>
| **Respiratory** | Smoking/exposure to second-hand smoke  
Influenza and pneumococcal vaccine status, particularly in those with advanced age, higher levels of paraplegia/tetraplegia, or neuromuscular compromise that influences muscles of respiration  
Fund of knowledge regarding management of early signs of chest congestion  
Access to assistance with secretion mobilization  
Posture/kyphosis/truncal spasticity/abdominal distention/obesity problems that can have potential impact on chest expansion/respiratory function  
Aspiration risk/gastroesophageal reflux  
Forced vital capacity and forced expiratory volume in 1 s  
Aging-related phenomena |
| **Cardiovascular** | Risk factors: family history, physical inactivity/decreased exercise capacity, dyslipidemia, abnormal carbohydrate metabolism, obesity, smoking, impaired peripheral circulation, hypertension  
Exercise practices in individuals capable of exercise/adapted exercise  
Fund of knowledge regarding diet, weight control, physical activity options, and smoking cessation resource options |
| **Skin** | Fund of knowledge on risk factors for skin breakdown and how to resolve early problems  
Frequency of visual skin inspections  
Frequency of pressure reliefs  
Evolving latex sensitivity/allergic reactions  
Moisture issues/fecal or urinary incontinence/friction  
Age and condition of wheelchair cushion/seating system  
Posture/pelvic obliquity/spasticity-induced shearing and pressure issues  
Weight/nutritional status and impact on skin vulnerability  
Transfer skills  
Pedal edema/shoe or orthotic trim lines and fit  
Palmar skin protection  
Cigarette smoking  
Sunscreen utilization  
Aging-related changes |
| **Neuromusculoskeletal** | Upper-extremity neuromusculoskeletal pain issues with or without current functional impact rotator cuff pathology; shoulder, elbow, wrist, or phalangeal joint contractures; ulnar and median nerve entrapment syndromes  
Lower-extremity/axial skeleton degenerative changes with impact on ambulation skills  
Osteoporosis/fractures  
Spasticity/tone-related impact on transfer safety, posture, or shearing phenomenon on skin  
Stability of sensory/motor profile  
Charcot joint arthropathy and impact on trunk stability/dysreflexia patterns/pain issues  
Neuropathic pain phenomena (new onset vs. stable pattern; functional impact)  
Motor coordination  
Activities to enhance strength, endurance, and flexibility |
| **Genitourinary** | Bladder hygiene/bladder management technique and equipment/supplies/adjunct therapies  
Infection/incontinence rates  
Urinary tract stone formation (upper and lower tracts)  
Genitourinary system surveillance history to date  
Gynecologic history: including access to Papanicolaou smear/bimanual examination/amenorrhea/breast examination habits/postmenopausal hormone replacement  
Prostate health assessment  
Fertility and sexuality concerns |
| **Gastrointestinal** | Bowel evacuation technique/duration/predictability  
Early identification of hypomobility patterns  
Latex allergy or evolving sensitivity  
Nutrition and hydration |
Physiologic and body composition differences in many individuals with chronic disability influence the reliability and utility of standard nutritional assessment tools. Basic biochemical and clinical assessment can be pursued based on principles described elsewhere in this text. Data relating to dietary history, social situation history, and functional status components of the nutritional health assessment can be obtained through the interdisciplinary treating team’s collaborative efforts. Action plans can be proposed based on the assembled data.

**Physical Fitness**

The cycle of disability has been characterized as a vicious cycle wherein physical disability and dysfunction through physical inactivity and deconditioning leads to additional/perpetuated physical disability and dysfunction (92) (Fig. 68.3). Health promotion and related educational efforts for those with disabilities would therefore be incomplete without the provision of a physical fitness component. Exercise regimens adapted to the characteristic limitations or physiologic vulnerabilities created by their specific impairments/disabilities are a must.

Physical fitness encompasses the physiologic attributes of (a) cardiopulmonary fitness, (b) muscular strength, (c) muscular endurance, (d) flexibility, and (e) body composition. Training principles and programs for those with disabilities have steadily been investigated and/or developed over the past 20 years since the 1996 Surgeon General’s Report on Physical Activity and Health included an unprecedented section specifically directed to persons with disabilities (93). Benefits of physical activity, including improvements in stamina, muscle strength, promotion of general feelings of well-being, and modulation of joint swelling and pain associated with arthritis, were cited. The importance and positive correlation of social support and the regular physical activity in those with disabilities were noted.

However, for many with disabilities, physical fitness cannot be defined in terms of the traditional conceptions of fitness in the general population. Rather, it must be defined in a broader sense that encompasses not only appropriately modified descriptions of endurance, strength, and flexibility but also (a) conscientious nutritional health practices to

---

**TABLE 60.3 Health Promotion Assessment and Health Maintenance Activities in Those with Physical Disabilities: Secondary Impairment and Secondary Disability Screening and Early Detection Activities (Continued)**

<table>
<thead>
<tr>
<th>Physical Health Functions</th>
<th>Assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Functional status</td>
<td>Anticholinergic use for detrusor hyperreflexia</td>
</tr>
<tr>
<td></td>
<td>Gallstone disease</td>
</tr>
<tr>
<td></td>
<td>Colon cancer (age specific)</td>
</tr>
<tr>
<td></td>
<td>Symptomatic hemorrhoids (bleeding: dysreflexia-inducing)</td>
</tr>
<tr>
<td></td>
<td>Interval changes in function secondary to changes in strength; endurance, balance, vision, hearing, hypotension, fatigue levels, pain status, cognition, polypharmacy, access to personal care assistance/transportation/health care resources in community, or life stressors</td>
</tr>
<tr>
<td>Mobidity</td>
<td>Cognitive intellectual</td>
</tr>
<tr>
<td></td>
<td>Communication</td>
</tr>
<tr>
<td></td>
<td>Social attitudinal</td>
</tr>
<tr>
<td></td>
<td>Depression</td>
</tr>
<tr>
<td></td>
<td>Knowledge base regarding physical activity/exercise options appropriate for the nature of their disability</td>
</tr>
<tr>
<td></td>
<td>Family and social support systems</td>
</tr>
<tr>
<td>Nutritional assessment</td>
<td>History and physical examination</td>
</tr>
<tr>
<td></td>
<td>Body weight</td>
</tr>
<tr>
<td></td>
<td>HDL and LDL cholesterol, fasting serum glucose</td>
</tr>
<tr>
<td></td>
<td>Swallowing difficulties</td>
</tr>
<tr>
<td></td>
<td>Functional, economic, or environmental influences on dietary habits/adequacy of nutrient intake</td>
</tr>
<tr>
<td></td>
<td>Medication effect on appetite/bowels</td>
</tr>
<tr>
<td>Tobacco, alcohol, and sub-stance use evaluation</td>
<td>Nutrition intake/diet composition</td>
</tr>
<tr>
<td></td>
<td>Smoking/second-hand smoke exposure</td>
</tr>
<tr>
<td></td>
<td>CAGE screening questionnaire/brief MAST screening tool</td>
</tr>
<tr>
<td></td>
<td>Adult immunizations</td>
</tr>
<tr>
<td></td>
<td>Dental health</td>
</tr>
<tr>
<td></td>
<td>Ocular health</td>
</tr>
</tbody>
</table>

nourish the body, (b) health practices and medical follow-up to minimize the risk of secondary impairments and disabilities, and (c) a balanced commitment to maximizing functional capabilities in both necessary and discretionary activities of daily living (94).

Well-planned, clinically sound physical exercise can improve cardiovascular and/or peripheral muscle endurance, enhance strength and coordination, and improve flexibility in many individuals with chronic disabilities. Therefore, despite a variety of practical challenges often articulated, the interdisciplinary team should systematically evaluate and recommend physical exercise activities for most individuals with chronic disability. In those with neuromuscular disease, there is a common concern that strength and endurance exercises have the potential to create exercise-induced weakness. This concern can be tempered through specific recommendations for nonfatiguing (intensity and duration of) exercise activities and with monitoring of functional capacity after introduction of regular exercise activities (95). For those with arthritis, isometric exercise programs can often maintain strength if pain and inflammation with movement are particularly problematic. However, attention must be directed to maintaining functional range of motion. Overall, strength and endurance programs for individuals with arthritis have been shown to result in better disease outcomes (96). Exercise programs for individuals with the most common forms of multiple sclerosis can be customized to address issues related to fatigue and heat intolerance. Attention should be directed to balancing frequency, duration, and time of day so that the exercise activity does not compromise the individual’s ability to perform activities of daily living (97–99). For those with spinal cord injury, the level of injury will influence the cardiovascular response to exercise. Generally speaking, the higher the level of injury, the more likely a significant reduction in cardiopulmonary capacity and fitness, as compared with those of the nondisabled population (100). However, the gains in peripheral muscle strength and endurance often enhance functional capacity and should therefore be encouraged. Recommendations for adaptation of exercise equipment and regimens to accommodate weakness, sensory deficits, or orthopedic limitations created by various disabilities are increasingly more available through a variety of resources.

**Screening for Substance Abuse**

The importance of screening for drug and alcohol abuse in the population with disabilities is necessitated by both the impact of these substances on bodily functions and by the behavioral aberrations associated with excessive use (101). A large percentage of individuals admitted to rehabilitation services with neurotrauma were intoxicated at the time of injury and often had preinjury histories of hazardous ethanol consumption patterns (102–104). Additionally, for some individuals with chronic disabilities, maladaptive coping may include in temperate or frankly abusive alcohol consumption. For others, excessive alcohol consumption may be a form of maladaptive self-medication for chronic pain phenomena. Finally, given the high prevalence of alcoholism found in the community-based general population studies and the morbidity and mortality associated with alcohol abuse and dependency, public health agencies recommend that screening for alcoholism should be a routine part of every medical evaluation (105). One first seeks to help the problem drinker acknowledge the problem, understand its consequences, and recognize the need for treatment. Thereafter, attention shifts to negotiating and carrying out an acceptable, customized treatment plan (106). Most screening instruments have been designed for substance abuse treatment populations. However, the 4-item CAGE (107) and the Alcohol Use Disorders Identification Test (AUDIT) (108) have been extensively tested and utilized in primary care settings. The CAGE questionnaire was designed as an easy, expedient instrument to evaluate a patient’s alcohol usage and to determine if further assessment is necessary (107). Although a limitation of this instrument is that it relies on self-report, it appears that those who drink intermittently are more inclined to give accurate responses to CAGE questions when the questions are part of a series of lifestyle questions that include diet, exercise habits, smoking, and safe sex practices (108). The CAGE interview questions are shown in Table 60-4. The World Health Organization AUDIT (108,109) attempts to identify drinkers whose consumption patterns place them at risk for direct or indirect medical problems and alcohol dependency, before frank dependency has developed. The AUDIT and the guidelines for its use can be downloaded from the National Institute of Alcohol Abuse and Alcoholism Web site www.niaaa.nih.gov.

The diagnosis of alcohol dependence or abuse involves identifying the quantity and duration of consumption, any physiologic manifestations of ethanol addiction, loss of control over drinking, and damage to physical health and social functioning. Routine screening may facilitate early detection of hazardous consumption patterns before frank dependency or abuse.

### TABLE 60.4 The CAGE Questions

<table>
<thead>
<tr>
<th>CAGE Questions:</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. In the last 3 mo, have you felt you should cut down or stop drinking? Yes  No</td>
</tr>
<tr>
<td>2. In the last 3 mo, has anyone annoyed you or gotten on your nerves by telling you to cut down or stop drinking? Yes  No</td>
</tr>
<tr>
<td>3. In the last 3 mo, have you felt guilty or bad about how much you drink? Yes  No</td>
</tr>
<tr>
<td>4. In the last 3 mo, have you been waking up wanting to have an alcoholic drink (eye-opener)? Yes  No</td>
</tr>
</tbody>
</table>

Each affirmative response earns one point. One point indicates a possible problem. Two points indicate a probable problem.
develops (110). Hazardous consumption has been defined as four or more standard drinks per day (or >14 per week) in men and three or more in a day (or >7 per week) in nonpregnant women. A standard “drink” is defined as approximately 12 ounces of beer, 5 ounces of wine, or 1.5 ounces of distilled liquor, the equivalent of 0.6 ounces (14 gm) of ethanol.

Once a problem is identified, a multifaceted, personalized, long-term management plan operationalized in collaboration with available community resources and family should be pursued. Transportation issues, architectural and attitudinal barriers within community-based facilities, and appropriate social support are specific challenges to systematic management in those individuals with both chronic disability and substance abuse problems.

**PRIMARY MEDICAL CARE IN THE GENERAL POPULATION**

Earlier in this chapter, several definitions of primary care were reviewed. The definitions were notable for the broad spectrum of services a primary care physician must address, including disease prevention, health education, and risk reduction. Additionally, this chapter presented differences regarding the ongoing health care needs of people with disabilities relative to those without disabilities. Several of those differences included people with disabilities having an increased vulnerability to certain health problems, an increased risk that a new health problem may impose significant secondary functional losses, and concerns that a given health condition may result in unusually complicated and prolonged treatment make the attention to health prevention and risk-factor modification in people with disabilities of paramount importance. Arguably, primary care for people with disabilities takes on a more critical role than for people without disabilities.

**The United States Preventive Services Task Force**

The United States Preventive Services Task Force (USPSTF) was convened by the United States Public Health Service in 1984. The USPSTF consists of 15 experts from the specialties of family medicine, pediatrics, internal medicine, obstetrics and gynecology, preventive medicine, public health, behavioral medicine, and nursing. Its mission is to systematically review the evidence of effectiveness of clinical preventive services and create age-, gender-, and risk-based recommendations about services that should routinely be incorporated into primary medical care (111). The initial efforts culminated in the Guide to Clinical Preventive Services in 1989, with a second edition published in 1996. The current Guide to Clinical Preventive Services, 2007, is available online at http://www.ahrq.gov/clinic/pocketgd07.

In addition to the USPSTF, there are many other major health organizations that propose recommendations. These organizations, which include the American College of Physicians, the American Heart Association, the American Academy of Family Physicians, and the American Cancer Society (ACS), oftentimes disagree on recommendations as the literature available for review may be inconclusive.

Most primary care organizations focus their recommendations on the following services, which will provide an outline for the following discussion:

- Immunizations
- Screening tests
- Counseling (discussed earlier in the chapter)
- Chemoprevention

**Immunizations**

Immunizations are one of the easiest and most effective ways to prevent potentially catastrophic illnesses. Unfortunately, they are often overlooked during patient encounters that are already full of multiple issues, especially in adults. However, it is incumbent upon primary care physicians to address the need for vaccinations, as one component of preventive services. The following is a list of immunizations, and the appropriate dosing schedules, that are recommended by the Center for Disease Control’s Advisory Committee on Immunization Practices (ACIP) (112).

**Childhood Vaccines**

For childhood immunization schedule, see Figure 60-3.

**Diphtheria, Tetanus, and Pertussis**

A child needs five diphtheria, tetanus, and pertussis immunization (DtaP) shots for maximum protection. The first three shots should be given at 2, 4, and 6 months of age. The fourth (booster) shot is given between 15 and 18 months, and a fifth shot (another booster) is given when the child is about to enter school, at 4 to 6 years of age. When DTAp vaccine is given according to this schedule, it protects most children from all...
three diseases. If a child does get one of the diseases in spite of the vaccine, it will probably be milder than it would have been otherwise.

**Polio**

There are two types of polio vaccines: inactivated (killed) polio vaccine (IPV), which is an injection, and live oral polio vaccine (OPV). The Centers for Disease Control and Prevention recommends only IPV, except in very limited circumstances. Children should get a total of four doses of IPV. The first two shots should be given at 2 and 4 months of age. The third shot should be given between 6 and 18 months, and a fourth shot (booster) is recommended when the child is about to enter school, usually at 4 to 6 years of age.

**Measles, Mumps, and Rubella**

Most children get measles, mumps and rubella vaccines all together in a single injection called Measles-Mumps-Rubella (MMR). All three of these vaccines work very well and will protect most children for the rest of their life. Children should get two doses of MMR vaccine. The first is given between 12 and 15 months of age. The second may be given at any time, as long as it is at least 28 days after the first. It is usually given at 4 to 6 years of age, before the child enters kindergarten or first grade.

**Haemophilus influenzae type B (HiB)**

Haemophilus influenzae type B (HiB) vaccine has had a dramatic impact on Haemophilus influenzae type B infection. As soon as the first vaccine was introduced in 1985, the disease began to disappear. Several improved vaccines have been licensed since then, and the age for the first shot has been lowered from 24 months to 2 months. There were an estimated 20,000 cases of HiB disease a year in the mid-1980s, but now there are only a few hundred cases a year.

Children should get either three or four doses, depending on which companies’ vaccine is used. All children should get the vaccine at 2 and 4 months of age, and a booster dose between 12 and 15 months. Some children should get an additional dose at 6 months. Children who have passed their fifth birthday do not need HiB vaccination.

**Pneumococcal**

Until very recently, pneumococcal vaccine was recommended mostly for adults over 65 years old and was not licensed at all for children under 2. The lack of licensing for children was because the only type of vaccine that was available (pneumococcal polysaccharide vaccine) did not work very well for young children. Now, a different type of vaccine (pneumococcal conjugate vaccine) has been licensed that is efficacious for children under 2. Thus, it is finally possible to prevent pneumococcal disease in that vulnerable age group.

Children should routinely get four doses of the vaccine, one dose each at 2, 4, 6, and 12 to 15 months of age. Children who begin the series later may not need as many doses. Children over 5 years old will generally not get the pneumococcal conjugate vaccine.

**Hepatitis B**

Children can be protected from hepatitis B via immunization with three doses of hepatitis B vaccine. Babies can get their first shot between birth and 2 months of age, the second at 1 to 4 months of age, and the third at 6 to 18 months of age. After the third injection, most children will be protected and will not need booster doses.

**Varicella**

Varicella vaccine is a live-virus vaccine. It has been used in some parts of the world, such as Japan, for over 20 years. It was licensed in the United States in 1995.

Two doses of varicella vaccine are recommended for children. The first dose should be given between 12 and 15 months of age, while the second dose should be given between 4 and 6 years of age. A child who has already had chickenpox disease does not need to get the shot. Occasionally, even children who respond to the vaccine get a very mild case of chickenpox (about 1% to 2%).

**Influenza**

Influenza vaccinations for very young children have been a recently added recommendation by the CDC. All children between 6 months and 5 years of age should get yearly influenza vaccinations.

**Rotavirus**

Rotavirus is the most common cause of gastroenteritis in infants and young children worldwide. Nearly every child in the United States is infected with rotavirus by 5 years of age. In 2006, a live, oral vaccine was licensed for use. Three doses are recommended at ages 2, 4, and 6 months of age.

**Hepatitis A**

Hepatitis A vaccine is an inactivated (killed) vaccine. All children should receive the vaccination. The first of two injections should be given at 12 months with the second one given 6 to 12 months after the first.

**Adult Vaccinations**

Some adults incorrectly assume that vaccines they received as children will protect them for the rest of their lives. Generally, this is true, except that

- Some adults were never vaccinated as children
- Newer vaccines were not available when some adults were children
- Immunity can begin to fade over time
- As we age, we become more susceptible to serious disease caused by common infections (e.g., flu, pneumococcus) (113)

The following vaccines are recommended for adults.

- Varicella (chicken pox) vaccine
- Hepatitis B vaccine (adults at risk)
- MMR vaccine
• Tetanus, diphtheria, and pertussis
• Influenza vaccine
• Pneumococcal vaccine
• Human papillomavirus (HPV)
• Shingles

**Varicella (chicken pox)**

The varicella vaccine is recommended for persons of any age without a reliable history of varicella disease or vaccination or who are seronegative for varicella. It is also recommended for persons who live or work in environments in which transmission of varicella is likely, such as teachers of young children, day care employees, college students, or military personnel. For persons 13 years of age and older, two doses separated by 4 to 8 weeks are the appropriate schedule.

**Hepatitis B**

Hepatitis B vaccine is recommended for adults with multiple sexual partners, a recent diagnosis of a sexually transmitted disease, occupational risk of exposure to blood or body fluids, household contacts and sexual partners of those with chronic hepatitis B infection, injection drug users, and all unvaccinated adolescents. The schedule is three doses, with the second dose 1 to 2 months after the first and the third dose 4 to 6 months after the first.

**Measles-Mumps-Rubella**

MMR is recommended for adults born after 1956 without written documentation of immunization, health care personnel born after 1956, and adolescents entering college. Women should be asked if they are pregnant before receiving the vaccine. At least one dose should be given. A second dose should be given a month after the first for people entering college or in the health care profession.

**Tetanus, Diphtheria, Pertussis**

Since the 1980s, the incidence of pertussis infection in adolescents and adults has been increasing. As a result of this increase, a new booster vaccine (Tdap) has been developed to include pertussis immunization with the previously standard booster shot of tetanus and diphtheria (Td). Adults between the ages of 19 and 64 should receive a Tdap vaccine every 10 years. For adults who will be in close contact with infants and for health care workers who have yet to receive this new Tdap formulation, the vaccine should be given as soon as 2 years after the last Td booster.

**Influenza**

Influenza is a viral illness that occurs yearly from November through March. The elderly, infirmed, and many people with disabilities are most susceptible to an influenza infection that could cause serious morbidity and mortality. Therefore, it is recommended that all adults 50 years of age and older; people with chronic medical conditions, including diabetes, asthma, renal insufficiency, and human immunodeficiency virus (HIV); residents of nursing homes and other chronic care facilities; and health care professionals receive a vaccination yearly in the late fall or winter.

**Pneumococcal Vaccine**

Pneumococcus is the most common pathogen responsible for pneumonia in adults of all ages. Those people most at risk for serious infection and death include the elderly, chronically ill, and immunosuppressed. Therefore, it is recommended that all adults 65 years of age and older, and people with chronic medical conditions, splenic dysfunction, asplenia, multiple myeloma or HIV, receive one dose of the pneumococcal vaccine. Those people vaccinated before the age of 65, or who are at high risk of fatal infection, should receive a second vaccination at least 5 years after the first dose.

**Human Papillomavirus**

Genital HPV is the most common sexually transmitted disease in the United States. While many infections are relatively asymptomatic, certain strains of HPV can lead to cervical cancer in women. Given this concern over cervical cancer, in March of 2007, the ACIP recommended that females between the ages of 9 and 26 receive the HPV vaccine. The vaccine is given as a series of three doses, with the second dose 2 months after the first and the third dose 6 months after the first.

**Varicella-Zoster (Shingles)**

Shingles is a painful skin rash that is caused by the varicella-zoster virus, the same virus that causes chickenpox. Shingles most commonly develops in the elderly and people who have compromised immune systems from disease or immunosuppressive medications. The ACIP recently recommended that people over the age of 60 receive one injection of the zoster vaccine.

**Screening for Cardiovascular Risk**

**Hypertension**

In addition to being a major risk factor for coronary artery disease and cerebrovascular disease, hypertension is also a major risk factor for congestive heart failure, renal disease, abdominal aortic aneurysm rupture, and retinopathy. Multiple studies over the last 30 years have demonstrated that reduction in elevated blood pressure leads to a reduction in risk of these diseases regardless of ethnicity, gender, or age (115). The current
screening recommendations from all major authorities call for routine blood pressure measurements at least once every 1 to 2 years in adults beginning between the ages of 18 to 21 (116). In 2003, The Seventh Report of the Joint National Committee on Detection, Evaluation and Treatment of High Blood Pressure (JNC VII) redefined high blood pressure categories to include a prehypertension stage of 130 to 139 systolic or 85 to 89 diastolic (117). This change was made to emphasize that blood pressure readings in the prehypertension category should not be considered normal. People with prehypertensive blood pressure readings should be counseled to lower their blood pressure through lifestyle modification, which includes weight loss, exercise, reducing sodium intake, increasing intake of fruits and vegetables, and limiting alcohol intake.

Lifestyle modification is not only recommended for the prehypertensive category but is considered the initial intervention for all elevated blood pressure readings. If lifestyle changes are not effective at lowering the blood pressure to the desired range, then medications should be started. There are many medications available for the treatment of hypertension. The JNC VII suggests that in uncomplicated stage I hypertension (systolic 140 to 159 or diastolic 90 to 99) a diuretic is most often the first medicine to use. In stage II hypertension (systolic greater than or equal to 160 or diastolic greater or equal to 100), a combination medicine that combines a diuretic with another agent is recommended as initial therapy. However, in the setting of specific clinical conditions, other choices for initial or subsequent therapy could be considered. Examples of these situations include the use of angiotensin-converting enzyme or angiotensin-receptor blockade medicines in patients with diabetes and congestive heart failure and the use of β-blockers in patients with systolic heart failure or who have suffered a myocardial infarction. A comprehensive list of such indications and medication choices is listed in the JNC VI (118).

**Diabetes Mellitus**

Diabetes mellitus is an established risk factor for both macrovascular (coronary artery disease, cerebrovascular disease) and microvascular diseases (retinopathy, nephropathy, nephropathy). Careful control of the blood sugar has been proven to reduce the development of these diseases (119,120). Despite this information regarding reduced risk of disease in the diabetic population, most major authorities, including the American College of Physicians, the American Academy of Family Physicians, and the American Diabetes Association, recommend against routine screening for diabetes among asymptomatic adults. The USPSTF finds insufficient evidence to recommend for or against routine screening. However, most authorities agree that selective screening in high-risk individuals is warranted. High-risk people would include those with obesity, hyperlipidemia, a strong family history of diabetes, or women with a history of gestational diabetes.

According to the American Diabetes Association, the best screening test for detecting diabetes is the fasting plasma glucose (FPG). A person with an FPG greater than 125 has diabetes. A person with an FPG between 100 and 125 has impaired glucose tolerance or prediabetes and is considered to be at an increased risk of developing diabetes in the future. Normoglycemia is defined as an FPG less than 100 (121).

Treatment options for diabetes, both insulin dependent and noninsulin dependent, have expanded significantly over the last several years. Oral agents are available that sensitize cells to the effects of insulin, as well as stimulate insulin secretion. Additionally, the array of insulin preparations available have increased to include longer acting agents that provide a more stable basal level, as well as agents that act quickly to provide better postprandial effects. Although a comprehensive review of these medications is beyond the scope of this chapter, the goal of treatment is to achieve a hemoglobin A1C less than 7.0% (see Chapter 65 for further discussion of the above medications).

**Hyperlipidemia**

There is a significant body of literature that demonstrates lowering cholesterol levels results in a reduction in primary cardiovascular events (122–124). As a result, most major authorities agree that people should be screened routinely for lipid abnormalities at least every 5 years. Additionally, these same authorities agree that appropriate screening entails obtaining a nonfasting total cholesterol and high-density lipoprotein (HDL) levels. However, there is disagreement as to the age to begin the screening process. The USPSTF and the American College of Physicians recommend screening men age 35 years and older and for women age 45 years and older. For people with other risk factors for coronary artery disease, these organizations recommend beginning the screening process as early as 20 years of age.

The National Cholesterol Education Panel-Adult Treatment Panel III (NCEP-ATP III) recommends that screening begin for all people at the age of 20 by obtaining a fasting lipid profile. The fasting lipid profile should include total cholesterol, low-density lipoprotein (LDL), and HDL levels. Subsequently, a person’s cardiovascular risk-factor profile should be determined to help determine their target LDL level. The NCEP-ATP III guidelines were released in 2001, with updated recommendations in 2004. Based on the updated recommendations, people with known coronary artery, peripheral vascular, or cerebrovascular diseases, as well as people with diabetes, should have their LDL value less than 70. In people with two or more cardiovascular risk factors (including hypertension, cigarette smoking, obesity, family history of premature coronary artery disease, and HDL cholesterol less than 40, a man older than 45 years of age or a woman older than 55 years of age), the LDL cholesterol level should be less than 130. However, a person with two or more risk factors that are in poor control should have the level less than 100. For people with less than two risk factors, the LDL cholesterol level should be less than 160 (125).

The initial approach to a person with elevated cholesterol is dietary modification. A diet low in fat and cholesterol should be tried for several months in attempts to
reach the target LDL level. If this level is not reached, then pharmacologic therapy should be instituted. The most common class of medications to treat elevated LDL cholesterol is the statin class. It is important to monitor the effects of these drugs on the cholesterol levels 2 to 3 months after initiating therapy or when increasing the dose. At the same time, careful monitoring of liver function tests is recommended as the statin medications have a small risk of causing elevated liver transaminases.

**Obesity**

During the past 20 years, there has been a dramatic increase in the rate of obesity in the United States. In 1987, there was no state with an obesity rate greater than 15%. However, as of 2006, there were 46 states with obesity rates greater than 20%, of which 22 states were above 25% and two states were above 30% (126). Furthermore, an estimated 66% of the adult population of the United States is either overweight or obese (127). Persons who are overweight or obese are more likely to have adult-onset diabetes, hypertension, and risk factors for other diseases (128). Observational studies have established a clear association linking overweight and obesity with hypercholesterolemia and suggest an independent relationship with coronary artery disease (129). Obesity has been associated with an increased risk of certain cancers (including those of the colon, rectum, prostate, gallbladder, biliary tract, breast, cervix, endometrium, and ovary) and with disorders such as cholelithiasis, obstructive sleep apnea, venous thromboembolism, and osteoarthritis (130). Finally, obesity can affect the quality of life by limiting mobility, physical endurance, and other functional measures, as well as through social, academic, and job discrimination. Arguably, obese people with disabilities may suffer even greater loss of mobility and endurance than obese people without disabilities. As a result of the above information linking obesity to a myriad of diseases, the USPSTF recommends that all people be screened for obesity. The most often used modality is the body mass index (BMI). The BMI is calculated through height and weight measurements. A normal BMI is between 18.5 and 25. Overweight is defined as a BMI between 25 and 29.9. Obesity is defined as a BMI greater or equal to 30.

**Cancer Screening**

**Breast Cancer**

Breast cancer is the most common malignancy in women in the United States. A women’s lifetime risk of developing breast cancer is approximately one in eight (131). The incidence of breast cancer has actually declined since 2002 presumably due to the decreased use of hormone replacement therapy (HRT) in postmenopausal women (132). From a peak of 203,500 new cases of breast cancer in women in 2002, there were only 178,480 new cases in 2007. Unfortunately, in 2007 there were still 40,460 deaths due to breast cancer. Treatment for breast cancer, especially when detected early, is effective at decreasing mortality. As a result, the ACS, American College of Radiology, American Medical Association, American College of Obstetricians and Gynecologists (ACOG), and a number of other organizations recommend screening mammography every 1 to 2 years and annual clinical breast examination beginning at the age of 40 and annual mammography and clinical breast examination beginning at age 50 (133).

**Cervical Cancer**

Approximately 11,150 new cases of cervical cancer will be diagnosed in the United States in 2007, with 3,670 deaths (134). The disease is usually preceded by an asymptomatic phase that is preinvasive and therefore ideal for early detection and treatment. Cervical cancer is caused by human papillomavirus, a sexually transmitted disease in 95% of the cases (135).

A consensus recommendation that all women who are sexually active, or who have reached the age of 18, should have annual Pap smears has been adopted by the ACS, National Cancer Institute, ACOG, American Medical Association, and the American Academy of Family Physicians. The recommendation permits Pap testing less frequently after three or more annual smears have been normal (136).

**Colorectal Cancer**

Colorectal cancer is the second leading cause of cancer mortality in the United States, with an estimated 52,180 deaths in 2007 (137). Over 90% of colorectal cancer begins as an adenomatous polyp that can often be detected through screening procedures. The ACS, the American College of Physicians, and the USPSTF agree that colon cancer screening should begin at age 50 for men and women at average risk. The ACS recommends screening with fecal occult blood testing plus sigmoidoscopy every 5 years or a colonoscopy every 10 years (138). Increasingly, physicians are recommending colonoscopy over sigmoidoscopy. For patients with a family history of colon cancer, screening should begin either 10 years before the family member was diagnosed with colon cancer or age 50, whichever age comes first.

**Prostate Cancer**

Prostate cancer is the most common cancer in men and accounts for the second highest number of cancer deaths in men (139). Public awareness of this disease and the opportunity to be screened with a blood test (prostate specific antigen or PSA) have grown tremendously in the last few decades. Unfortunately, there has been no research to date demonstrating that systematic early screening for prostate cancer with PSA or digital rectal examination would save lives. As a result, the USPSTF concludes that there is insufficient evidence to recommend for or against screening for prostate cancer at this time (140). However, the ACS and American Urological Association do recommend annual digital rectal exam (DRE) beginning at age 40, with an annual serum PSA beginning at age 50. PSA screening should begin at age 45 for those men at high risk for prostate cancer, including African American men and those with a family history of prostate cancer (141).
Chemoprevention
Aspirin Therapy
Coronary artery disease is associated with well-established risk factors, such as hypertension and smoking. In addition to the commonly discussed risk factors, coronary artery disease has also been related to platelet activity and inflammation. The fact that aspirin possesses both antiplatelet and anti-inflammatory properties has made it an interesting topic regarding coronary artery disease prevention.

There have been five major randomized trials of aspirin therapy for the primary prevention of coronary events (142–146). Four out of the five studies showed a reduction in the rates of cardiovascular events, most notably myocardial infarction. The reduction in relative risks ranged from 4% to 44%.

The USPSTF recently strongly recommends that clinicians discuss aspirin chemoprevention with adults who are at increased risk for coronary artery disease. Increased risk is considered for men older than 40, postmenopausal women, and younger people with risk factors for atherosclerotic disease. Other authorities recommend calculating the risk of coronary artery disease over 5 or 10 years and utilizing aspirin in people who reach a certain threshold of risk.

Unknown in the recommendations is the appropriate dose of aspirin to recommend. The five primary prevention studies used aspirin doses ranging from 75 to 500 mg, and there have been no studies comparing different doses of aspirin.

Postmenopausal Hormone Replacement
HRT for postmenopausal women has been a controversial topic for the last 20 years. In the 1980s and early 1990s, observational studies suggested that women who were treated with HRT had a lower risk of developing coronary heart disease, had more favorable lipid profiles, and had a better quality of life. Unfortunately, as with all observational studies, there were significant confounding factors. Some of these factors included the fact that women treated with HRT were usually healthier at baseline, more physically active, and were of higher socioeconomic status (147). Further complicating the controversy over HRT was the fact that there was an increased risk of breast cancer in women receiving HRT. The relative risk of developing breast cancer in women receiving HRT was 1.3 compared to the risk in those women who did not receive HRT (determined from meta-analysis). The Women's Health Initiative is a long-term, national health study looking at the effects of HRT, dietary modification, and calcium and vitamin D supplementation in 161,108 women between the ages of 50 and 79. The HRT arm of the trial enrolled women to receive one of three treatments: estrogen and progesterone if they had a uterus, estrogen alone if they did not have a uterus, or placebo. The study was due to be completed in 2005, but due to emerging, concerning trends, the estrogen plus progesterone and estrogen arms were stopped early in July 2002 and March 2004, respectively.

The major findings from the trial were that estrogen plus progesterone increased the risk of heart disease, breast cancer, strokes, and pulmonary emboli. There was a decreased risk of colorectal cancer and hip fractures. Overall, health risks exceeded benefits of this combination therapy in postmenopausal women (148). Concerning estrogen alone, there were increased risks of strokes and deep venous thromboses, decreased risk of hip fractures, and no clear effect on heart disease or breast cancer (149). Similar to the combination therapy, the health risks exceeded the health benefits.

Osteoporosis
Osteoporosis is characterized by low bone mass, leading to bone fragility and an increased susceptibility to fractures (150). One out of two women and one out of eight men older than the age of 50 will have an osteoporosis-related bone fracture in their lifetime. The estimated direct expenditure for osteoporosis and its related fractures is $14 billion annually (150). Specifically concerning hip fractures, there is 15% to 20% risk of dying within 1 year of suffering the fracture (151). Risk factors for osteoporosis include age, Caucasian and Asian women, small, thin body habitus, family history of osteoporosis, cigarette smoking, excessive alcohol intake, medications such as steroids and anticonvulsants, and extended bed rest or inactivity.

The diagnosis of osteoporosis is made by a bone mineral density (BMD) study. Dual energy x-ray absorptiometry (DEXA) is the BMD examination that is most reliable. Osteoporosis is defined as a BMD more than 2.5 standard deviations below the average BMD of young women, while osteopenia is defined as a BMD between 1 and 2.5 standard deviations below the BMD of young women (151).

Many people with disabilities suffer from a loss of mobility and weight-bearing activity. As a result, osteoporosis can develop at an earlier age than average. Obtaining DEXA scans and discussing prevention and treatment of osteoporosis in people with disabilities should be of primary importance.

There are several ways to help maintain a healthy BMD and prevent osteoporosis. First, it is important that people take adequate amounts of calcium and vitamin D. For most people, 1,000 to 1,200 mg of calcium a day and 400 to 800 IU of vitamin D are recommended (152). Furthermore, counseling regarding smoking cessation and moderation of alcohol consumption is important. When possible, weight-bearing exercise should be encouraged.

Medications are available to treat osteoporosis. The most common class of medications used is bisphosphonates. Bisphosphonates work by inhibiting osteoclast-mediated bone resorption. Recently, there has been concern that these medications may cause osteonecrosis of the jaw. However, a recent review of all available studies contends that the risk remains extremely low (153).

Low Back Injuries
A leading cause of missed days in the workplace, acute low back pain has not yet received the scientific inquiry and educational priority that it deserves. A report from the North Carolina Back Pain Project revealed that treatment from family physicians, orthopedists, and chiropractors had equal success in resolving uncomplicated acute low back pain (154). However,
ambulatory patients surveyed in the study were significantly more satisfied with the care they received from chiropractors, as they reported more satisfaction with the chiropractic clinician’s examination and patient education efforts. This study’s conclusions should not be surprising because undergraduate and resident medical education programs do not always adequately stress the musculoskeletal examination and even more importantly, patient education and prevention techniques. The AHCPR published clinical guidelines for the evaluation and management of acute low back pain (155). Authorities propose that prevention programs emphasizing weight reduction, proper lifting techniques, ergonomic design of the home and workplace, and strengthening and flexibility enhancement will reduce the incidence of this common condition. However, a more recent update by the USPSTF found no new good evidence for or against the use of back-strengthening exercises or risk factor modification (e.g., increased physical activity, smoking cessation, or reduced alcohol consumption) for the primary prevention of low back pain in adults. There was limited evidence that educational sessions in occupational settings (e.g., back schools) produce modest short-term benefits in adults with recurrent or chronic low back pain, but no evidence that such education prevents back pain in healthy individuals or those at risk for back pain (156).

**Overuse Injuries**

Sports-related and work-related overuse injuries present principally in ambulatory medical settings. Therefore, primary medical care providers should be familiar with their evaluation and management.

**Sports**

Exercise is an important risk reducer for the chronic diseases mentioned above. Providers should know how to manage the adverse reactions (overuse injuries) of an exercise prescription, just as they should know how to manage the adverse reactions of medication. Common running and walking overuse injuries include metatarsal and tibial stress fractures, metatarsalgia, Morton neuroma, shin splints, patellofemoral pain syndrome, infrapatellar tendinitis, Achilles tendinitis, pes anserine bursitis, and iliotibial band friction syndrome. Although overuse injuries of the lower limbs are usually precipitated by improper training techniques and excessive exercise duration or intensity, sports medicine practitioners now know that the root cause is often a biomechanical imbalance such as overpronation and inflexibility (157). Thus, a working knowledge of strength and flexibility assessment, biomechanical foot disorders, and proper athletic shoe prescription is an essential requirement of the modern primary care provider.

**Repetitive Injuries in the Workplace**

Because primary care providers aim to view patients in the full context of their home and work, occupational medicine issues are an integral component of comprehensive health care. For example, our “keyboard society” has hastened a dramatic increase in the reporting of carpal tunnel syndrome cases. Primary care providers must become well versed in the assessment, conservative treatment, and knowledgeable referral of patients with carpal tunnel syndrome and other work-related repetitive injuries.

**Trauma-Related Injuries**

**Motor Vehicle Injuries**

In 2006, motor vehicle crash (MVC)-related injuries and other unintentional injuries (MVC was number one among this group) were the fifth leading cause of death in the United States (158). In 2006, the cost of unintentional injuries totaled $652.1 billion, with motor vehicle injuries being responsible for almost half of this cost (158). Motor vehicle fatality rates are highest for young and elderly adults, whereas injury rates peak in young adulthood (158,159). Although alcohol-related traffic fatality rates have declined by more than one third since 1979 (159), alcohol use remains an important risk factor for motor vehicle injuries. In drivers 15 to 20 years of age involved in fatal crashes between 1992 and 2002, the incidence of intoxication dropped by 7%. Still, 27% of the young male drivers (11% of females) involved in fatal crashes had been drinking at the time of the crash. Evidence also indicates that impairment with drugs other than alcohol also may play an important role in traffic injuries and deaths, although the relationship is not as well-defined as for alcohol. In addition to driving while impaired by alcohol or drugs, failing to use occupant protection (e.g., safety belts, child safety seats, motorcycle helmets) is also an important risk factor for motor vehicle injury (159,160). Substance abuse screening and intervention are likely to be efficacious in reducing motor vehicle injuries and fatalities, whereas the use of occupant restraints has been proven to reduce the risk of motor vehicle injury and death (161).

Recommendations (161) from the USPSTF for the prevention of motor vehicle–related injuries are listed as follows:

1. Clinicians should regularly urge their patients to use lap and shoulder belts for themselves and their passengers while riding in automobiles, including automobiles equipped with air bags.
2. Operators of vehicles carrying infants and toddlers should be urged to install and regularly use federally approved child safety seats in accordance with the manufacturer’s instructions and the child’s size.
3. Those who operate or ride on motorcycles should be counseled to wear approved safety helmets.
4. All patients should be counseled regarding the dangers of operating a motor vehicle while under the influence of alcohol or other drugs, as well as the risks of riding in a vehicle operated by someone who is under the influence of these substances.

**Household and Recreational Injuries**

Unintentional injuries accounted for nearly 120,000 deaths in the United States in 2006, more than half of these being related to household, recreational, and other etiologies.
unrelated to motor vehicles (158). Falls, poisoning, choking/aspiration, fires and burns, drowning, mechanical suffocation, and firearms caused nearly 50% of these deaths (158). Almost 90% of deaths relating to sports and recreation occur during swimming, boating, bicycling, riding off-road vehicles such as all-terrain vehicles, or using firearms (158). All of the above are also common causes of nonfatal injuries, with falls being the most common (161).

Recommendations (161) from the USPSTF for the prevention of household and recreation related injuries are listed as follows:

1. Parents should be counseled on measures to reduce the risk of unintentional injuries to their children from residential fires and hot tap water, drowning, poisoning, bicycling, firearms, and falls.
2. Homeowners should install smoke detectors and carbon monoxide detectors in appropriate locations and test the devices periodically to ensure proper operations. Encourage the use of flame-resistant nightwear during sleep and reduce or cease smoking in the home.
3. Households are advised to keep a 1-ounce bottle of ipecac, display the telephone number of the local poison control center, and place all medications, toxic substances, and matches in child-resistant containers.
4. Bicyclists and parents of children who ride bicycles should be counseled about the importance of wearing approved safety helmets and avoiding riding bicycles in motor vehicle traffic.
5. Families should be encouraged to install fences with gates around swimming pools.
6. All windows that pose high risk for falls should have window guards.
7. All residents of homes with swimming pools, young children, or elderly persons should be encouraged to learn cardiopulmonary resuscitation and maneuvers to manage choking incidents.
8. All firearms should be removed from the home or at a minimum kept unloaded in a locked compartment separated from the ammunition.
9. Elderly patients should be counseled on measures to reduce the risk of falling, including exercise (particularly training to improve balance), safety-related skills and behaviors, and environmental hazard reduction, along with monitoring and adjusting medications.

**Organization of the Medical Care Team**

Managed care has dramatically and irrevocably changed the delivery of health care in the United States. Managed care organizations now carefully define the roles of their primary care providers and specialty providers. Powerful practice management computer databases are being developed that track patient populations, case mix and levels of service acuity, resource utilization, prescribing patterns, referral patterns, individual practice profiles, and treatment outcomes. Individual practices, regional health care systems, and Fortune 500 health care conglomerates are implementing continuous quality improvement programs.

A concept in managed care delivery that depends on continuous quality improvement for its success is called disease state management (162). This health care organization model targets high-cost and high-volume diagnoses and those conditions that tend to have wide variations in provider practices and resource utilization. The intention of disease-state management is to improve outcomes while maintaining or lowering overall costs. The process of disease-state management must include evidence-based clinical policies, an accurate and complete database that can measure the well-defined outcomes, and a team-oriented, multidisciplinary approach. The multidisciplinary approach will require careful and extensive communication between generalist and specialist members of the team (163). Advocates of this organizational system recognize that its success critically depends on a sophisticated information system and the rapid development of evidence-based clinical policies that focus on patient outcomes. Unfortunately, outcome studies in primary care need further development. Thus, outcome-based studies will be important research objectives for primary care scholarship.

**Family and Community Issues**

Primary care providers assess and treat patients within the context of their family and community. Indeed, comprehensive longitudinal care is incomplete and ineffective without addressing the meshwork of stresses and supports that surround individual patients (162,164). Quality primary care addresses and explores the following family/community issues:

- The family circle
- Patient birth order
- Family genogram
- Childhood traumas, expressed and suppressed
- Family violence
- Marital history and marital stresses
- Chronic illness in other family members
- Family organization and identity (chaotic? crisis-oriented? nurturing? etc.)
- Educational level of the patient
- Educational resources available for the patient and family
- Child day care services
- Health belief models expressed by the patient and family members
- Timing of illnesses in a patient’s life-cycle stage
- Patient’s work/employment history
- Co-worker stress/support
- Patient hobbies and stress relievers
- Patient’s pets
- Home safety
- Neighborhood safety
- Family and nonfamily members living in the same household
- Transportation availability
- Services and support for persons with disabilities
Model Programs for Provision of Primary Care to Persons with Disabilities

The need to provide primary care services to persons with disabilities is obvious, yet the process by which care can be provided remains unclear. Given that barriers to care exist, some models of care may be more beneficial to persons with disabilities seeking primary care services. Pediatric programs that provide primary and rehabilitation care to children with disabilities are well established; however, innovative models of care for adults are not currently available in most communities.

Barriers to Primary Care Services

Persons with disabilities experience many barriers to receiving primary care services. Physical barriers to care include problems with transportation and inaccessible parking lots, offices, and exam rooms (23,30,165). Those with physical disabilities have also described difficulty transferring to examination tables, as well as inaccessible equipment for routine tests such as mammograms or pelvic exams (166). Attitudinal barriers to care have been described, as persons with disabilities have reported instances of providers refusing to provide medical care due to unfounded disability concerns or have felt that primary care physicians lack adequate knowledge of their specific disabilities (165,167).

In many cases, physicians themselves impose barriers to care. Many physiatrists, while comfortable treating rehabilitation problems, are not comfortable providing primary care services (51). Beyond the intern year, most PM&R residency programs do not include primary care or basic internal medicine topics in their curriculum. Similarly, general practitioners, who are trained to provide preventive care and to treat acute medical problems, are not adequately trained in how to provide care to persons with disabilities (11,168). Low reimbursement rates for longer office visits amplify a negative stigma, which is often unfairly attached to persons with disabilities (23). Thus, primary care providers who would otherwise be willing to treat persons with disabilities are often hindered by both their lack of training and the cost of providing care. The following will describe four models of primary care for persons with disabilities.

General Practitioner Model

In a traditional primary care setting, a family medicine physician, internist, or subspecialist physician assumes responsibility for all primary care issues. This model is most commonly used in cases of spinal cord injury or traumatic brain injury, particularly in younger patients who did not have a primary care physician at the time of injury (11). Under the physiatrist model of care, the physiatrist treats all chronic problems and acute medical issues, in addition to any rehabilitation-specific conditions. Due to lack of support staff and difficulty making expedited appointments, patients may utilize emergency room services for urgent medical problems. All routine screening and health maintenance care become the responsibility of the physiatrist, thus potentially leading to unmet needs as many rehabilitation providers do not have
the time, knowledge, staff, or resources to provide necessary preventive services (11,51).

An example of a successful program where physiatrists provide quality comprehensive care is within the Department of Veterans Affairs (VA) Hospital System (170). Persons with spinal cord injuries receive acute, chronic, and preventive care services from physicians trained in spinal cord injury. Comprehensive care is provided to persons with spinal cord injury in one of 23 regional centers. The spinal cord injury specialist runs a multidisciplinary team that coordinates annual comprehensive evaluations. All inpatient hospital admissions are directed to the spinal cord injury service regardless of the medical problem. Even after undergoing surgery, patients are transferred to the spinal cord injury floor for medical care. Preventive health care, education, emergency services, subspecialty care, and long-term care are all provided through the VA. Additionally, all necessary equipment, vocational programming, social services, and psychological services are offered throughout the veteran’s lifetime.

**Collaborative Practice Model**

Collaborative practices allow persons with disabilities the opportunity to receive all necessary medical care from multiple medical personnel in a single office setting. Access to primary care, gynecology, urology, and rehabilitation services is made available within the same practice. Medical information is typically communicated in one chart, and patients rarely require referrals outside of the practice. In some programs, patients may receive treatment for all acute problems including urgent and emergency care and inpatient hospital services. Additionally, routine preventive services are provided and chronic conditions are monitored.

The initial concept for a collaborative practice model was developed at the Rehabilitation Institute of Michigan, where persons with disabilities on Medicaid could receive comprehensive care (30,171). The basic team consisted of an internist, physiatrist, and nurse. Specialty physicians such as urologists and surgeons were consulted when needed, and all visits were provided on-site. Urgent care services were provided in clinic, and an answering service was available for off-hours. Any urgent needs were coordinated with the emergency room by the interdisciplinary team. Unfortunately, the program was forced to close when Michigan Medicaid shifted from a fee for service plan to a capitated managed care organization.

Rancho Los Amigos Rehabilitation Center in California is another example of a collaborative practice model, providing primary care to persons who have had spinal cord injuries or strokes (172). At present, a complex system of outpatient clinics provides primary care services, rehabilitation services, annual exams, and subspecialty follow-up at the rehabilitation center. A physiatrist transitions patients from inpatient rehabilitation to any pertinent outpatient clinics. On initial follow-up after discharge from inpatient services, patients are evaluated by therapists in the outpatient clinic as part of care plan development. An internist or nurse practitioner staffs specialized primary care clinics. Clinics for wound management, pulmonary care, gynecology, orthopedics, urologic care, and spasticity are available in the same building, and all medical records are kept in a single outpatient chart to facilitate information sharing. Although primary care services are predominately used for persons with spinal cord injury, if needed, those with stroke or traumatic brain injury may also receive specialized primary care.

Currently, the Rehabilitation Institute of Chicago (RIC) offers a comprehensive program for primary and rehabilitative care (173). Primary care clinics are available 4 days a week and are staffed by two internists. Patients discharged from RIC with primary care needs may be followed up in the clinic for all primary care services and preventive care. A women’s gynecology clinic is also available. One of the major goals of the program is to prevent secondary health conditions and functional decline. Services from therapists, physicians, and allied health professionals combine to provide comprehensive health care, and the same chart is used for all clinic and therapy notes.

**Coordinated Care Model**

Previously, separate health care coordination and home-care models of primary care have been described. At its inception, the health care coordination model emphasized the use of nurse health coordinators who performed initial health assessments; developed patient-centered plans of care, coordinated with the primary care physicians and specialists; and provided 24-hour urgent care assistance. The nurse health coordinator facilitated all communication between providers and helped to integrate all services. Similarly, the home-care model described comprehensive patient care based out of the participant’s home.

More recently, the health care coordination and home-care models have essentially been combined into a model of coordinated care (174). Disability coordination care organizations have been described as a combination of a home-based model and community nursing services for people with long-term health care needs. A common theme to all coordinated care programs is that of patient autonomy and patient-directed care. Most programs permit medical care to be provided in an outpatient clinical setting or in a participant’s home. A nurse or social worker serves as a case manager to coordinate all needed medical and social services. In many instances, case managers will make patient appointments, arrange for transportation to visits, and attend appointments if needed. At the time of enrollment, each participant receives an extensive medical and social evaluation. Although care focuses on treatment and prevention of medical problems, an emphasis is also placed on provision of any services that might enhance quality of life and functional independence.

Most coordinated care models operate under a capitated prepaid managed care system with state Medicaid agencies (175). Some programs also enroll Medicare patients. The goal of disability coordination care organizations is to reduce costs while providing better clinical outcomes. With increased preventive services, the number of inpatient admissions and emergency room visits diminish (174). Start-up costs for coordinated care programs are often high because participants
may have significant unmet need at the time of enrollment. However, after services have been established and medical problems stabilize, costs tend to be contained (174). Some programs receive additional grant support from outside sponsors such as the Robert Wood Johnson Foundation.

The first program of this type originated from the Urban Medical Group in Boston, which provided primary care to inner-city nursing home residents and frail homebound elderly persons (30,174–176). In the 1980s, the Boston Center for Independent Living and the Urban Medical Group created Boston’s Community Medical Group with the intent to provide primary care services to persons with spinal cord injury or advanced HIV disease. Currently, Boston’s Community Medical Group provides primary care to persons with disabilities as part of the Commonwealth Care Alliance, and approximately 450 patients are enrolled. At the core of the program are nurse practitioners and physician assistants, who serve as gatekeepers to care after initial evaluations are completed. An individualized comprehensive medical plan is provided to participants in their homes, clinics, long-term care facilities, or hospitals. The physician extenders promptly evaluate new medical problems and provide 24-hour emergency assistance. Interdisciplinary meetings with the physician extender, primary care physician, social worker, therapists, and behavioral health specialist lead to coordination of services. The initial program started as a Medicaid fee-for-service program but later converted to a capitated prepaid managed care organization in conjunction with Massachusetts Medicaid. Participants may be dually enrolled in Medicare.

The Wisconsin Partnership Program encompasses separate organizations throughout Wisconsin under a Medicare and Medicaid capitated program (174,177). Some of the Partnership organizations provide interdisciplinary care to persons with disabilities less than 65 years of age. Most enrollees are at a Medicaid nursing home level of care. A nurse practitioner coordinates care with the primary care provider, nurse, social worker, and other team members. Medical care, long-term care, and community-based services are provided in participants’ homes or in the setting of their choice. Emergency care, transportation, and personal care services are also covered. Emphasis is placed on preventive services and patient education. Patient autonomy is an important component of care, and patients are encouraged to choose participating physicians and to direct personal care services. A plan is in place to expand programs so that all sites will be able to provide services to persons with physical disabilities and to those with developmental disorders. As of January 2008, almost 2,700 participants were enrolled in the Partnership program.

The AXIS Health care program in Minnesota is a nontraditional Medicaid managed care program (174,178). It was created in 1997 through a grant from the Center for Health Care Strategies, to help find an innovative way to provide quality, cost-effective care to persons with disabilities. Similar to other coordinated care models, nurses serve as case managers and perform complex assessments on all patients as they enter the program. Nurses are on call 24 hours a day for assistance with medical issues, and all patient information is available in a portable medical record. Participants are referred to their primary care physicians and specialists as needed, and nurse practitioners are available for episodic care. All program providers have experience in treating persons with disabilities. Care can be received at home, in nursing facilities, or in outpatient clinics. The program provides participants with all needed health care and support services, including physical and occupational therapy, mental health, and substance abuse treatment. Presently 1,100 persons with disabilities are enrolled.

CONCLUSION

All of us, including persons with disabilities, want the same things for ourselves and our families: a compassionate, knowledgeable, and available physician to help us maintain our health when we are well and to manage our minor illnesses when we are sick, and the best specialists available when we have a major acute or chronic medical problem (11). In this chapter, we have reviewed some of the obstacles that those with disabilities endure to achieve the above ideal. It is not expected that the field of PM&R will be declared a primary care specialty nor would this be a viable solution to providing primary care services for all of those with disabilities. However, one of the goals of writing this chapter was to provide those physiatrists who currently perform continuity of care and gatekeeper functions with some guidelines, as well as resources, to continue to provide state-of-the-art primary care. The majority of practicing physiatrists, however, do not fulfill this role.

Our health care system is changing rapidly, and in order to assist our patients with disabilities to achieve optimal health, we need to frequently reevaluate and modify our methods of health care delivery. The last section of this chapter provided alternatives and models for the provision of primary care services in this population. Some of the programs have become very successful within their own health care systems but may not directly apply to all practitioners interested in providing these services. At a minimum, rehabilitation care providers should collaborate with existing primary care providers, including generalist physicians and physician extenders, to achieve the goal of providing adequate, high-quality primary care for persons with disabilities. This collaboration could include educating primary care providers on psychiatric issues as well as team ventures with them to provide primary care for this population.

REFERENCES

38. Pender N. *Health Promotion in Nursing Practice.* Norwalk, CT: Appleton & Lange; 1982.


166. Smeltzer SC. Preventive health screening for breast and cervical cancer and osteoporosis in women with physical disabilities. Fam Community Health. 2006;29(15):355-43S.


Management Methods
The term **therapeutic** relates to the treatment of disease or physical disorder, and **exercise** refers to bodily exertion for the sake of training or improvement of health. This chapter on therapeutic exercise therefore addresses the use of activities requiring physical exertion in the prevention, treatment, and rehabilitation of illness and disabling conditions. The pertinent exercises considered in this chapter include those to develop endurance, strength, flexibility, and proprioception.

The use of therapeutic exercise in the treatment of injuries is not a new concept. Hippocrates (460 to 370 B.C.) reportedly advocated exercise as an important factor in the healing of injured ligaments, and the Hindus and Chinese used therapeutic exercise in the treatment of athletic injuries as early as 1,000 B.C. Today, the various types of exercise probably account for the most commonly used treatments in the field of physical medicine and rehabilitation. Therefore, it is important for clinicians in the field to have a thorough understanding of this area.

The concepts forming the basis for therapeutic exercise come from studies in basic physiologic science and applied exercise physiology. In recent years, epidemiologic investigations have provided additional insight into the importance of exercise in prevention of disease. Consequently, much of this chapter is devoted to this basic information that provides the foundation for the clinical use of exercise.

### METABOLIC FUNDAMENTALS

Part of adaptation to regular exercise is the development of the energy systems most involved in the type of training being used. Therefore, to condition an individual optimally for participation in a particular activity or sport, an exercise program should be designed to increase the physiologic capacity of the energy systems most important to that activity. For this reason, it is valuable to have an understanding of how energy is derived for muscular contraction.

The immediate usable form of chemical energy for all muscular contraction is adenosine triphosphate (ATP). ATP is supplied to the muscle through three systems: (a) the adenosine triphosphate–creatine phosphate (ATP-CP) system, (b) the anaerobic glycolysis system, and (c) the aerobic system. The relative utilization of the three systems depends on the intensity and duration of the exercise (Table 61-1) (1).

### Anaerobic Metabolism

Anaerobic metabolism refers to a series of chemical reactions that do not require the presence of oxygen. Two of the systems that supply energy for muscular contraction are anaerobic.

### The ATP-CP System

ATP and CP are high-energy phosphagen compounds stored within the muscle and ready for immediate use. The breakdown of ATP produces adenosine diphosphate (ADP), inorganic phosphate, and energy used in muscular contraction. CP is broken down to create energy that is used to reform ATP.

This system provides an immediate source of energy for the muscle and has a large power capacity. In other words, a large amount of energy per unit time can be supplied through this system. However, because of the small stores of ATP and CP, the total capacity for work with the ATP-CP system is limited. In fact, the energy resources from the ATP-CP system will be exhausted in 30 seconds or less during an all-out bout of exercise (2,3).

### Anaerobic Glycolysis System

Glycolysis refers to a series of reactions resulting in the breakdown of carbohydrate into pyruvate or lactate. Anaerobic glycolysis means that this breakdown of carbohydrate is performed in the absence of oxygen.

During maximal exercise that lasts 1 to 2 minutes, lactic acid produced by the skeletal muscles accumulates in the muscles and blood. This is accompanied by an increase in the proton release causing acidosis. Interestingly, lactate acts as a solid indicator of the changes in the acid-base balance but is not the cause of the acidosis. In fact, lactate acts as a buffer for the acidosis that would otherwise have a more rapid onset and produce fatigue even more quickly (4). At any extent, when the concentration of lactic acid is high enough, nerve endings are stimulated, resulting in the sensation of pain. In addition, the lactic acid within the muscle cell inhibits the production of more ATP (5) and the binding of calcium to troponin (6), which is part of the series of events leading to muscle contraction. Therefore, the amount of energy obtained from the anaerobic glycolysis system is limited by these effects. Nevertheless, the anaerobic glycolysis system is extremely important because it can provide a rapid supply of energy.
## TABLE 61.1 Characteristics of the Three Metabolic Systems

<table>
<thead>
<tr>
<th>Metabolic System</th>
<th>Substrate (Fuel)</th>
<th>Oxygen Required</th>
<th>Speed of ATP Mobilization</th>
<th>Total ATP Production Capacity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anaerobic metabolism</td>
<td>ATP—CP system Stored phosphagens</td>
<td>No</td>
<td>Very fast</td>
<td>Very limited</td>
</tr>
<tr>
<td>Glycolysis</td>
<td>Glycogen/glucose</td>
<td>No</td>
<td>Fast</td>
<td>Limited</td>
</tr>
<tr>
<td>Aerobic metabolism</td>
<td>Glycogen/glucose, fats</td>
<td>Yes</td>
<td>Slow</td>
<td>Virtually unlimited</td>
</tr>
</tbody>
</table>


### Aerobic Metabolism

In the presence of oxygen, glycolysis produces pyruvate, which is further metabolized through the tricarboxylic acid (TCA) cycle (also known as the *Krebs* or *citric acid cycle*) and electron transport system to yield carbon dioxide, water, and energy. Relative to anaerobic glycolysis, the energy produced from a given amount of carbohydrate is about 13 times greater through aerobic metabolism. Furthermore, there are no fatiguing or painful byproducts through aerobic metabolism, and not only carbohydrate but fats and proteins may be metabolized aerobically. Although the ability to metabolize fat means that this system provides a virtually unlimited source of energy, aerobic metabolism provides energy at the slowest rate of the three energy systems.

### The Energy Continuum

All three energy systems supply a portion of energy to the body at all times. However, one energy system may predominate during a particular activity. Which energy systems are predominant during a given activity depends on the rate of energy (power) requirement during the activity (Fig. 61-1) (7). In activities performed at maximal intensity for only a few seconds, most of the ATP is supplied by the ATP—CP system. Activities of lower intensity, such as those at a maximal effort that can be sustained for 1 to 2 minutes, primarily rely on the anaerobic glycolysis system. Longer-duration, lower-intensity activities that may last several minutes or hours are supplied almost entirely through aerobic metabolism.

During a graded exercise test, blood lactate concentration remains relatively constant until a critical exercise intensity is reached, at which time lactate begins to accumulate in the blood (Fig. 61-2) (8). This accumulation does not indicate the onset of anaerobic glycolysis; rather, it is the result of the rate of lactate production exceeding its rate of removal.

### Fuel Utilization

The three fuels that may be used to generate ATP for muscular contraction are carbohydrate, fat, and protein. These fuels differ in the amount of oxygen required for metabolism to the end products of carbon dioxide and water (Table 61-2). Thus, the amount of carbon dioxide produced relative to the amount of oxygen used (the respiratory quotient [RQ]) differs among fuels. During exercise, this ratio is referred to as the *respiratory exchange ratio* (RER) rather than the RQ. This distinction is made because the rate of carbon dioxide exhalation increases out of proportion to metabolism whenever metabolic acidosis exists, as occurs at high exercise intensities. Metabolic acidosis is buffered in the blood by the bicarbonate buffering system with nonmetabolic carbon dioxide produced as a byproduct and in the muscle cell by the phosphate system. Therefore, at maximal exertion during dynamic exercise, an RER greater than 1.1 is expected. In fact, an RER greater than 1.1 is frequently used as a criterion to assess whether maximal exertion was achieved with graded dynamic exercise testing.

Because fat serves as the primary form of stored energy in the body, it is fortunate that it has a caloric density that is much higher than carbohydrate (9.3 kcal/g, compared with 4.1 kcal/g). However, slightly less energy (4.7 kcal/L, compared with 5.0 kcal/L) is produced from every liter of oxygen when fat is used than when carbohydrate is used. Amino acids can also be metabolized to produce energy for muscular actions, although this contribution is generally 10% or less.

The fuel used during exercise is influenced by several factors, including the exercise intensity and duration, pre-exercise...
diet, mode of exercise, and level of fitness. As the intensity of the exercise is increased, the predominant fuel source shifts toward carbohydrate. This is partly because ATP production shifts toward anaerobic metabolism during high-intensity exercise and carbohydrate is the only fuel available for anaerobic glycolysis. Carbohydrates are made available to the contracting muscle through mobilization of muscle and liver glycogen stores as well as through ingested carbohydrates that are circulating in the bloodstream.

Exercise duration also has an effect on the fuel utilization pattern. Fat usage gradually increases during long bouts of exercise. Free fatty acids are made available to the contracting muscle through lipolysis of triglycerides within extramuscular (e.g., adipose) and muscular stores. The type of diet consumed before exercise affects fuel utilization during exercise. During prolonged exercise, carbohydrate is more likely to be used when one has been eating a diet rich in carbohydrates (Fig. 61-3) (9).

The mode of exercise can also influence the fuel source used during exercise. Exercise that localizes the work to a small muscle mass will tend to use a greater proportion of carbohydrate as fuel.

Finally, training status can affect the composition of the fuel used during exercise. Adaptations to endurance training increase the ability of muscles to use free fatty acids as a fuel while sparing glycogen (10,11).

<table>
<thead>
<tr>
<th>Fuel</th>
<th>Energy Content (kcal/g)</th>
<th>Oxygen Equivalent (kcal/L)</th>
<th>Respiratory Quotient (RQ)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carbohydrate</td>
<td>4.1</td>
<td>5.0</td>
<td>1.00</td>
</tr>
<tr>
<td>Fat</td>
<td>9.3</td>
<td>4.7</td>
<td>0.70</td>
</tr>
<tr>
<td>Protein</td>
<td>4.3</td>
<td>4.4</td>
<td>0.80</td>
</tr>
</tbody>
</table>
Oxygen Uptake

Oxygen uptake ($V_{O2}$) is a measure of the rate of oxygen utilization for the production of energy. This measure is typically reported in units of liters per minute (L/min) or milliliters per kilogram body mass per minute (mL/kg/min). The latter provides for better comparison among individuals. Another commonly used measurement unit for oxygen uptake is the metabolic equivalent (MET). One MET is equal to 3.5 mL/kg/min, which is approximately the resting metabolic rate. Maximal oxygen uptake ($V_{O2}\text{max}$) or maximal METs represents the maximal rate at which an individual can use oxygen. Classically, $V_{O2}\text{max}$ is defined as the rate of oxygen uptake at which no further increase occurs despite an increase in dynamic work rate by the individual (see Fig. 61-2). The term peak $V_{O2}$ is often referred to when it is recognized that the highest attainable $V_{O2}$ is may not have been reached by the individual because of the mode of exercise being used, the testing protocol, physical limitations other than cardiorespiratory, or inadequate motivation. Clinically, assessment of peak or maximal $V_{O2}$ (mL/kg/min or MET units) is useful in assessing cardiorespiratory fitness as well as in counseling patients on tolerance for undertaking home, leisure, or vocational activities within a reasonable work intensity.

MUSCLE PHYSIOLOGY

Structure and Function of Muscle

Morphology

Skeletal muscle is made up of structural and functional subunits as displayed in Figure 61-4 (12). The largest subunit of a muscle, from a morphologic standpoint, is the fascicle. The fascicle contains anywhere from one to hundreds of muscle fibers, the individual muscle cells. At each of these structural levels is a different connective-tissue covering.

Myofibrils are found within the muscle cells. The smallest functional subunit of the myofibril is the sarcomere. Sarcomeres are aligned end to end to form a myofibril. Myofibrils contain two basic protein filaments, a thicker one, called myosin, and a thinner one, called actin. These proteins are arranged in such a way as to give skeletal muscle its striated appearance. Sarcomeres run from one Z-line to the next Z-line.

![FIGURE 61-4. Structural and functional subunits of skeletal muscle. (With permission from Lamb DR. Physiology of Exercise: Responses and Adaptations. 2nd ed. New York: Macmillan; 1984.)](image-url)
Sliding-Filament Theory
The primary function of muscle is to shorten and develop tension. The sliding-filament theory (Fig. 61-5) (13) provides an explanation of how muscle fibers shorten and so develop tension.

On stimulation from the motor axon, calcium ions are released from storage in the sarcoplasmic reticulum, exposing active binding sites on the actin and allowing actin-myosin cross-bridge formation. Once the actin-myosin bonds are formed, a conformational change in the angle of the myosin head occurs, causing the actin filaments to be pulled over the myosin filaments and to shorten the sarcomere.

For more shortening to occur, ATP is required to break the actin-myosin bond and allow binding of myosin to another actin site closer to the Z-line. The conformation change of the myosin head that is essentially the enzyme myosin ATPase then occurs, resulting in further shortening.

Relaxation of the muscle occurs when stimulation ends. This triggers the active pumping of calcium back into the sarcoplasmic reticulum. The result is a loss of active binding sites on the actin, so that actin-myosin crossbridges are broken, and the muscle relaxes.

Mechanical Model
Figure 61-6 shows a useful model to understand the mechanical properties of muscles (14). This model consists of the contractile element and two elastic elements. (From Roberts TDM. Neurophysiology of Postural Mechanics. London: Butterworth; 1978.)

FIGURE 61-5. Postulated mechanism of crossbridge formation and cycling by actin and myosin filaments. The head of the myosin filament is stippled and forms the crossbridge between the two filaments. Active sites on the actin (thin) filament are outlined by broken circles. 1: No bonds between filaments. 2: Initial attachment of myosin head to one of the active sites on the actin filament. This attachment takes place only in the presence of Ca\(^{2+}\) ions. 3: Formation of strong actin-myosin bond. This process causes a conformational change in the angle of the myosin head, which produces a relative movement of actin and myosin filaments across one another. ADP and phosphate ions are lost from the myosin. 4: If ATP is available, the actin-myosin bond can be broken. Subsequent hydrolysis of the ATP by the myosin ATPase returns the cycle to stage 1. If no ATP is available, the actin-myosin bond remains intact, as in muscular rigor. (From Gordon AM. Muscle. In: Ruch T, Patton H, eds. Physiology and Biophysics. Vol. IV. Philadelphia, PA: WB Saunders; 1982:170–260. Reprinted by permission of WB Saunders Co.)
contractile element with both series and parallel elastic elements. The contractile element actively generates force and represents the interaction between actin and myosin filaments. The elastic elements are purely passive components acting as mechanical springs. The series elastic element represents the tendinous insertions of muscle, and the parallel elastic element represents the connective tissue surrounding the various sub-units of the muscle.

Muscle Fiber Types
Skeletal muscles contain a mixture of muscle fiber types that can be distinguished by their physical and biochemical characteristics. At present, the most often-used classification system is the myosin ATPase histochemical profile for muscle. This array of fiber types allows classification of muscle fibers along a continuum from the most oxidative fiber to the least oxidative phenotype (Table 61-3). Muscle fiber types can be identified using histochemical or gel electrophoresis methods. Histochemical staining of the ATPases is used to separate fibers into type I (slow twitch), type IIA (fast twitch), and type IIB (fast twitch) forms. However, the work of Smerdu et al. (15) indicates that type IIB fibers contain type IIX myosin heavy chain (gel electrophoresis fiber typing). For the sake of continuity, and to decrease confusion on this point, most journals and scientists have now adopted the term IIX to designate IIB fibers. While type IIB muscle fibers are found in humans, they are very rare and are part of the rodent fiber array. Thus, a continuum of muscle fibers exists from the most oxidative type I to the least oxidative type IIX as shown in Table 61-3. A motor unit typically contains type I or type II muscle fibers, thereby giving it specific characteristics important to recruitment order and “size principle” whereby motor units are typically recruited from the low threshold to high threshold motor units that is related to the force production required by the external demands of the task.

Recognition that there may be differences in muscles dates back to observations that fowl has meat that is “red” and “white.” It has since been learned that the different properties of muscles extend beyond the muscle fiber to the level of the motor unit. Based on contraction speed following a stimulus to the nerve axon, two main categories of motor units can be distinguished—fast units and slow units. Peak tension and relaxation are achieved more rapidly for fast units than for slow units (Fig. 61-7). Units with a fast contraction time are composed of muscle fibers with relatively large fiber diameters and are innervated by large, fast-conducting motor neurons. These units also have a larger number of muscle fibers. Compared with the slow units, the fast units produce higher tensions. These fast units are further subdivided into fast-fatiguable (FF) units, which fatigue relatively easily, and fast-fatigue-resistant (FR) units, which have a high resistance to fatigue, similar to the slow units.

Although the number of fibers per unit and the fiber diameter are larger with fast units, these factors do not fully explain the differences in force production of fast and slow units. It is thought that there are also differences in the mechanism of force production within the fibrils (16).

All fibers from a given motor unit have the same histochemical characteristics. Slow units innervate type I slow twitch fibers and these fibers are characterized by high activities of succinic dehydrogenase and nicotinamide adenine dinucleotide dehydrogenase, enzymes involved in the major pathways for oxidative metabolism. The type I fibers are also richly supplied by capillaries. As a result, these fibers are well suited for performance of low-intensity, long-duration activities.

Those fibers innervated by fast motor units (fast-twitch or type II fibers) have a high activity of myofibrillar ATPase, the enzyme that breaks down ATP to release energy for contraction. They also have a high capacity for anaerobic metabolism as demonstrated by the high levels of glycogen and

<table>
<thead>
<tr>
<th>Type I</th>
<th>IC</th>
<th>IIC</th>
<th>IIA</th>
<th>IIX</th>
</tr>
</thead>
<tbody>
<tr>
<td>Highly oxidative</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High endurance</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High mitochondrial density</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low contractile speed</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Least oxidative
High force
Lower mitochondrial density
High contractile speed

FIGURE 61-7. Schematic representation of rate of tension generation and force production with different fiber types. The speed of contraction and force are greatest for type II fibers.
phosphorylase. Phosphorylase is an enzyme that is involved in the breakdown of glycogen. As a result, these fibers are suited for performance of high-intensity, short-duration work. Hybrid subtypes exist linking the major fibers of I and IA and IX on a continuum from type I to IC to type IIC, IIA, and IIXA to IIX. With activation of the motor units, a transition of the ATPase isoforms will move from the type IIX to the type IIA subtype (17). The continuum of hybrids fibers goes from the most oxidative isoform type I to the least oxidative isoform type IIX. However, again, when a motor unit is activated, oxidative processes increase and exercise training will even leave type IIX fibers that remain with higher concentrations of oxidative enzymes. Thus, the transition to the primary fiber type as visualized by the histochemical staining is related to the oxidative status of the fiber while it keeps its characteristic size and function. Capillary development also follows this transition to greater density as oxidative needs increase. However, due to the fact that type I motor units are orderly recruited first in almost all activities and are involved for the repetitive endurance activities, their oxidative, mitochondrial, and capillary profiles are always greater than type II or fast motor units.

Again, there is a continuum of myosin ATPase within muscle fibers ranging from that found in the least oxidative type IX fibers and extending to the most oxidative fiber, the type I fiber. It has not been shown that any type of voluntary physical training can change fiber types from type II to type I, or vice versa. In reality, a continuum of fibers, including hybrids expressing more than one type of myosin simultaneously, exists in humans, extending through types IIX, IIXA, IIA, IIC, IIIC, and I. The difference in protein composition within human muscle fibers, compared with the less complex situation in animals, appears to be due to a more intricate regulation by the nervous system to achieve functional changes at the level of the contractile unit.

The advantage of having different types of muscle fibers within a muscle is that the characteristics of the muscle are extended beyond that of any single fiber type. It is the proportion of muscle fiber types within a muscle that gives muscles the properties that make them suitable for different functions.

Factors Affecting Muscle Function
A number of factors affect muscle function. It is well recognized that these factors include the state of training and degree of fatigue of the muscle. The specific adaptations of muscle to exercise training and the relationship of fatigue with muscle function are discussed in later sections. In this section, the interaction of other muscular, neural, and mechanical factors affecting muscle function is discussed. The reader is also referred to Chapter 3: Assessment of Human Muscle Function.

Muscular Factors
Cross-Sectional Area
Muscle size is one of the most obvious factors affecting strength. In isolated as well as intact muscles, maximal strength is related to the cross-sectional area of the muscle. This relationship probably is related to the greater quantities of actin and myosin, and therefore greater numbers of crossbridges that can be activated to produce force when the muscle is larger. Muscles can produce 10 to 20 N/cm² of cross-sectional area (18).

Numerous studies have demonstrated changes in muscle cross-sectional area that occur in conjunction with strength changes produced by resistance training or detraining. Nevertheless, strength and muscle cross-sectional area do not change in parallel. It is clear that the cross-sectional area of the muscle does not fully account for differences in strength among individuals.

Muscle Fiber Type
As described earlier, the maximal force and the power generation of a muscle are related to the percentage of fast-twitch fibers. In other words, muscle that has a high percentage of fast-twitch fibers will generate a greater maximal force than the same-size muscle with a lower percentage of these fibers (19).

Neural Factors
Motor Learning and Recruitment
The importance of neural factors in affecting muscular strength has been recognized through the dissociation between changes in strength and muscle size during a strength-training program. The gains in strength during the first few weeks of such training occur without any change in muscle size (20,21). The general consensus is that these early changes reflect neural adaptations that may include improved muscle activation and improved task performance from motor learning and coordination (22).

Inhibitory Reflexes
Neural factors may also play a role in inhibition of muscle contraction. A protective reflex mechanism is thought to operate through the Golgi tendon organs that may be of particular importance when large amounts of force are being generated (23,24). Reflex inhibition of muscle contraction can also result through other sensory nerve endings. For instance, it has been demonstrated that quadriceps muscle inhibition is mediated through afferent activity of intracapsular receptors (25). It is also likely that muscle contraction is inhibited through pain reflexes (26–28).

Protective reflex mechanisms are also thought to be involved in what has been called the bilateral deficit. The force developed during bilateral contractions of a given muscle group is less than the sum of the forces developed by each limb independently (29,30). Reduced motor unit stimulation is associated with this bilateral deficit (31).

It may be possible to reduce the influence of protective reflexes through strength training. Hypnosis was shown to increase maximal force produced during forearm flexion by 17% among non–resistance-trained individuals, whereas there was no significant change in a strength-trained individual (32). It was concluded that strength training may induce an
inhibition of the protective reflex mechanisms. Furthermore, the bilateral deficit has been shown to be reduced through training with bilateral contractions (33).

The protective reflexes may be reduced in another way. Strength of a muscle group is increased when its activity is immediately preceded by contraction of the antagonist muscle group (34). The precontraction is thought to reduce the neural protective mechanism, allowing a greater force production.

**Mechanical Factors**

**Force-Velocity Relationship**

The maximum force a muscle can exert depends on the speed at which it is contracting. The maximal isometric force of a muscle is always greater than the force that can be exerted during shortening, and the maximal force exerted during lengthening is always greater than that exerted during isometric contraction. This relationship is displayed in Figure 61-8.

It is thought that the shape of the force-velocity curve is explained on the basis of the sliding-filament theory of muscle contraction (16). During a maximal isometric contraction, all crossbridges are formed. However, during shortening, there is an increase in the rate of detachment of crossbridges and an increase in the number of attached crossbridges that exert negative force. The result is a decrease in the total force exerted when the muscle is shortening. During lengthening contractions, the rate of detachment of crossbridges is slower than during a shortening contraction at the same velocity. The effect is that crossbridges are forcibly detached, and so a greater force is produced compared with shortening contractions.

**Positive Versus Negative Work**

Besides the difference in maximal force production through concentric and eccentric contractions, there is a difference in energy cost for performing work through concentric (positive work) and eccentric (negative work) contractions. The energy cost for negative work is dramatically less than that for performing the same amount of positive work (Fig. 61-9) (35,36). This phenomenon is thought to be linked to the requirement for ATP for the detachment and resetting of the crossbridges in concentric work but not in eccentric work (Fig. 61-10) (37).

**Short-Range Stiffness**

When a maximally activated muscle is forcibly lengthened, the force produced by the muscle is greater than that produced isometrically (see Fig. 61-8). However, the situation may be different when the muscle is contracting submaximally. Forced lengthening of a partially contracted muscle results in an initial increase in resistance greater than that produced isometrically, but the resistance may then fall below that produced isometrically (38). This phenomenon is referred to as short-range stiffness and is thought to be accounted for as follows. The rapid rise in tension at the beginning of forced lengthening results from the stretch of attached actin-myosin crossbridges. The stiffness of the muscle is very high while the crossbridges are still attached. However, once the crossbridges begin to dissociate, it is thought that the formation of new crossbridges occurs more slowly in a partially activated muscle, and so the number of attached crossbridges and the force produced are diminished.

Short-range stiffness is important in the initial part of the response of a limb to a disturbing force. For instance, partial contraction of agonist and antagonist muscles about a joint.
increases the mechanical stiffness of the joint and can provide considerable resistance to perturbing forces such as those that might be encountered while walking across a crowded room with a full glass. This same phenomenon may also play a role in protecting joints from traumatic forces.

**Muscle Orientation and Attachment**

The distance a tendon is inserted from the axis of rotation affects the torque generated by that muscle. For a given tension developed by a muscle, a tendon insertion farther from the center of rotation will allow greater torque production, although angular range will be reduced (Fig. 61-11). This anatomic effect allows some muscles to be more suited for production of large forces than others. Small anatomic variations may also account for some of the performance differences among individuals.

**Length-Tension Relationship**

The tension produced by a muscle is affected not only by the contractile elements but also by passive stretch of the elastic elements (see Fig. 61-6). When relaxed muscle is passively stretched beyond its resting length, tension progressively develops (Fig. 61-12) (39). Maximal contraction of the muscle at different lengths yields another length-tension curve. Subtraction of the passive tension from the total tension of the contracting muscle yields a closer representation of the actual tension produced by the contractile elements. The greatest tension generated by the contractile mechanism is at the resting length of the muscle, and the greatest total tension is at a length slightly longer than resting length.

The influence of length on the force produced by the contractile mechanism is related to the way in which the actin and
myosin filaments interact at the sarcomere level. Figure 61-13 shows the length-tension relationship for a single muscle fiber and the overlap of the actin and myosin filaments in a single sarcomere (40). As the overlap between actin and myosin increases, so does the tension production. Maximal tension is developed at lengths yielding maximal contact of actin and myosin filaments. As the sarcomere length decreases further, the actin filaments begin to overlap. It is believed that this interferes with crossbridge formation and causes a decline in tension development. Variations among sarcomeres and muscle fibers cause the length-tension curve to be more rounded for a whole muscle.

An example of the effect of muscle length on force production is apparent by the position of the wrist during a handgrip. While gripping, the wrist is maintained in extension by contraction of the wrist extensor muscles of the forearm. This allows the finger flexor muscles to be at a more optimal part of the length-tension curve. In this way, a stronger handgrip is produced.

**Leverage Effect**

The leverage effect relates to the mechanical advantage offered by the angle of tendon insertion. The torque produced by a muscle is dependent on the sine of the insertion angle. When the insertion angle is 90 degrees, the torque production is greatest for a given tension in the muscle.

Torque production is the net result of the length-tension relationship and the leverage effect. As a result of the force-velocity relationship, the torque production for a given movement is also dependent on the velocity of movement. Figure 61-14 displays an isokinetic torque curve for a
movement commonly tested in the clinical setting (41). This figure demonstrates how maximal torque varies across movements when angular velocity is constant.

**Force Transmission**

One way in which muscle force transmission is thought to vary without changing the tension developed in the muscle fibers is through a change in the elastic elements described in the mechanical model discussed previously (see Fig. 61-6). A decrease in the elasticity of these elements would allow a greater proportion of the force generated by the sarcomere to be transmitted to the skeletal system. With training, tensile strength of connective tissue is known to increase (42). Such changes may improve force transmission of a muscle.

**Elastic Storage and Recovery**

Storage and recovery of elastic energy in the muscle-tendon unit occurs when an active prestretch immediately precedes a shortening contraction. This combination of eccentric and concentric contractions is a natural type of movement that has been referred to as the stretch-shortening cycle (43,44) and allows a greater concentric force production or power output than when the prestretch does not occur. In effect, this phenomenon modifies the length-tension curve so that at a given muscle length, the force produced is greater than that without the prestretch. The precise location and mechanism of the elastic storage is not clear, but it has been attributed to compliance of the crossbridges and connective tissue.

The greater force from a concentric contraction when immediately preceded by an eccentric contraction is a common feature of normal movement. An example of the use of elastic storage in this manner is the knee and hip flexion that occurs immediately before jumping. The most dramatic example of the use of elastic storage and recovery to affect the energy cost of movement is seen with the big red kangaroo. This animal actually uses less energy per unit time as its speed increases as a result of greater use of elastic storage and recovery (45).

**Optimal Phenomena**

The mechanical properties of muscles are important in accounting for what has been referred to as optimal phenomena (35). One such phenomenon is how the aerobic demand of riding a bicycle at a given power output is altered by the seat height. A height can be identified that minimizes aerobic demand and maximizes power output (Fig. 61-15) (46–48). Another example of optimal phenomena in cycling relates to pedaling rate. There is a pedaling frequency at which the aerobic demand to generate a given power output is minimized (see Fig. 61-15). Individuals also have a stride length that optimizes the aerobic demands of running (Fig. 61-16) (49) and a walking speed...
that optimizes the aerobic demands for walking a given distance (Fig. 61-17) (50).

The critical observation to understanding these optimal phenomena has come from the work of Hill (51). He showed that the force-velocity curve could be used to generate a power-velocity curve that shows a point of optimality (Fig. 61-18). Then, by considering the energy cost of developing muscle tension under various conditions, a muscle efficiency-velocity curve was generated that also showed a point of optimality. Thus, the changing of pedaling or stride rate can be considered as moving along the velocity axis of the muscle-efficiency curve. The interaction of the length-tension relationship is also important for some of the optimal phenomena.

**Muscular Fatigue and Endurance**

Endurance and exercise intensity are related by hyperbolic functions, as demonstrated in Figures 61-19 and 61-20 (52–54). At high intensities, exercise can be continued for only short durations, whereas at low intensities, exercise can be continued much longer. The portion of the curve approaching the time axis is predominantly determined by the capacity for aerobic metabolism, whereas the portion of the curve approaching the intensity axis is predominantly determined by the capacity for anaerobic metabolism. The portion with greatest curvilinearity is determined by a combination of aerobic and anaerobic capabilities.

**FIGURE 61-17.** Optimal phenomena during walking. A speed exists at which energy cost per unit distance is minimized. Energy cost is shown here in units of kilocalories per kilogram body mass per meter traveled. (From Cavanagh PR, Kram R. Mechanical and muscular factors affecting the efficiency of human movement. *Med Sci Sports Exerc.* 1985;17:326–331. Reprinted by permission of Williams & Wilkins. Figure based on data of Ralston JH. Energy–speed relation and optimal speed during level walking. *Int Z Angew Physiol.* 1958;17:277–283.)

**FIGURE 61-18.** Optimal phenomena can be explained by derivations from the force-velocity relationship. The power-velocity curve is obtained from the product of force and velocity and demonstrates an optimal point. Energetic data allow generation of the efficiency-velocity curve, which also shows an optimal point. (From Cavanagh PR, Kram R. Mechanical and muscular factors affecting the efficiency of human movement. *Med Sci Sports Exerc.* 1985;17:326–331. Reprinted by permission of Williams & Wilkins. Figure based on Hill AV. The maximal work and mechanical efficiency of human muscles and their most economical speed. *J Physiol.* 1922;56:19–41.)

**FIGURE 61-19.** Isometric endurance as a function of percentage of maximal strength. (Adapted from Rohmert W. Ermittlung von erholungspausen fur statische arbeit des menschen. *Int Z Angew Physiol.* 1960;18:123–164.)
Most exercise is performed at submaximal levels. At the onset of exercise, there is little sense of effort, but as the exercise is continued, performance is eventually reduced. This has led to the concept that fatigue is delayed in onset. However, maximal force generation and endurance capacity may begin to decline from the onset of even submaximal work with the same muscle group (Fig. 61-21) (55). Therefore, it may be preferable to think of fatigue as “any reduction in the maximal force-generating capacity” (56).

Another conceptual issue relates to the common belief that fatigue is a failure of normal physiologic function. Perhaps it is more appropriate to consider fatigue as a protective mechanism for survival. Fatigue prevents the onset of irreversible muscle rigor and protects the subsequent recovery process.

The causes of fatigue have received considerable attention but have not been clearly established. It is evident that multiple factors are involved, and the relative importance of each is dependent on the fiber type composition of the contracting muscle; the intensity, type, and duration of the contractile activity; and the individual’s fitness and motivation level. For instance, the fatigue experienced from high-intensity short-duration exercise such as weight lifting is dependent on factors different from those causing fatigue during low-intensity long-duration endurance exercise.

In daily life, a reduction in power output is frequently limited by central neural drive. Nevertheless, when motivation is high, the primary sites of muscular fatigue are thought to be within the muscle cell rather than the central nervous system or the neuromuscular junction. Specifically, it is thought that fatigue may result from disturbances in the surface membrane, excitation-contraction coupling, or metabolic events (57).

**ACUTE PHYSIOLOGIC RESPONSES TO EXERCISE**

**Dynamic Exercise**

Acute physiologic adjustments occur in most bodily systems with dynamic exercise. Collectively, these adjustments increase the availability of oxygen and nutrients to the active muscle cells, remove exercise-induced metabolic byproducts (e.g., carbon dioxide, lactate, heat), and maintain an appropriate internal milieu (pH, body fluid, etc.) for bodily function.
Dynamic exercise effort is typically expressed in either absolute units (e.g., \( V\text{O}_2 \) in liters per minute) or relative units (e.g., percentage of an individual’s \( V\text{O}_{2\text{max}} \)). Absolute units provide a measure of work performed per unit time, whereas relative units reflect the degree of effort or how strenuous the exercise feels. Most physiologic changes with exercise are more proportional to relative than to absolute work units within a certain intensity range (58). This includes heart rate, ventilation, sympathetic/parasympathetic nerve outflow, circulating hormones, and core temperature. The one parameter that increases more in proportion to absolute work performed than relative intensity is cardiac output, which increases about 5 to 6 L/min for each rise in \( V\text{O}_2 \) of 1 L/min (59). Several factors may influence the magnitude of changes produced by dynamic exercise, such as body posture, age, gender, fitness level, disease state, and mode of exercise (e.g., leg vs. arm).

Many of the cardiovascular adjustments to dynamic exercise are regulated by changes in autonomic nervous activity outflow (60–63). Parasympathetic tone exists at rest, and its withdrawal at the onset of exercise allows heart rate to rise. When work intensity reaches about 50% \( V\text{O}_{2\text{max}} \), parasympathetic withdrawal appears to be exhausted, and any further rise in heart rate is totally dependent on increased sympathetic nerve activity. In addition to increasing heart rate, sympathetic nerve activity increases myocardial contractility, mobilizes nutrients, influences several circulating hormone levels, and contributes to blood flow redistribution by vasoconstriction in inactive regions. Although muscle sympathetic nerve activity appears to increase in the active muscles, metabolic byproducts override this vasoconstriction effect to produce vasodilation. Control of the autonomic nervous system during exercise originates from both central and peripheral receptors located in the motor cortex, ergoreceptors, and arterial and cardiopulmonary baroreceptors (63–65).

The rise in cardiac output (between about four- and sixfold to sevenfold at maximal effort) with upright leg exercise stems from a rise in heart rate (two- to threefold) and stroke volume (about 1.5- to 2-fold) (59,66). The rise in stroke volume results from increased myocardial preload and contractility (59,67,68). Preload increases as a result of enhanced venous return, which is brought on by vasoconstriction and muscle contraction. An increase in contractility leads to more complete emptying of the heart (i.e., decreased left ventricular end-systolic volume), whereas an increase in preload increases left ventricular end-diastolic volume. The net effect is increased stroke volume.

Up to 80% of cardiac output can be distributed to the active muscles at maximal effort, compared with only about 20% of cardiac output being distributed to the muscles at rest. As illustrated in Figure 61-22, this marked blood flow redistribution is accomplished by arterial vasodilation in the active muscles and arterial vasoconstriction in other vascular regions (e.g., splanchnic, inactive muscle, renal) (61,69,70). Total systemic vascular resistance declines progressively with increasing work intensity. The precise mechanisms leading to vasodilation in the active muscle remain debatable but likely stem from changes in several local factors, including potassium, hydrogen ion, endothelium-relaxing factor, adenosine, osmolarity, and others (71,72). The oxygen extraction rate is high within the active muscle. This, combined with the increased percentage of blood flow directed to the muscles, leads to an approximately
threelfold increase in arteriovenous oxygen difference at maximal exercise (59).

Systolic blood pressure rises progressively with increased dynamic workload, whereas diastolic blood pressure generally remains relatively unchanged. The net effect is a modest increase in mean arterial blood pressure (usually <20 mm Hg).

Ventilation rises linearly with \( V\text{O}_2 \) up to the anaerobic threshold (73,74). At and above the anaerobic threshold, ventilatory volume and carbon dioxide output \( (V\text{CO}_2) \) rise out of proportion to metabolism because of the \( \text{CO}_2 \) produced from the bicarbonate buffering system (73,74). The changes in the relationship among \( V\text{O}_2 \), ventilation, and \( V\text{CO}_2 \) with graded exercise testing are used to assess the anaerobic threshold noninvasively (73,74). The significance of determining the anaerobic threshold is that it provides an index of tolerance for sustained work and can be used in prescribing exercise intensity for aerobic training.

Considerable amounts of heat can be produced during exercise, as about 75% of the energy with aerobic metabolism is released as heat. Heat is transported to the skin surface via the cardiovascular system and dissipated via convection, radiation, conduction, and evaporation. Body core temperature rises with exercise, which aids in heat dissipation by increasing the heat flow gradient from the core to the skin (75,76). Reflex and locally mediated arterial vasodilation allow greater blood flow to be directed to the cutaneous vascular bed (63,77,78). This increased blood flow is accomplished in part by blood flow redistribution away from splanchnic and renal arterial vascular beds (63,78). Increased cutaneous venous compliance permits increased cutaneous blood volume and surface flow, which enhance heat dissipation at the skin surface (78,79).

Increased cutaneous blood volume may lead to reduced venous return, left ventricular end-diastolic volume, and stroke volume. Heart rate can rise to compensate for the lower stroke volume as long as maximal heart rate is not attained. The sequence of events associated with heat stress during exercise is referred to as cardiovascular drift (77,79,80). Some cardiovascular drift typically occurs with prolonged (e.g., >60 minutes) exercise in a thermoneutral environment, although the magnitude is much greater in hot and/or humid environments because of the greater cutaneous blood flow and blood volume circulatory demands for heat removal. In a hot environment, the skin-to-environment temperature gradient for dissipating heat via convection, radiation, and conduction is reduced or even reversed if air temperature is greater than skin temperature. Similarly, the ability to dissipate heat via evaporation is proportional to the magnitude of the water pressure gradient existing between the skin and environment.

Sustained levels of high work intensity, especially when performed in combination with heat stress, can lead to high rates of sweat loss reaching 2 to 3 L/h. These rates of sweating can lead to dehydration and subsequent sequelae of decreased total blood volume, blunted cardiac output reserve, reduced thermoregulatory capacity, and decreased work tolerance (79,81–84). To help prevent serious dehydration, fluids should be consumed during sustained exercise. This should occur even in the absence of thirst because a person can lose up to 2% of body water before feeling thirsty (77,81). A number of hormones involved in fluid regulation are altered during exercise, including increases in plasma renin activity, aldosterone, arginine vasopressin, and atrial natriuretic peptide (82,84–89).

**Static Exercise**

The hemodynamic responses to static exercise are related to the percentage of maximal voluntary contractions (MVC) and the amount of muscle mass involved in the contraction (90). Increases in \( V\text{O}_2 \), cardiac output, and heart rate are typically modest during static exercise compared with dynamic exercise (Fig. 61-23) (91). Additionally, total peripheral vascular resistance does not decrease, and stroke volume typically fails to rise as occurs with dynamic exercise (66). Blood flow through the active muscle is dependent on a balance between metabolically induced vasodilation and mechanical restriction of flow associated with contraction of the surrounding muscle. At high static efforts, blood flow through the active muscle is restricted and may be completely occluded (92). Reduced muscle blood flow relative to metabolic demands results in greater reliance on anaerobic metabolism and consequently earlier onset of fatigue than occurs with dynamic exercise. Mechanical and metabolic activation of skeletal muscle afferent nerve fibers during static exercise evokes a pressor response that leads to a significant increase in blood pressure, especially mean and diastolic blood pressures (93). For this reason, static exercise is often viewed as placing primarily a pressure load on the left ventricle, whereas dynamic exercise is viewed as placing more...
of a volume load on the left ventricle (94,95). Although initial studies suggested that the magnitude of the pressor response to static exercise was primarily related to the percentage of MVC, most (but not all) subsequent studies suggest that the amount of muscle mass used also impacts positively on the pressor response.

**HEALTH BENEFITS DERIVED FROM A REGULAR PROGRAM OF EXERCISE**

Many health benefits have been reported with long-term, regular participation in physical activity. The optimal exercise prescription required for primary and secondary prevention of various disease states and for reducing overall mortality remains under investigation (96–99). For many years, the primary recommendation was to participate in an aerobic exercise program. More recently, many health organizations have revised their recommendations to acknowledge the benefit of physical activities that may not meet the aerobic criteria (96,100). The new recommendations encourage all people to accumulate a minimum of 30 minutes of at least moderate exercise most, if not all, days of the week (96,100). This 30 minutes can be accumulated from activities performed throughout the day, such as climbing stairs, gardening, and playing with children. This change in recommendations stems from (a) epidemiologic studies showing reduced mortality and/or morbidity in physically active people, including those who were not necessarily participating in an aerobic exercise program; (b) the failure of many people to adopt long-term aerobic exercise programs because of a variety of factors, including discomfort with higher-intensity effort; and (c) the low levels of physical activity required in our normal daily activities (96). Although the magnitude of benefits in terms of prevention is likely to be dose related, benefits are first contingent on long-term compliance with being physically active. As illustrated in the Activity Pyramid in Figure 61-24, the recent recommendations regarding accumulation of at least 30 minutes of moderate intensity of physical activity most, if not all, days of the week and the recommendations for aerobic and strength training provide a continuum of potential health and fitness benefits that can be used in prescribing exercise to people who differ in exercise readiness, health, exercise preferences, and goals. The concept is to individualize the exercise prescription to get more people physically active on a long-term basis to gain health benefits. Some of the health benefits that have been reported with regular participation in physical activity and/or aerobic exercise training are discussed next.

**All-Cause Mortality**

Several studies indicate a reduction in all-cause mortality (101–103) with long-term, regular exercise participation. Although the minimal and/or optimal intensity, frequency, and duration for reducing mortality remain uncertain, most studies indicate an inverse linear dose-response relationship between the volume of physical activity and all-cause mortality (104). Some data suggest that the threshold level of physical activity required to have a measurable impact on mortality is an expenditure of 500 to 1,000 kcal/wk (104,105). Expending 1,000 calories per week in physical activity has been reported to reduce all-cause mortality by 20% or more (104,105). An inverse relationship has also been reported between laboratory-determined cardiorespiratory fitness and all-cause mortality, with an asymptote often reported at the upper level of the fitness distribution.

**Primary and Secondary Prevention of Cardiovascular Disease**

Physical inactivity is considered a major risk factor for the development of cardiovascular disease (96,106–108). Many epidemiologic studies (105,109–114) have shown that people who are physically inactive have a higher incidence of heart disease than those who are physically active. Based on meta-analysis studies, the estimated risk of heart disease is approximately two times greater for inactive than for active people (115,116). In fact, this risk approaches that for hypertension and hypercholesterolemia. Because physical inactivity affects a greater number of people than any other single heart disease risk factor, the significance of physical inactivity is especially noteworthy.

In secondary prevention, meta-analysis of randomized trials of cardiac rehabilitation indicates a 20% to 25% reduction in mortality in those who participated in cardiac rehabilitation versus those who did not (117,118). The independent influence of exercise conditioning on mortality in this population remains uncertain, as several of these programs included education and advice on modifying other risk factors. For additional
Reduced incidence of cardiovascular disease through physical activity certainly could have important societal implications. The estimated cost of cardiovascular diseases and stroke in the United States for 2002 was $329 billion (119). If the prevalence rates remain the same, future health care costs could escalate with the projected rise in number of Americans older than 65 years. A regular program of exercise, combined with other healthy lifestyle habits, has tremendous potential in curtailing future health care costs. The nation’s health care goals for the year 2010, published in the Healthy People 2010 Program, are aimed at encouraging Americans to adopt healthier lifestyle habits, including greater participation in a regular program of exercise (120).

The mechanism by which physical activity reduces the risk of heart disease is not entirely clear. Some of the reduced risk stems from its impact on improving other risk factors, such as lipids, blood pressure, obesity, diabetes, and psychological stress (96,121). Physical inactivity remains, however, an independent risk factor for coronary artery disease after statistical adjustment for other risk factors.

**Blood Pressure Regulation**

Hypertension is a potent risk factor for coronary artery disease, stroke, congestive heart failure, renal disease, and peripheral vascular disease. Both acute and chronic exercise can lower blood pressure (96,104,122–125). Meta-analysis of aerobic exercise training studies indicates an average decrease in systolic and diastolic blood pressures of 7 and 6 mm Hg, respectively, in those with hypertension and 3 and 2 mm Hg in systolic and diastolic pressures, respectively, in those with normal blood pressure with aerobic training (126). Debate still exists regarding the type of exercise program that produces the best blood pressure-lowering effect (126). Recently, aerobic exercise was reported to be associated with a more favorable blood pressure level than resistance exercise training among 10,000 participants evaluated from 1988 to 1994 in the Third National Health and Nutritional Examination Survey (127). Although there is some suggestion that moderate-intensity dynamic exercise may have a more favorable effect on blood pressure than high-intensity exercise, the data remain inconsistent (126).

**Lipid Management**

Most studies indicate that aerobic exercise training lowers plasma triglycerides and may raise high-density lipoprotein (HDL) cholesterol (96,128–131). Elevated triglycerides and low HDL cholesterol are common lipid abnormalities in those with the metabolic syndrome that, in turn, is associated with increased prevalence of diabetes, hypertension, and coronary artery disease. Questions remain regarding the threshold of exercise needed to produce changes in blood lipids (130,132–135). Some data suggest a dose-dependent relationship (130,133–135), but Leon and Sanchez (129) stated that there was insufficient evidence to conclusively establish a dose-response relationship. Although lipid changes are associated with long-term participation in an exercise program, a single bout of exercise can also produce acute positive lipid changes (104,136,137).

**Weight Control**

Percentages of populations within industrialized countries who are considered overweight and obese are rising at epidemic proportions (138–140). In the U.S. population, more than 30% are considered obese (body mass index [BMI] ≥ 30 kg/m²) and another nearly 30% of women and 40% of men fall within the overweight category (BMI: 25 to 29.9 kg/m²) (139–142). Especially disconcerting is the nearly fourfold increase from 1963 to 2004 (<5% to <20%) in the number of obese children. Obesity carries many health risks, including heart disease, type 2 diabetes, hypertension, dyslipidemia, stroke, gallbladder disease, osteoarthritis, and certain types of cancer (139,143). When the BMI increases above 27, mortality increases sharply (143,144).

Most population studies show an inverse relationship between obesity and physical activity (96). Although many overweight and obese people periodically lose significant amounts of weight through diet alone, most regain this weight (145). In fact, many of these individuals undergo a weight loss/weight regain cycle (yo-yo weight pattern) several times in a lifetime. Rather than trying to lose weight through diet alone or exercise alone, a combination of moderate diet restrictions and increased regular exercise results in the best long-term weight management plan (146). Exercise seems to be especially important in long-term maintenance of weight loss (96,139,144). The most appropriate combination of intensity, duration, frequency, and mode of exercise to recommend for weight loss remains uncertain but will likely vary, depending upon percentage of body fat, age, and presence/absence of orthopedic or other medical complications (146–148). In addition to fat loss, increased physical activity may help maintain or increase lean body tissue during weight loss.

**Type 2 Diabetes Mellitus Prevention**

Diabetes mellitus (DM) can lead to devastating microvascular complications and is a major risk factor for development of coronary artery disease (149). Morbidity and mortality are also higher in those with coronary artery disease when DM is present. Type 2 DM is increasing at an alarming rate in developed countries and is occurring at younger ages (138). Much of this increased incidence is attributable to lifestyle behaviors such as limited physical activity, excessive food intake, and greater prevalence of obesity.

Many studies have shown that exercise is beneficial for those with type 2 DM. A major benefit of exercise is that it increases insulin sensitivity. This effect is seen after a single bout of exercise as well as with chronic exercise. Exercise may act through several mechanisms to improve glucose regulation (104,150,151). Regular exercise may further benefit those with DM by improving body weight, blood pressure, and lipids.

Because of the devastating potential effects of diabetes, primary prevention of type 2 DM is important. Two recent studies (152,153) found that people with impaired glucose tolerance
(prediabetic state) reduced their incidence of DM during follow-up with a lifestyle diet/exercise intervention program designed to lose weight compared with those receiving normal care. This benefit was even more effective than that resulting from the administration of an oral antihyperglycemic agent.

The optimal intensity, frequency, and duration of exercise to recommend for protection against the development of type 2 DM or improved control of type 2 DM remain uncertain. Most studies showing benefits have used an aerobic exercise program. One study suggested that, for each 500 kcal/d increase in energy expenditure, the age-adjusted risk of type 2 DM could be reduced by 6% (154). Another study found that moderate-to-high-intensity exercise (≥5.5 METs) performed more than 40 min/wk reduced the incidence of type 2 DM, whereas lower levels of exercise, regardless of duration, did not provide protection (155). These results suggest that a threshold of exercise must be achieved in terms of intensity and duration.

**Improved Psychological Well-being and Quality of Life**

Mood state is improved immediately after aerobic exercise among regular exercisers (156). Furthermore, it is well recognized that regular physical activity improves general sense of well-being and quality of life (157–162). Changes that have been postulated to contribute to this include reduced psychological stress and improved tolerance for activities of daily living. In addition, regular exercise may help improve quality of life by protecting people from development of disabling diseases such as heart disease, diabetes, cancer, and cognitive decline as well as enabling people with diseases to regain functional work tolerance. Freedom from disease and ability to function independently into old age are important factors in quality of life. Shephard (163) estimated that remaining physically active into old age could allow one to maintain functional independence for 10 to 20 years longer than if one is inactive (Fig. 61-25).

**Maintenance of Bone Density**

Osteoporosis is an important health problem. It causes considerable societal disability among the elderly and is a major contributor to health care costs. Regular physical activity is recommended for enhancing bone density or reducing an age-related decline in bone density (164–172). Exercise habits during the peak bone-forming years may impact on bone density years later (166,171–173). The mechanism by which exercise impacts on bone density is incompletely understood, and the optimal intensity or threshold level of exercise for enhancing bone density remains uncertain. An interaction may exist between an adequate intake of calcium and exercise-induced benefits (174). Weight bearing appears to be an important stimulus, although muscle contraction without weight bearing may also promote bone density (164,165,167–169). In women, very high-intensity training can lead to amenorrhea and reduced bone density (175), an effect believed to be related to decreased estrogen levels.

**Increased Fibrinolytic Activity**

Regular exercise may reduce the risk of thrombotic events by exerting an effect on coagulation and fibrinolytic factors such as lowering fibrinogen, plasminogen activator inhibitor 1, and platelet aggregation while raising tissue plasminogen activation (96,104,176–180). Most types of chronic exercise appear beneficial. However, acute strenuous exercise may increase platelet adhesiveness and aggregability in some individuals, more so in those who are sedentary than in those who are physically active (104,179,181).

**Decreased Inflammatory Marker**

Associations have been reported between C-reactive protein and increased risk of cardiovascular disease. Regular exercise may have an anti-inflammatory effect, as evidenced by reduced levels of C-reactive protein with exercise training (180,182) as well as among those who are physically active (183).

**Improved Endothelial Function**

Improved myocardial perfusion has been reported in patients with known coronary artery disease after a period of exercise training (184). One proposed mechanism is improved endothelial function (184). This may help attenuate a paradoxical
coronary arterial vasoconstrictor response observed in atherosclerotic segments and thereby improve coronary artery blood flow. Aerobic conditioning has been recently reported to help prevent and restore age-related declines in endothelial-dependent vasodilation in healthy people (185).

**Nonpharmacologic Antiarrhythmic Intervention**

By enhancing vagal tone and reducing sympathetic activation, aerobic exercise training may exert an antiarrhythmic effect and thereby reduce the incidence of sudden death (186).

** Improved Sleep**

Epidemiologic studies indicate that regular physical activity may be beneficial in improving sleep quality and reducing sleepiness during normal waking hours (187). The mechanism(s) by which exercise impacts on sleep remains uncertain.

**Possible Enhanced Immune Function**

The impact of exercise training on the immune system remains uncertain. Some studies suggest moderate exercise training may enhance the immune system, whereas heavy training or arduous endurance events may depress the immune system. Much of this evidence stems from studies suggesting a protective effect against respiratory illnesses in those participating in regular moderate exercise training compared with sedentary individuals, but increased risk in athletes participating in high-level competitive events compared with sedentary individuals.

Although several studies have reported changes in many components of the immune system (e.g., cytokines, natural killer cell, immunoglobin) with exercise and/or exercise training, further study is required to ascertain whether these changes impact significantly on immunity and, if so, whether they are beneficial or detrimental (188–192).

**Reduced Cancer Risk**

A growing number of studies are showing that increased physical activity is associated with reduced risk of colon and perhaps breast, prostate, and lung cancer (110,193–195). Some of the potential direct and/or indirect pathways by which exercise may be beneficial include reduced bowel lining exposure to mutagens via accelerated movement of food through the intestines, reduced breast tissue exposure to circulating estrogen, lowered circulating concentrations of blood insulin and growth factors, and improved body weight management. Many questions remain regarding the optimal dose of exercise for prevention of cancers, but, in general, exercise guidelines used for prevention of coronary artery disease are recommended (195).

**AEROBIC EXERCISE**

**Physiologic Adaptations to Aerobic Conditioning**

Aerobic exercise training produces many physiologic adaptations (Table 61-4). An important adaptation is increased work capacity, or $\dot{V}O_2_{max}$. Most studies indicate that sedentary people within diverse populations (age, gender, income, ethnic background, health status) will experience $\geq 15\%$ improvement in $\dot{V}O_2_{max}$ within 3 months of starting aerobic training (59,66,96,196,197). This increase is caused about equally by central cardiovascular adaptations that raise maximal cardiac output and peripheral adaptations that enhance oxygen extraction from the circulating blood (66).

Increased oxygen extraction with aerobic training stems from changes within the trained muscle, including increased capillary density, capillary-fiber ratio, tissue myoglobin, size and number of mitochondria per muscle cell, and respiratory enzyme capacity per mitochondrion (198–200). These muscular adaptations are believed to raise the anaerobic threshold and improve tolerance for sustained work.

The rise in maximal cardiac output with aerobic conditioning resides in increased stroke volume. Maximal heart rate does not rise with aerobic conditioning and may even be lower in well-conditioned endurance athletes than in sedentary individuals. The mechanism by which maximal stroke

<table>
<thead>
<tr>
<th>Table 61.4</th>
<th>Physiologic Adaptations to Aerobic Exercise Training as Observed in Resting and Exercise States</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Rest</td>
</tr>
<tr>
<td>Aerobic power</td>
<td>No change</td>
</tr>
<tr>
<td>Heart rate</td>
<td>Decrease</td>
</tr>
<tr>
<td>Stroke volume</td>
<td>Increase</td>
</tr>
<tr>
<td>Cardiac output</td>
<td>No change</td>
</tr>
<tr>
<td>Myocardial $O_2$ demand</td>
<td>Decrease</td>
</tr>
<tr>
<td>Ventilation</td>
<td>No change</td>
</tr>
<tr>
<td>Arteriovenous $O_2$ difference</td>
<td>No change</td>
</tr>
<tr>
<td>Blood lactate concentration</td>
<td>No change</td>
</tr>
<tr>
<td>Muscle blood flow</td>
<td>No change</td>
</tr>
<tr>
<td>Splanchic blood flow</td>
<td>No change</td>
</tr>
<tr>
<td>Systolic blood pressure</td>
<td>Decrease</td>
</tr>
<tr>
<td>Diastolic blood pressure</td>
<td>Decrease</td>
</tr>
</tbody>
</table>
volume increases with training is not entirely clear but involves increased cardiac preload and probably enhanced myocardial contractility and relaxation (66,201). Aerobic training causes an increase in total blood volume that partially accounts for the increased cardiac preload. The extent of cardiac adaptations appears to be related to such training factors as the length, intensity, duration, and mode of training and the time of life at which training was initiated.

A characteristic finding among elite male endurance athletes is an increased heart size (“athlete's heart”) characterized by increased left ventricular end-diastolic volume and a proportional increase in left ventricular mass and normal wall tension (66,202). It is important to note that the enlarged heart of the athlete differs from the enlarged heart in hypertension and congestive heart failure (203). In the elite athlete, left ventricular hypertrophy is eccentric rather than concentric, and ventricular dilation is proportional to wall thickness. Indices of left ventricular diastolic function are typically normal or increased in athletes but impaired in pathologic states (204). The increased heart size in the athlete is believed to be important in permitting high levels of maximal stroke volume and thereby high functional work tolerance. Although highly trained women frequently show cardiac dimensional adaptations, they rarely demonstrate cardiac dimensional changes outside normal limits (205).

Physiologic adaptations to aerobic training may be restricted to the trained muscles when the amount of muscle mass used in the exercise is small (206–208). For instance, the physiologic benefits of upper-body endurance training appear to be primarily limited to the periphery (206,208). This is attributed to the lower blood flow and cardiac output requirements, which lessen the stimulus for central (i.e., heart) adaptations.

**Principles of Aerobic Conditioning**

Intensity, duration, and frequency recommendations for aerobic exercise training have been previously provided by various groups. In 1998, the American College of Sports Medicine (ACSM) (209) recommended an exercise intensity range of 55% to 90% of maximal heart rate, or 40% to 85% of maximal reserve or heart rate reserve, with the lower end of this range being most applicable to individuals who are quite unfit. This typically translates to rating of perceived exertion (RPE) levels between about 11 to 12 (“fairly light”) and 16 (“hard”) (210,211). The RPE scale on which these values are based is shown in Table 61-5 (212). To achieve these recommended intensities, exercise modes that incorporate a large muscle mass such as walking, running, cycling, swimming, and cross-country skiing are optimal (211,213). Exercise durations of at least 20 minutes, and exercise frequency of at least three times per week was also recommended.

The ACSM has recently updated and clarified the earlier recommendations (100). This group now indicates that in order to promote and maintain health, all healthy adults need to engage in moderate-intensity aerobic physical activity for a minimum of 30 min/d on 5 d/wk or vigorous-intensity aerobic activity for a minimum of 20 min/d on 3 d/wk. The activity can be accumulated in bouts of at least 10 minutes, and combinations of moderate- and vigorous-intensity activities can be used to meet these guidelines. They also point out that larger amounts of physical activity, including more activity at higher intensities, provide additional health benefits.

The optimal rate of progression to follow in implementing an exercise program depends on several factors, including the individual's current activity levels, physiologic limitations, health, age, and exercise goals (99,209). The primary focus should be on adopting a progression that will result in long-term participation. Attempting to do too much too fast can lead to increased dropout rates as a result of perceived excessive discomfort and/or injuries. The ACSM suggests a progression rate subdivided into three phases, as illustrated in Table 61-6. Exercise programs can be tailored according to personal preferences by selecting various combinations of frequency, duration, intensity, and mode of exercise.

**Assessment of Aerobic Capacity**

The $\text{VO}_{2\text{max}}$ provides a reliable, reproducible measure of dynamic work capacity and cardiovascular fitness. It also provides information regarding medical prognosis in patients with heart disease and can aid in evaluating work resumption after recovery from a cardiac event (95,99,214). Many factors influence $\text{VO}_{2\text{max}}$, including age, gender, chronic levels of exercise, genetics, and disease (95,99). With increasing age, $\text{VO}_{2\text{max}}$ declines about 5% to 10% per decade after age 20 (200,215). The age-related loss is attributed to several factors, including a progressive decline in maximal heart rate, body composition changes (e.g., loss of muscle), decreased physical activity, myocardial and vascular stiffening, and disease with increased age (200,215–217). Gender has an impact on $\text{VO}_{2\text{max}}$ with men having higher values than women when compared at a given age and activity level (95,200). The gender difference

---

**TABLE 61.5 Borg RPE 6–20 Scale**

<table>
<thead>
<tr>
<th>Intensity</th>
<th>RPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very, very light</td>
<td>6</td>
</tr>
<tr>
<td>Very light</td>
<td>7</td>
</tr>
<tr>
<td>Fairly light</td>
<td>8</td>
</tr>
<tr>
<td>Hard</td>
<td>9</td>
</tr>
<tr>
<td>Somewhat light</td>
<td>10</td>
</tr>
<tr>
<td>Hard</td>
<td>11</td>
</tr>
<tr>
<td>Very, very hard</td>
<td>12</td>
</tr>
<tr>
<td>Very hard</td>
<td>13</td>
</tr>
<tr>
<td>Very, very hard</td>
<td>14</td>
</tr>
<tr>
<td>Hard</td>
<td>15</td>
</tr>
<tr>
<td>Very, very hard</td>
<td>16</td>
</tr>
<tr>
<td>Very hard</td>
<td>17</td>
</tr>
<tr>
<td>Very, very hard</td>
<td>18</td>
</tr>
<tr>
<td>Very hard</td>
<td>19</td>
</tr>
<tr>
<td>Very, very hard</td>
<td>20</td>
</tr>
</tbody>
</table>

is largely attributed to the smaller muscle/whole-body weight ratio, lower hemoglobin, and smaller stroke volume in women (95). The high levels of \( \dot{V}O_2 \)max observed in elite endurance athletes (e.g., 70 to 85 mL/kg/min) compared with average young men (e.g., 40 to 50 mL/kg/min) probably stem from prolonged intense training plus genetic factors that enhance responsiveness to training (200,218,219).

The \( \dot{V}O_2 \)max can be mathematically defined as the product of maximal cardiac output and maximal arteriovenous oxygen difference. Considerable debate has existed over the years as to what physiologic system limits \( \dot{V}O_2 \)max. Most current evidence points to the cardiovascular system (62,200,220–222). A cardiovascular limitation could reside in maximal cardiac output, maximal muscle blood flow, or maintenance of an appropriate arterial perfusion pressure.

Aerobic capacity can be assessed through direct measurement of \( \dot{V}O_2 \) during maximal exercise or through estimates from maximal or submaximal testing. Direct measurement of \( \dot{V}O_2 \)max is achieved with open-circuit spirometry techniques. In the clinical laboratory setting, \( \dot{V}O_2 \)max is more often estimated, based on the peak workload attained, than directly measured. If \( \dot{V}O_2 \)max is estimated, it is important to discourage significant handrail support with treadmill testing, as this can lead to a marked overestimation (99,223). Reasonably accurate estimates are possible with treadmill and cycle ergometer testing, as work efficiency is relatively constant among people with these testing modes. Conditions in which work efficiency is altered (e.g., walking with a prosthesis) can significantly reduce the accuracy of estimating \( \dot{V}O_2 \)max based on workload (224).

Treadmill or leg cycle ergometer testing is primarily used to evaluate \( \dot{V}O_2 \)max, although arm ergometry testing may be used for those unable to perform leg exercise. Among these test modalities, \( \dot{V}O_2 \) at peak effort is generally highest with the treadmill, intermediate with leg cycle ergometer, and lowest with the arm ergometer (225–227). Differences in peak \( \dot{V}O_2 \) between ergometers are largely attributed to differences in amount of muscle used.

Some individuals do not achieve \( \dot{V}O_2 \)max with symptom-limited testing, as the test is stopped before fatigue at the appearance of an adverse sign and/or symptom of another limiting factor (e.g., claudication). Lack of motivation is another common reason for not achieving a true \( \dot{V}O_2 \)max.

Supervision of exercise tests requires knowledge of when to stop a test for unreasonable risk and of which individuals should not undergo testing because of contraindications for exercise (99,228). Most clinical laboratories have established end points for graded dynamic exercise testing and contraindications for exercise testing (95,99,228).

A less formal method for assessing tolerance for daily activities or for estimating functional work tolerance is to measure distance walked/jogged in a set time period (e.g., 12 minutes) (99).

### STRENGTHENING EXERCISE

#### Importance of Resistance Training

Strength development through resistance training is important for maintaining functional capacity, preventing and recovering from injuries, and improving sports performance (229). With the many physiologic conditions that promote catabolic breakdown of the muscle and connective tissues (e.g., aging, injury, disease), resistance training presents the only natural method to offset such wasting conditions.

#### Physiologic Adaptations to Resistance Training

#### Nervous System Adaptations

Maximal force production from a muscle requires the maximal recruitment of all motor units (230). It has been theorized that untrained individuals may not be able to recruit all motor units.

---

**TABLE 61.6** Example of the Progression of an Aerobic Exercise Program for a Presumably Healthy Individual

<table>
<thead>
<tr>
<th>Program Stage</th>
<th>Week</th>
<th>Exercise Frequency (sessions/wk)</th>
<th>Exercise Intensity (% Heart Rate Reserve)</th>
<th>Exercise Duration (min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial stage</td>
<td>1</td>
<td>3</td>
<td>40–50</td>
<td>15–20</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>3–4</td>
<td>40–50</td>
<td>20–25</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>3–4</td>
<td>50–60</td>
<td>20–25</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>3–4</td>
<td>50–60</td>
<td>25–30</td>
</tr>
<tr>
<td>Improvement stage</td>
<td>5–7</td>
<td>3–4</td>
<td>60–70</td>
<td>25–30</td>
</tr>
<tr>
<td></td>
<td>8–10</td>
<td>3–4</td>
<td>60–70</td>
<td>30–35</td>
</tr>
<tr>
<td></td>
<td>11–13</td>
<td>3–4</td>
<td>65–75</td>
<td>30–35</td>
</tr>
<tr>
<td></td>
<td>14–16</td>
<td>3–5</td>
<td>65–75</td>
<td>30–35</td>
</tr>
<tr>
<td></td>
<td>17–20</td>
<td>3–5</td>
<td>70–85</td>
<td>35–40</td>
</tr>
<tr>
<td></td>
<td>21–24</td>
<td>3–5</td>
<td>70–85</td>
<td>35–40</td>
</tr>
<tr>
<td>Maintenance stage</td>
<td>24+</td>
<td>3–5</td>
<td>70–85</td>
<td>35–40</td>
</tr>
</tbody>
</table>

voluntarily. Thus, part of the adaptation to resistance training is the development of the ability to recruit all motor units (22). Such neural adaptations are thought to be responsible for the increase in strength that precedes any increase in muscle size during the early phase of a resistance training program.

The central nervous system is also capable of limiting force by engaging inhibitory mechanisms that may be protective in nature. Resistance training may result in changes in the order of fiber recruitment or reduced inhibition (231).

### Muscle Enlargement

During heavy-resistance training, motor units containing both type I and type II muscle fibers are recruited and presented with a stimulus for adaptation. As a result, resistance training typically induces increases in the cross-sectional area of both type I and type II muscle fibers. This fiber hypertrophy is translated to increases in the cross-sectional size of the intact muscle, which can be observed after several months of training (232,233).

An upper limit of muscle cell growth has not been determined, but it has been suggested that an ‘optimal size’ or ceiling of adaptation may exist for individual muscle fibers after a prolonged period of strength training (234). This upper limit may be due to a host of genetic factors, but it appears that the availability of myonuclei from satellite cells is important when cross-sectional increases skeletal muscle fiber area are greater than about 25% due to the need to maintain myonuclear domains (235,236). Alterations in the neural patterns of activation within the muscle to recruit all the available fibers may be of importance for the intact muscle to achieve maximal hypertrophy (237).

Muscle fiber hypertrophy is thought to occur through remodeling of protein within the cell and an increase in the size and number of myofibrils (238). Increases in the number of actin and myosin filaments along with sarcomere addition contribute to the increase in muscle fiber size. It has been suggested that the packing density of actin increases, but not myosin, as the contractile proteins are added to the outside of the myofibril without altering the crossbridge configurations (238,239).

The mechanisms and biochemical alterations that mediate the net changes in strength and muscle size observed with heavy-resistance training continue to remain an intense topic of study. Hormonal changes induced by resistance exercise appear to stimulate the uptake of amino acids, yet their incorporation into the contractile unit is not guaranteed. It is possible that contractile proteins accumulate in the muscle fiber either by an increased synthesis or by decreased rate of breakdown, or some combination of both. It is clear that muscle fibers are disrupted, and certain fibers damaged, with intense resistance exercise (240). The extent of this damage is less in trained individuals than in untrained individuals. The repair process of remodeling the muscle fiber may well involve a host of regulatory mechanisms (e.g., hormonal, cell signaling, and metabolic) interacting with the training status of the individual as well as the availability of protein.

Hyperplasia is considered to be a possible adaptive strategy (239), but the extent and frequency of this adaptive response remains a topic of debate. Interestingly, it has been speculated that most of the hyperplasia may well come from neural cell sprouting with the neural outgrowths taking on part of damaged fibers that have lost neural connections and would die without some type of reinnervation by a sprouting neuron (241). It may be that hyperplasia exists, but the magnitude of its contribution in even exceptional situations may not be great (<5%). In addition, hyperplasia may not occur equally in all individuals (238).

### Muscle Fiber Conversion

Transition within the muscle fiber subtypes from type IIX to type IIA appears to be quite typical with resistance training (17,242,243), and this conversion begins to occur within about 2 weeks of initiation of training (244). However, any transformation from type I to type II muscle fibers seems less probable (242,243).

It is not known to what extent this early conversion of muscle fibers may contribute to the initial changes in muscle strength. Nervous system alterations may have the most dramatic effects, mediating strength changes early in a training program, but other changes are taking place in the remodeling of the muscle fibers that may also affect strength.

### Other Adaptations Within the Muscle

It has been demonstrated that strength training can increase the activity of enzymes associated with the ATP-CP energy system (245,246), the anaerobic glycolysis system (245), and the aerobic system (245–247). However, increases in enzyme activity have not been a consistent finding among studies. The design of the strength-training program affects the magnitude of enzyme changes in the muscle. Changes in metabolic enzymes depend on the duration of individual sets rather than on the total amount of work performed. For practical application, normal heavy-resistance programs will have minimal effect on enzyme activities. However, a training program that calls for resistances that are tolerable for at least 30 seconds will most likely induce increases in the activity of muscle enzymes.

Muscular stores of ATP and CP and glycogen may increase with resistance training (248), although these changes are not always observed (249). Whether or not these changes occur with resistance training appears to depend on pretraining status, muscle group examined, and the type of program performed.

Improved capillarization has been observed with resistance training of untrained subjects (243,249–251), but the time required for this adaptation to take place can be greater than 12 weeks (249,252,253). It is thought that low-intensity/high-volume strength training is more likely to increase capillary density than high-intensity/low-volume training.

Myoglobin content in the muscle may be decreased following strength training (252). Thus, it has been postulated that long-term strength training may depress the ability of the muscle fibers to extract oxygen. Again, the initial state of training as well as the specific type of program that is used
may influence the effect of resistance training on myoglobin content.

Few studies have examined the effect of resistance training on mitochondrial density, but the observation of decreased mitochondrial density by one study (254) is consistent with the limited demands for oxidative metabolism placed on the musculature during most resistance-training programs. Yet, this report of decrease in mitochondrial density may be due to volume dilution in the analytical methods. With larger fibers, the relative number of mitochondria in the measured area decreases, when in fact absolute mitochondria and oxidative potential may well increase to address the greater oxidative needs of the activated motor unit (255).

**Body Compositional Changes**

Body compositional changes can occur during short-term (6 to 24 weeks) resistance-training programs. Increases in fat-free mass normally mirror increases in muscle tissue weight, but because of concomitant decreases in fat, total body weight generally increases little over short training periods. The largest gains in fat-free mass that can be expected are a little more than 3 kg (6.6 lb) in 10 weeks of training (229). Though some athletes desire rapid gains in body weight, this is not possible to achieve through gains in muscle mass.

**Endocrine System Adaptations**

It is apparent that the endocrine system plays a major role in the adaptational responses of skeletal muscle to resistance training (256). Serum testosterone concentrations have been demonstrated to increase in the first 6 weeks of resistance training and then to return to pretraining values (244). During the time that serum testosterone concentrations are elevated, significant changes in the type of muscle proteins (e.g., myosin ATPase hybrids, or myosin proteins) occur. Thus, the endocrine system appears to play an important role in the mediation of protein synthesis or reduced degradation that occurs during some phases of training.

**Adaptations to Connective Tissue, Bone, and Cartilage**

The adaptations from resistance training allowing greater tension development by the muscles make it important for the strength of ligaments and tendons to also increase in order to avoid damage to these structures. It is now acknowledged that the dense fibrous tissues that make up tendons and ligaments are adaptable, but no research has been done specifically to examine the effects of heavy-resistance exercise on these structures (257,258).

Physical activity causes increased metabolism, thickness, weight, and strength of ligaments (42,259). Damaged ligaments regain their strength at a faster rate if physical activity is performed after the injury (42,259). Research involving laboratory animals has demonstrated that endurance-type training increases the amount of force necessary to cause separation at the attachment site of a ligament or tendon to a bone and the musculotendinous junction (42). There is reason to believe that resistance training would produce similar results.

The connective-tissue sheaths that surround the entire muscle (epimysium), groups of muscle fibers (perimysium), and individual muscle fibers (endomysium) also adapt to resistance training. These sheaths are of major importance in the tensile strength and elastic properties of muscle, as they form the framework that supports an overload on the muscle. It has been found that muscle hypertrophy is accompanied by an increase in the absolute collagen content of these connective-tissue sheaths (260,261), but the amount of connective tissue appears to increase at the same rate as the muscle tissue (238).

Bone adapts to resistance training, but much more slowly than muscle, requiring 6 to 12 months for adaptations to be observed (262). Bone is sensitive to compression and strain. Such forces are common in resistance training and are related to the type of exercise used, the intensity of the resistance, and the number of sets performed. Training characterized by high-power exercise movements, heavy resistances, and multiple sets appears to be most likely to produce changes in bone metabolism.

Resistance training has been found to increase the thickness of hyaline cartilage on the articular surfaces of bone (263,264). One major function of hyaline cartilage is to act as a shock absorber between the bony surfaces of a joint. Increasing the thickness of cartilage could facilitate the performance of this shock absorber function.

**Peak Oxygen Consumption**

The $\dot{V}O_{2max}$ is normally not considered to be significantly affected by heavy-resistance training (265,266). However, circuit weight training consisting of performing sets of exercises of 12 to 15 repetitions at 40% to 60% of the one-repetition maximum with short rest periods of 15 to 30 seconds between exercises may produce small increases in $\dot{V}O_{2max}$ (267). A resistance training program designed to increase $\dot{V}O_{2max}$ should consist of a high volume of training and relatively short rest periods between sets and exercises. Even then, the maximal increase in $\dot{V}O_{2max}$ brought about by resistance training is substantially less than the 15% to 20% increase associated with traditional aerobic training programs. Therefore, if a major goal of a training program is to increase $\dot{V}O_{2max}$, some form of endurance training should be included in the program.

**Principles of the Exercise Prescription for Enhancing Strength**

Program design is a highly individualized process based on a sound understanding of the basic principles of resistance training. Therefore, one must have a process that can be used to design a resistance training program based on the understanding of a specific paradigm that can be used to develop, prescribe, and modify a resistance training workout over time. The considerations in this program design process will be discussed.
Program Design Variables

The key to successful program design is the identification of specific variables that need to be controlled to better predict the training outcomes. In fact, one study showed that using a personal trainer to better control the progression of intensity led to more optimal gains in strength, compared with allowing the average individual performing resistance training to make such decisions alone (268). This underscores that the most challenging aspect of the resistance training prescription is making decisions related to the manipulation of the program variables. The development of individual training goals for specific phases of cycles of training also becomes paramount in program design. Therefore, the dynamics of planning and changing the exercise prescription is vital for ultimate success of any resistance training program.

Understanding the factors that go into creating the training stimulus is crucial to the success of the process. The creation of an effective exercise stimulus starts with the development of a single training session directed at specific trainable characteristics (e.g., force production, power, hypertrophy). Over time, changes made in the program variables will create the progressions, variations, and overloads needed to observe physiologic adaptations and improved functional abilities.

The Needs Analysis

A needs analysis is a process that consists of answering a series of questions that will assist in the design of a resistance training program (269). The major questions in a needs analysis are:

1. What muscle groups need to be trained?
2. What are the basic energy sources (e.g., anaerobic, aerobic) that need to be trained?
3. What type of muscle action(s) (e.g., isometric, eccentric actions) need to be trained?
4. What is the prior injury history and what are the primary sites of injury for the particular activities in which the individual participates?
5. What are the specific needs for muscle strength, hypertrophy, endurance, power, speed, agility, flexibility, body composition, balance, and coordination?

Biomechanical Analysis to Determine What Muscles Need to be Trained

The first question requires an examination of the muscles and the specific joint angles that need to be trained. This involves a basic analysis of the movements performed and muscles involved with an activity. The decisions made at this stage will help define the choice of exercises, which is one of the program variables.

Because the principle of specificity is a major tenet in resistance training, understanding exactly what movement one is trying to mimic during resistance training is an important aspect of program design. Such analyses will allow one to use specific exercises that require the proper muscles and actions to be used in a manner specific to the activity for which one is being trained.

Resistance training for any sport or activity should include full range-of-motion exercises around all the major body joints. However, training designed for improving specific movements should also be included in the exercise protocol to maximize...
the contribution of strength training to function. The best way to select such exercises is to analyze the biomechanics of the activity and match it to exercises. The importance of movement-specific resistance training programs (e.g., functional training) has surfaced over the past decade and has led to the development of specific equipment (i.e., stability balls, wobble boards) and programs targeting core stability, rotational strength and power, balance, reaction, speed, acceleration, and agility (270).

The principle of specificity is an overriding rule in any of the choices made for the design of a resistance training program. Each exercise and resistance used in a program will have various amounts of transfer to another activity or sport. The amount of transfer will be related to the degree of specificity that can be achieved with the available equipment. When training for improved health and well-being, the specificity of the training will be related to picking the right exercises that can impact a given physiologic variable (e.g., bone mineral density). Other program variables (e.g., rest periods) will also interact to optimize the metabolic system for positive effects as well. Thus, each decision that is made for the choice of exercise will interact with other program variables to create an integrated stimulus of a specific workout protocol.

**Determination of What Energy Sources Need to be Trained**
Performance of any activity, whether self-care, work, or sport, relies on varying proportions of the three energy sources (1). The energy sources to be trained have a major impact on the program design. Resistance training typically focuses on the improvement of the utilization of energy derived from the anaerobic energy sources (ATP-CP and anaerobic glycolysis systems). However, resistance training can contribute to improvement in aerobic training effects by its synergistic effects on reductions in cardiovascular strain, more efficient motor unit recruitment patterns, and improved blood flow dynamics (e.g., enhanced reperfusion of muscles, tolerance of higher pressure loads, and higher percentages of MVC before shutting down venous return) under work stress.

**Determination of What Type of Muscle Actions Should Be Used**
Decisions regarding the use of isometric, dynamic concentric, dynamic eccentric, and isokinetic modalities of exercise are important in the preliminary stages of planning a resistance training program for sport, fitness, or rehabilitation. The basic biomechanical analysis described earlier is used to decide what muscles to train and to identify the type of muscle action involved in the activity. Most resistance training programs use several types of muscle actions. For example, one factor that separates elite power lifters from less competitive power lifters is the rate at which the load is lowered in the squat and bench press. Elite power lifters lower the weight at a slower rate than less competitive lifters, even though the former use greater resistances. In this case, some eccentric training may be advantageous for competitive power lifters. In another example, many holds in wrestling involve isometric muscle actions of various muscle groups. It has been seen that both isometric grip strength and “bear hug” isometric strength are dramatically reduced over the course of a wrestling tournament (271). Therefore, some isometric training may help in the conditioning of wrestlers. Similarly, improved isometric endurance for shoulder stabilization and elbow extension may enhance crutch use, and various functional activities in daily life, such as getting onto or off of a chair or toilet, require isometric strength. Thus, improvement in isometric endurance at high levels of force production may aid recovery and help performance.

**Assessment of the Primary Sites of Injury**
It is also important to determine the primary sites for potential injury in a work or sport activity. Furthermore, it is important to understand the injury profile of the individual. The prescription of resistance training exercises will be directed at enhancing the strength and function of tissue so that it better resists injury or reinjury, recovers faster when injured, and reduces the extent of damage related to an injury. The concept of prehabilitation refers to preventing initial injury by training the joints and muscles that are most susceptible to injury in an activity. The prevention of reinjury can also be an important goal of a resistance training program. Thus, understanding the typical sites of injury for the activity (e.g., knee joints in alpine skiing or low back in construction workers) and the individual’s prior history of injury can help in properly designing a resistance training program. Resistance exercise stress causes muscle tissue damage or disruption, which stimulates hypertrophy and is mediated in part by many of the same inflammatory, immune, and endocrine processes that are involved in the repair of injury. Resistance training most likely helps to condition and prepare these systems for more extensive repair activities needed for faster injury repair as well as less tissue damage due to stronger tissue components.

**Determination of the Need for Various Components of Muscular Fitness**
Determination of the magnitude of improvement needed for variables such as muscle strength, power, hypertrophy, endurance, speed, balance, coordination, flexibility, and body composition is a very important step in the overall resistance training program design. It may seem reasonable to assume that a resistance training program should be designed to optimize all these variables, but that may not necessarily be the case. For example, many sports require a high strength-to-mass or a high power-to-mass ratio. In such a case, resistance training programs should be designed to maximize strength and power while minimizing increases in body mass. This is evident in sports that use weight classes such as weightlifting, power lifting, and wrestling and for those sports that require maximal sprinting speed or jumping ability (e.g., high jump, long jump) where increasing body mass may be detrimental to the maximal height or distance attained, as well as speed. In addition, some sports, such as American football, benefit from...
increasing lean body mass, where the force of impact is greater for a given body mass, assuming power is increased accordingly. Thus, the need for these components of muscular fitness must be evaluated so that proper resistance training partitioning may be used.

Program Design
After the needs analysis has been completed, a specific workout is designed that leads to the development of a training program. These workout sequences should address the specific goals and needs of the individual. Program variables serve as the framework of one specific resistance training session. Changes in the program variables will make up the progression plan for an entire training period. Periods of training are now planned over a training cycle, and “periodization” of training becomes a vital concept in the manipulation of the program variables in chronic program design (229,272). In the rehabilitation setting, such periodization may mean that focused strengthening exercises are shifted to more functional exercises emphasizing the development of muscular endurance. In addition, neuromuscular recovery is often a vital factor in the rehabilitation setting, and using the concept of periodized training will allow for variation in both the exercise stimuli and the planned recovery.

Program Variables
A number of variables can be manipulated in any resistance training prescription (229). These program variables are capable of providing a general description of any single workout protocol.

Choice of Exercise
As described in the needs analysis, the choices of exercises will be related to the biomechanical characteristics of the targeted goals for improvement. The number of possible angles and exercises are almost as limitless as the body’s functional movements. A change in angle affects which muscle tissue is activated. Using magnetic resonance imaging (MRI) technology, investigators have shown that the type of resistance exercise (relative to joint angle and loading method) alters the activation pattern of the muscle (273). Therefore, it is important to understand that if the muscle tissue is not activated by using the most appropriate resistance load (periodized loadings) or joint angles, the desired tissue may not be affected and therefore the optimal rehabilitation progression will not occur.

Exercises can be arbitrarily designated as primary exercises and assistance exercises. Primary exercises are the exercises that train the prime movers in a particular movement. They are typically major muscle group exercises (e.g., squat, bench press, hang pulls). Assistance exercises are exercises that train smaller muscle groups and aid in the movement produced by the prime movers (e.g., triceps press, bicep curls).

Exercises can also be classified as structural (i.e., multiple-joint) or body part (i.e., isolated joint). Structural exercises include those whole body lifts that require the coordinated action of several muscle groups. Power cleans, power snatches, deadlifts, and squats are good examples of structural whole-body exercises (Fig. 61-26). Multiple-joint exercises are those that use more than one joint to perform the movement. An example is the bench press, which involves movement of both the elbow and the shoulder joints. Other examples of multiple-joint exercises include the lat pull-down, military press, and leg press.

Exercises that attempt to isolate a particular muscle group are known as body-part or single-joint exercises. Bicep curls, sit-ups, knee extensions, and knee curls are examples of isolated single-joint or body-part exercises (Fig. 61-27). Many assistance exercises may be classified as body-part or single-joint exercises.

Structural or multiple-joint exercises require neural communication between muscles and promote coordinated use of multiple-joint movements. This has recently been shown scientifically in a study that demonstrates that multiple-joint exercises require a longer initial learning or neural phase than single-joint exercises (274). It is especially important to include structural and multiple-joint exercises in a program when whole-body-strength movements are required for a particular activity. In fact, most sports and functional activities in everyday life depend on structural multiple-joint movements.

For individuals interested in basic fitness, structural exercises may only be advantageous when there is a limited amount of training time and it is necessary to train more than one muscle group at a time. The time economy achieved with
structural and multiple-joint exercises is an important consideration for an individual or a team with a limited amount of time per training session.

**Muscle Action**

Whether the muscle action is performed concentrically, eccentrically, or isometrically has an influence on the adaptation to the resistance exercise. Greater force is produced during eccentric muscle actions with the advantage of requiring less energy per unit of muscle force (275–277). It has been known for some time that there is a need for an eccentric component to optimize muscle hypertrophy (242,243). This is why techniques such as heavy negatives, forced negatives, and slow negatives have been used by individuals (e.g., bodybuilders) to maximize muscle hypertrophy. With pure eccentric resistance exercise, especially in untrained individuals, delayed-onset muscle soreness (DOMS) can be more prominent than with concentric actions (278). In addition, performing a high-intensity training session or performing new exercises at novel angles can result in greater muscle soreness when an eccentric action is involved. Nevertheless, dynamic strength improvements and hypertrophy are greatest when eccentric actions are combined with concentric actions in a repetition (242).

Isometric muscle actions are less metabolically demanding and less conducive to hypertrophy than dynamic muscle actions (279,280). Isometric strength increases are specific to the joint angles trained (i.e., angular specificity) but have shown carry-over of up to ±30 degrees of the trained angle (281). However, the magnitude of carryover appears greatest at joint angles corresponding to greater muscle lengths (282). Dynamic and isokinetic muscle actions can improve isometric strength, but not to the extent of isometric training itself, further supporting the concept of specificity (283,284). Interestingly, a major criticism of several strength training studies is that subjects were trained dynamically yet tested isometrically (285).

**Order of Exercise**

Traditionally, it has been recommended that resistance training workouts sequence large muscle group exercises before the exercises involving smaller muscle groups. Exercising the larger muscle groups first may result in a superior training stimulus being presented to all the muscles involved. This is thought to be mediated by stimulating a greater neural, metabolic, endocrine, and circulatory response, which may augment the training effects of subsequent muscles trained in the workout.

This concept was also used in the sequencing of structural or multiple- and single-joint exercises. In this system, the more complex multiple-joint exercises (e.g., power cleans) were performed initially followed by the less complex single-joint exercises (e.g., bicep curls). This sequencing rationale is that the exercises performed in the beginning of the workout require the greatest amount of muscle mass and energy for optimal performance (286). Thus, these sequencing strategies focus on attaining a greater training effect for the large muscle groups. If structural exercises are performed early in the workout, more resistance can be used due to a limited amount of fatigue.

Contrary to the preceding discussions pertaining to exercise order, different types of “pre-exhaustion” methods have been used by bodybuilders in the United States and by weight lifters in the former Soviet bloc countries in their training. These approaches reverse the order of the exercises so that the small muscle groups are exercised before the larger muscle groups, or a single-joint exercise (e.g., dumbbell flye) is performed before a multiple-joint exercise (e.g., bench press). It is theorized that the fatigued smaller muscles will contribute less to the movement, thereby placing greater stress or intensity on the large muscle groups. Nevertheless, the pre-exhaustion techniques do not appear to be as effective for maximizing strength as the traditional sequencing method (229).

Another consideration in the exercise order is placing exercises that are being taught or practiced (especially complex movements) in the beginning of the exercise order. For example, if an athlete was learning how to perform power cleans, this exercise would be placed in the beginning of the workout, so learning motor skills of an exercise would not be inhibited by prior fatigue.

The order of exercises used in various types of circuit weight training protocols is another issue. The issue relates to whether one follows a leg exercise with another leg exercise or whether it is appropriate to proceed to another muscle group. The concept of pre-exhaustion can come into play here. Arm-to-leg ordering allows for some recovery of the arm muscles while the leg muscles are exercised. This is the most common order used in designing circuit weight training programs. Beginning lifters are less tolerant of pre-exhaustion and arm-to-arm and leg-to-leg exercise orders or stacking exercises for a particular muscle group due to high blood lactate concentrations (10 to
14 mmol/L), especially when rest periods between exercises are short (10 to 60 seconds) (287,288). In contrast, stacking exercises is a common practice among elite bodybuilders in an attempt to bring about muscle hypertrophy.

For strength training, it is recommended that basic strength exercises such as the squat and the bench press be performed initially during the workout. Training for enhanced speed and power entails performance of total-body explosive lifts such as the power clean and the jump squats in the beginning of a workout. Bodybuilding entails performance of many exercises under various conditions where training through fatigue is paramount. Improper sequencing of exercises can compromise the lifter’s ability to perform the desired number of repetitions with the desired load. Therefore, exercise order needs to correspond with specific training goals. A few general methods for sequencing exercises for both multiple or single muscle group training sessions are:

- Rotation of push/pull exercises for total-body sessions.
- Rotation of upper-/lower-body exercises for total-body sessions.
- Large muscles before smaller ones.
- Multiple-joint exercises performed before single-joint exercises.
- Weak-point exercises performed before stronger ones.
- Olympic lifts before basic strength and single-joint exercises.
- Most intense exercises performed before least intense (particularly when performing several exercises consecutively for the same muscle group).

One final consideration for exercise order is the fitness level of the individual. As discussed earlier, training sessions should never be designed that are too stressful for an individual, especially a beginning trainee.

**Number of Sets**

First, it should be noted that not all exercises in a training session need to be performed for the same number of sets. The number of sets is one of the factors in any volume-of-exercise equation (e.g., sets 3 reps). Typically, three to six sets are used to achieve the optimal gains in strength, and the physiologic responses appear to be different with three versus one set of exercises in a total body workout (289,290). It has been suggested that multiple-set systems work best for development of strength and local muscular endurance (291,292), and the gains made will be at a faster rate than gains achieved through single-set systems (293). In most training studies, one set per exercise performed at the 8 to 12 repetition maximum (RM) at a slow velocity has been compared with both periodized and nonperiodized multiple-set programs. In untrained subjects, several studies have reported similar-strength increases between single- and multiple-set programs, and some have reported superiority of multiple sets. Studies examining resistance-trained individuals have shown multiple-set programs superior for strength, power, hypertrophy, and high-intensity endurance improvements (292,294–297). Most studies of 14 weeks or longer (especially in trained individuals) demonstrate the superiority of varied, multiple-set programs for long-term improvement. These findings have prompted the recommendation from the ACSM for periodized multiple-set programs when long-term progression (not maintenance) is the goal (298). No study has shown single-set training superior to multiple-set training in either trained or untrained individuals. Therefore, it appears that both programs are effective for increasing strength in untrained subjects during short-term training periods (i.e., 6 to 12 weeks), yet multiple sets may provide the advantage for upper body exercises during this early phase of training (299).

However, some short-term and all long-term studies support the contention that higher volume of training is needed for further improvement and progression in physical development and performance (300). Still, the need for variation is also critical for continued improvement, and this includes the use of lower-volume training programs for certain phases of the overall training cycle. The key factor is the use of “periodization” of training volume rather than number of sets, which represents only one factor in a volume and intensity periodization model.

Considering the number of variables involved in resistance training, comparing single- and multiple-set protocols may be an oversimplification. For example, several of the previously mentioned studies compared programs of different set number regardless of differences in intensity, exercise selection, and repetition speed. In addition, the use of untrained subjects during short-term training periods has raised criticism (301), as untrained subjects have been reported to respond favorably to most programs (302). Part of this has evolved because of the popularity of a specific single-set program in which researchers desired to compare it with other protocols. The use of a single-set program may be effective for untrained individuals during the early phases for lower body strengthening (299,303). However, it appears that greater volume (within reasonable limits) is needed beyond this point in training to produce optimal improvements. In advanced lifters, further increases in volume may be counterproductive, but it appears that the correct manipulation of both volume and intensity produces optimal performance gains and avoids overtraining (304,305).

Multiple sets of an exercise present a training stimulus to the muscle during each set. Once initial fitness has been achieved, a multiple presentation of the stimulus (three to six sets) with specific rest periods between sets that allow the use of the desired resistance is superior to a single presentation of the stimulus. The concept that a muscle or muscle group can only perform maximal exercise for a single set has not been demonstrated. In fact, multiple sets at a 10 RM using the same resistance can be repeated with as little as 1 minute rest in highly trained bodybuilders (306) or athletes trained to tolerate short-rest-period protocols (295).

The importance of the exercise volume (sets × repetitions) is a vital concept of training progression. This is especially true in individuals who have already achieved a basic level of training or strength fitness. The time course of volume changes is important to the change in the exercise stimulus in periodized
training. Use of a constant volume program may lead to staleness and lack of adherence to training. Ultimately, the variation of training volume (i.e., using both high- and low-volume protocols) to provide different exercise stimuli over long-term training periods will be important to provide rest and recovery periods.

The number of sets performed per workout for multiple-set programs is highly variable and has not received much attention in the literature. In general, the total number of sets per workout will be affected by (a) the muscle groups trained and their consequent size (e.g., large- vs. small-muscle-mass exercises); (b) intensity (fewer sets for high intensity and vice versa); (c) training phase (whether the goal is strength, power, hypertrophy, endurance); (d) training frequency and workout structure (e.g., total-body vs. upper/lower body splits vs. muscle group split workouts); (e) level of conditioning; and (f) use of anabolic drugs (which enable lifters to tolerate higher than normal training volumes). It is common to see resistance training programs incorporating anywhere from 10 to 40 sets per workout. The correct number is based on the individual and depends on the preceding factors.

Rest Periods Between Sets, Exercises, and Repetitions

An understanding of the influence that rest periods have in dictating the stress of the workout and influencing the amount of resistance has been a topic of study over the past 15 years. Rest periods between sets and exercises determine the magnitude of ATP-CP energy source resynthesis and the concentrations of lactate in the blood. The length of the rest period significantly alters the metabolic, hormonal, and cardiovascular responses to an acute bout of resistance exercise, as well as performance of subsequent sets (287,288,295,306,307). Kraemer (295) reported differences in performance with 3- versus 1-minute rest periods. All subjects were able to perform 10 repetitions with 10-RM loads for three sets with 3-minute rest periods for the leg press and bench press. However, when rest periods were reduced to 1 minute, 10, 8, and 7 repetitions were performed, respectively.

For advanced training for absolute strength or power, rest periods of at least 3 to 5 minutes are recommended for structural exercises (e.g., squat, power clean, deadlift) using maximal or near-maximal loads, whereas less rest may be needed for smaller muscle-mass exercises or single-joint movements. For novices to intermediate-level lifters, 2 to 3 minutes of rest may suffice for structural lifts, as loading used during this stage of resistance training appears less stressful to the neuromuscular system (i.e., advanced lifters require resistances closer to his/her genetic potential where maximizing energy stores is crucial to attainment of that level of strength). Robinson et al. (308) reported a 7% increase in squat performance after 5 weeks of training when 3-minute rest periods were used, compared with only a 2% increase when 30-second rest periods were used. Pincivero et al. (309) reported significantly greater strength gains (5% to 8%) when 160-second rest intervals were used compared with 40-second ones. Strength and power performance are highly dependent on anaerobic energy metabolism, primarily through the phosphagen system. It appears that the majority of phosphagen repletion occurs within 3 minutes (310,311). In addition, removal of lactate and other metabolites may require at least 4 minutes (308). Therefore, performance of maximal lifts requires maximal energy substrate availability before the set with minimal or no fatigue. Stressing the glycolytic and ATP-CP energy systems may enhance training for muscle hypertrophy (e.g., bodybuilding). Thus, less rest between sets appears to be effective for hypertrophy, whereas the use of both types of protocols may be important for optimizing strength and size.

In studies by Kraemer et al. (287,288,307), a combination of workouts was used to compare the impact of rest changes on lactate responses. The 5/3 workout consisted of using a 5-RM resistance for all exercises with 3 minutes of rest between sets and exercises. The 5/1 workout consisted of using a 5-RM for all exercises with 1 minute of rest between sets and exercises. The 10/3 and 10/1 workouts consisted of using 10-RM with either a 3- or 1-minute rest between sets and exercises, respectively. Comparisons of the 5/1 to 5/3 and 10/1 to 10/3 demonstrated the dramatic effect rest periods have on blood lactate concentrations. Short rest periods significantly elevated blood lactate concentrations compared with longer rest periods. Comparisons of 5/1 to 10/1 and 5/3 to 10/3 demonstrated the effects on blood lactate concentrations of repetitions at an RM resistance and total work (repetitions × resistance × distance resistance moved) performed. These comparisons indicated that higher volumes of work (10-RM) resulted in higher blood lactate concentrations. The effect that rest periods between sets and exercises, and the total work performed have on blood lactate concentrations is similar for both genders.

These studies indicate that the heavier resistance does not necessarily result in higher blood lactate concentrations. It is the amount of work performed and the duration of the force demands placed on the muscle(s) that influence the blood lactate concentrations. In other words, to produce high concentrations of blood lactate the high-intensity-exercise stimuli must be maintained for a relatively longer time (>20 seconds) than for pure ATP-CP phosphagen metabolism. In addition, repeating the stimuli with limited rest continues to drive the accumulating amounts of lactate in the blood to higher levels. The classic 10 RM performed in a multiple-set fashion (e.g., three sets) provides for such stimuli (287). In addition, a significant amount of muscle mass (e.g., leg exercises vs. arm curls) must be stimulated. High force that produces a greater time under tension may also stimulate greater lactate responses than work-matched, higher-velocity, lighter-resistance, and higher-power protocols (312).

From a practical standpoint, it has been demonstrated that short-rest programs can cause greater psychological anxiety and fatigue (313). This might be due to the greater efforts, discomfort, muscle fatigue, and metabolic demands. The psychological ramifications of using short-rest workouts must also be carefully considered when designing a training session. The increased anxiety appears to be due to the dramatic metabolic demands characterized by short rest workouts (i.e., 1 minute.
or less). Although the psychological demands are higher, the changes in mood states do not constitute abnormal psychological changes and may be a part of the arousal process before a demanding workout.

The frequent use of such high-intensity workouts with short rest periods and heavy loading should be slowly introduced into a training program to enable a gradual improvement in tolerance to increased muscle and blood acid levels (decreased pH), and improvements in acid-base buffer mechanisms (314). When such adaptations are vital to an activity (e.g., wrestling or 400- to 800-m track events), the progression of training protocols from long to short rest period lengths is important to improve function. Usually, such a program is performed within the context of a more classical strength/power training program for a sport (e.g., two high-lactate workouts and two strength-power workouts in a week cycle) or as a pre-season program for 8 to 12 weeks before the start of the sport season. Short rest period length is also characteristic of circuit weight training, but the resistances used are typically lighter (i.e., 40% to 60% of 1 RM are used) (267). This type of training session does not result in blood lactate concentrations as high as those of short-rest-period 10-RM sessions.

Lactic acid may not be the “bad guy” we have often thought it to be (315). Although it may contribute to fatigue, it can be used as a source of energy. Furthermore, it provides a relative comparison of the stress and accumulative utilization of the anaerobic glycolysis system. The type of training session, including rest periods, will determine to a great extent the amount of lactic acid that is produced and removed from the body. Recent studies have shown that elevations in blood lactate concentrations may be important for increases in muscle strength and hypertrophy (316,317). Thus, the role of lactic acid during resistance training appears to be important, depending on the rehabilitation or training goals.

If a particular needs analysis identifies lactic acid as the primary energy source, the rest periods may be gradually shortened to allow the buildup of blood lactate, thus encouraging an increased tolerance and buffering of more acidic conditions. This type of training design (particularly for preseason training) may allow better tolerance for such anaerobic athletes as wrestlers, sprinters (400 to 800 m), and basketball players. Other anaerobic athletes, such as baseball players, rely primarily on the ATP-CP energy system to perform their skills. Consequently, resistance training programs that elevate lactic acid concentrations may not be necessary to improve performance of such sports. Careful manipulation of rest periods is essential to avoid placing inappropriate and needless stresses on the individual during training. Furthermore, because of fatigue created by a high-volume, short-rest workout, it may be inappropriate to place such a workout immediately before a training session designed to develop skill in the sport or activity. Exceptions to this rule of thumb may be in sports such as wrestling, where all skills are performed under conditions of high lactate concentrations.

Training for improved local muscular endurance implies the individual (a) performs several repetitions (or long-duration sets), (b) trains to and beyond the point of fatigue, and/or (c) minimizes recovery between sets (i.e., training in a semi-fatigued state). Therefore, many repetitions and shorter rest periods (30 to 90 seconds or less) for local muscular endurance training appear to be most effective (318).

The amount of rest taken between repetitions has only been partially addressed. Rooney et al. (319) had subjects train using 6 to 10 consecutive high-intensity repetitions or 6 to 10 repetitions separated by 30-second rest periods. They reported significantly greater strength improvement with consecutive repetitions (56%) than with extended rest between repetitions (41%). These findings demonstrate that fatigue may contribute to the strength training stimulus. However, if a high percentage of the peak velocity or power is desired for each repetition in the workout, longer rest periods with fewer repetitions for several sets may need to be performed. Future research will need to address this aspect of rest between repetitions, as the “quality” of each repetition begins to take on greater importance in producing gains in strength and power. Rethinking of the old “rest-pause” training system may be one of the new directions for research in optimizing quality of the training session. With new feedback systems on resistance exercise equipment allowing signaling of other performance factors beyond just lifting a weight, the quality of each repetition can be evaluated based on the percentage of peak velocity or maximal power. The concept of a quality repetition in a workout will start to see more scrutiny and evaluation. Interestingly, it remains to be determined where this “fatigue” stimuli and the later, more “full recruitment” stimuli requiring a longer rest period interface for optimal training effects. Most likely both styles of training may be needed to gain different aspects of strength, size, and power fitness.

**Resistance Used (Intensity)**

The amount of resistance used for a specific exercise is probably one of the key factors in any resistance training program (293). It is the major stimulus related to changes observed in measures of strength and local muscular endurance. When designing a resistance training program, a resistance for each exercise must be chosen. Either RM or the specific resistance that only allows a specific number of repetitions to be performed probably provides the easiest way to determine a resistance. Typically, one uses a training RM target (a single RM target such as 10-RM) or RM target zone (a range such as 3 to 5 RM). As the strength level of the lifter changes over time, the resistance is adjusted, so a true RM target or RM target zone resistance is used.

Research has supported the basis for an RM continuum (291,293,318,320). This continuum simply relates RM resistances to the broad training effects derived from their use. An inverse relationship exists between the amount of weight lifted and the number of repetitions performed. Several studies have indicated that training with loads corresponding to 1- to 6-RM was most conducive to increasing maximal dynamic strength (321,322). Although significant strength increases have been reported using loads corresponding to 8- to 12-RM
(244,295,323) and this load range appears most effective for increasing muscular hypertrophy (324), loads lighter than this (i.e., 12- to 15-RM and lighter) have only had small effects on maximal strength in previously untrained individuals (318). However, they have been shown to be very effective for increasing local muscular endurance (325). Contrary to early studies in resistance training, it appears that using a variety of training loads is most conducive to increasing muscular fitness as opposed to performing all exercises with a 6-RM load (326). Therefore, periodized training in which great load variation is included appears most effective for long-term improvements in muscular fitness.

Moving away from the 6-RM or less strength stimulus zone, the gains in strength diminish until they are negligible. The strength gains achieved above 25-RM resistances are typically small to nonexistent in untrained individuals (291,318) and may be related to enhanced motor performance or learning effects when they occur. Various individual responses due to genetic predisposition and pretraining status affect the training improvements observed. But after initial gains have been made as a result of learning effects, heavier resistances will be needed to optimize muscle strength and size gains.

Another method of determining resistances for an exercise involves using a percentage of the 1-RM (e.g., 70% or 85% of the 1-RM). If the trainee’s 1-RM for an exercise is 100 lb, 80% of the 1-RM would be 80 lb. This method requires that the maximal strength in various lifts used in the training program be evaluated regularly. If 1-RM testing is not done regularly (e.g., each week), the percentage of 1-RM used in training decreases and therefore the training intensity is reduced. From a practical perspective, use of percentage of 1-RM determination for many lifts may not be feasible because of the amount of testing time required. Use of an RM target or an RM target zone allows the individual to adjust resistances in response to an ability to perform greater repetitions in order to stay at the RM target or within the RM target zone.

The use of a percentage of the 1-RM resistance is warranted for lifts related to the competitive Olympic lifts of the clean and jerk, snatch, and variations. Since these lifts require coordinated movements and optimal power development from many muscles to result in correct lifting technique, the movements cannot be performed at a true RM or to momentary failure. The reduction in velocity and power output experienced in the last repetition of an RM set is not conducive to such structural lifts. When determining the resistance for power movements, a percentage of the 1-RM loading should be used rather than failure to lift the weight or RM.

In two studies by Hoeger et al. (327,328), the relationship between the percentage of 1-RM and the number of repetitions that can be performed was studied in both trained and untrained men and women. It was demonstrated that this relationship varies with amount of muscle mass needed to perform the exercise (i.e., leg press uses more muscle mass than knee extensions). When using machine resistances with 80% of the 1-RM, previously thought to be primarily a strength-related prescription, the number of repetitions that could be performed was typically greater than 10, especially for large-muscle-group exercises such as the leg press. The larger-muscle-group exercises appear to need much higher percentages of the 1-RM to keep them in the strength RM zone of the repetition continuum. Shimano et al. (329) examined this relationship using free weights in trained and untrained men. At 60% of 1-RM, significantly more repetitions were performed during the back squat than the bench press or arm curl by both groups of men. At 80% and 90% of 1-RM, there were significant differences between the back squat and other exercises; however, differences were much less pronounced than at 60% of 1-RM and these intensities were within the 5 to 10 RM zone needed for strength development. This was different from the machine based data of Hoeger et al. (327,328). No differences in number of repetitions performed at a given exercise intensity were noted between trained and untrained (except during bench press at 90% 1 RM). Thus, using percentages of the 1-RM are dependent upon exercise equipment (free weights vs. machines) and exercise choice, and the RM zone should be checked to be sure the percentage is yielding the proper number of repetitions to meet training related goals. Initial data in men appear to indicate that training status of the individual has a minimal impact on the number of repetitions performed at relative exercise intensity with free weights.

In a study by Kraemer et al. (330), it was shown that power lifters could lift 80% of their 1-RM in the leg press for 22 repetitions (e.g., 80% of the 1-RM equated to the 22-RM resistance) and the untrained controls could only perform 12 repetitions at 80% of their 1-RM (e.g., 80% of the 1-RM equated to the 12 RM). Such data, along with the data presented in the two important studies by Hoeger et al. (327,328), clearly indicate that the method used to determine the resistance to be used for specific exercises in a training program must be carefully considered for each muscle group and for each specific type of lift and the exercise mode used (e.g., free weight squat vs. leg press machine). In general, a certain percentage of the 1-RM with free weight exercises will allow fewer repetitions than the same percentage of 1-RM on a similar exercise performed on a machine. This is probably due to the need for greater balance and control in space when using free weights.

Charts have been developed to predict the 1-RM from the number of repetitions performed with a submaximal load or to help determine an RM (e.g., from 1 to 10) from the 1-RM resistance that can be lifted. Unfortunately, most of these charts assume a linear relationship between these variables, which is not the case. Thus, such charts and the resulting values should only be used as rough estimates of a particular resistance to use for an RM resistance or to predict an individual’s 1-RM. A variety of prediction equations are available to predict 1-RM, but these equations have the same inherent weaknesses as the prediction charts.

The amount of weight lifted per repetition or set is highly dependent on other variables, such as exercise order, volume, frequency, muscle action, repetition speed, and rest period length (331). Altering the training load can significantly affect
the acute metabolic (332,333), hormonal (287,288,334), neural (231,335), and cardiovascular (336) responses to training.

The load required to increase maximal strength may depend on training status. A load of at least 45% to 50% of 1-RM is needed to increase dynamic muscular strength in beginners (337). However, greater loads are needed in more experienced subjects. Häkkinen et al. (335) reported that at least 80% of 1-RM was needed to produce any further neural adaptations during resistance exercise. Neural adaptations are crucial to resistance training, as they precede hypertrophy during intense training periods. Thus, it appears that a variety of loads conducive to increasing both neural function (i.e., increased motor unit recruitment, firing rate, and synchronization) and hypertrophy is most effective for strength training.

**Repetition Speed**

The speed used to perform dynamic muscle actions affects the adaptations to resistance training. Repetition speed is dependent on training loads, fatigue, and training goals, and has been shown to significantly affect neural (276,335), hypertrophy (338,339), and metabolic (340) adaptations to resistance training. Force production and repetition speed directly interact during exercise performance. Generally, force production is greatest at slower speeds and lowest during high-speed movements. This relationship is graphically represented as the force-velocity curve discussed earlier in this chapter. The implications of the force-velocity relationship are that training at slow velocities with maximal tension will be effective for strength training and training with high velocities will be effective for power/speed enhancement. This generally is the case; however, an interaction between both extremes of the velocity spectrum may be most effective for both strength and power enhancement.

Generally, moderate-to-rapid speeds (i.e., 1 to 2 seconds concentric, 1 to 2 seconds eccentric) are most effective for enhancing gains in strength and power performances compared with slow speeds, which appear most effective for increasing local muscular endurance and isometric strength for a specific number of repetitions. LaChance and Hortobagyi (341) reported significantly greater average power, work, and total number of repetitions performed using a self-selected (rapid) speed, compared with 2-second concentric and 2- or 4-second eccentric repetition speeds. Thus, improving set performance (i.e., number of repetitions or load) may be best accomplished with use of moderate-to-fast speeds.

A distinction needs to be made when examining intentional and unintentional slow-speed repetitions. Unintentional slow lifting speeds are used during high-intensity repetitions (i.e., strength training) in which either the loading and/or fatigue are responsible for the longer repetition duration. For example, the positive phase of a 1-RM bench press and the last repetition of a 5-RM set may last 3 to 5 seconds (342). This may be considered slow by comparison; however, lifting the weight faster is not possible during maximal efforts. These types of unintentionally slow lifting speeds are crucial to maximal strength development.

Intentional slow-speed repetitions are used with submaximal loads, where the lifter has greater control of the speed. Keogh et al. (343) showed that the force produced during the concentric phase was significantly lower for intentionally slow (5-second concentric, 5-second eccentric) lifting compared with traditional (i.e., moderate) speed with a corresponding lower integrated electromyographic (IEMG) activity for the slow speed. This study shows that motor unit activity may be limited when attempting to intentionally lift slowly. In addition, the lighter loads required to lift slowly may not provide an optimal stimulus for strength enhancement. Hatfield et al. (344) demonstrated that when intentionally slow speeds (10-second concentric and 10-second eccentric) were used, a significant and dramatic reduction in the number of repetitions, power output, and force output that could be performed was observed making such programs less than optimal for typical strength training goals. Therefore, intentionally slow lifting appears to be most suitable for increasing local muscular endurance, where the time under tension is greater than moderate and fast speeds.

In comparison, both fast and moderate lifting speeds can increase local muscular endurance, depending on the number of repetitions performed and rest between sets. Training with fast speeds is the most effective way to enhance muscular power and speed, and is effective for strength enhancement (345) but not as effective for hypertrophy at slow or moderate speeds (335). High-speed repetitions impose fewer metabolic demands in exercises such as the leg extension, squat, row, and curl than slow and moderate speeds (340). In addition, training for power is best accomplished through lighter loads (30% of the 1-RM) performed at maximal speeds (<1:1.5 ratio) (346).

A popular technique used for both strength and power training is compensatory acceleration (347,348). Compensatory acceleration requires the lifter to accelerate the load maximally throughout the motion (regardless of momentum) during the positive phase, thus striving to increase bar velocity to maximal levels. A major advantage is that this technique can be used with heavy loads and is quite effective, especially for multiple-joint exercises (349). Hunter and Culpepper (350) and Jones et al. (349) reported significant strength and power increases throughout the range of the movement when compensatory acceleration was used, and the increases were greater than those for training with a slower speed (349).

**Rest Periods between Workouts (Training Frequency)**

The number of training sessions performed during a specific period time (e.g., 1 week) may affect subsequent training adaptations. Frequency also includes the number of times certain exercises or muscle groups are trained per week and depends on several factors, such as volume and intensity, exercise selection, level of conditioning and/or training status, recovery ability, nutrition, and training goals. Training with heavy loads increases the recovery time needed before subsequent sessions, especially for multiple-joint exercises (337). The use of extremely heavy loads (90% to 100% of 1-RM)
may require 72 hours of recovery, whereas low and moderate loads (60% to 85% of 1-RM) may require less recovery time (48 and 24 hours, respectively) (351). In addition, reduced frequency is adequate during maintenance training. Training 1 or 2 d/wk is adequate for muscle mass, power, and strength retention (337). However, this appears effective for short time periods, as such long-term maintenance training (i.e., reduced frequency and volume) may lead to detraining.

Heavy eccentric training requires greater recovery time between workouts. Loading during eccentric training may be substantially more than concentric (i.e., 120% to 130% of 1-RM). Studies show that eccentric exercise is more conducive to DOMS (278,352,353). Eccentric training causes greater muscle fiber and connective-tissue disruption, enzyme release, DOMS, and impaired neuromuscular function that limits force production and range of motion (354). Recovery times of at least 72 hours are required before initiating another session requiring several heavy sets or supramaximal eccentric lifts (351). Thus, frequency modulation is needed for heavy eccentric training. A recent study in untrained subjects compared frequencies of 1 d/wk to 2 or 3 d/wk (355). Each session consisted of seven sets of 10 1- or 2-second eccentric-only contractions for the quadriceps muscles. Both training groups showed improvement following training. However, the results showed that eccentric training once per week was effective for maintenance, whereas two times per week was more effective for strength increases. Thus, the inclusion of heavy eccentric repetitions may necessitate a change in frequency (or the muscle groups trained per session) to accommodate the greater muscle damage.

Numerous resistance training studies have used frequencies of 2 or 3 alternating days per week in untrained subjects (242,356). This has been shown to be a very effective initial frequency. If the resistance training is not excessive, only moderate amounts of delayed muscular soreness should be experienced 1 day after the session. Some studies have shown 3 d/wk to be superior to 2 d/wk (357), whereas 3 to 5 days were superior in other studies (358,359). The progression from beginning to more aggressive training does not necessitate a change in frequency but may be more dependent on alterations in other acute variables, such as exercise selection, volume, and intensity. However, it is common to see three or four training sessions per week among those individuals training more aggressively. Increasing training frequency allows for greater specialization (i.e., greater exercise selection per muscle group and/or volume in accordance with more specific goals). To achieve a higher frequency of training, more detailed workouts should be developed, as simply performing the same exercises four times rather than three times per week is not the optimal approach to increasing frequency. Programs should use exercises that involve similar muscle groups but use different angles for particular movements (referred to as split programming) (229). For example, a 4 d/wk routine that involves performing the bench press all 4 days would be better designed by having the individual perform a regular bench press on 2 days and an alternate type of bench press (e.g., incline bench press) on the other 2 days. Thus, split programming allows for more variety in the exercise choices, because of the increase in the number of training days available. In addition, total-body or split-body-part workouts have been used to allow more training variety. Both styles of training (total-body or split-body-part workouts) have been shown to produce improvements in muscle strength and size (360). However, it is recommended that similar muscle groups or selected exercises not be performed on consecutive days during split-routine workouts to allow adequate recovery and to minimize the risk of overtraining.

Training frequency for advanced or elite athletes may vary considerably (depending on intensity, volume, and training goals) and is typically greater than that in intermediate lifting. One study demonstrated that football players training 4 or 5 d/wk achieved better results than those self-selected frequencies of 3 and 6 d/wk (361). Weight lifters and bodybuilders typically use high-frequency training (i.e., four to six sessions per week). The frequency for elite weight lifters and bodybuilders may be greater. Double-split routines (i.e., two training sessions per day) are common (351,362,363) during preparatory training phases, which may result in 8 to 12 training sessions per week. Frequencies as high as 18 sessions per week have been reported in Bulgarian weight lifters (351). The rationale for this high-frequency training is that frequent short sessions followed by periods of recovery, supplementation, and food intake allow for high-intensity training via maximal energy utilization and reduced fatigue during exercise performance (337). Häkkinen and Kallinen (362) reported greater increases in muscle size and strength when training volume was divided into two sessions per day as opposed to one among female athletes. In addition, exercises performed by Olympic lifters (i.e., total body lifts) require technique mastery that may increase total training volume and frequency.

Elite power lifters typically train with frequencies of four to six sessions per week (364). The superior conditioning of these athletes allows them to tolerate such high-frequency programs. However, training at these high frequencies would result in overtraining in most individuals. It is important that the individual be able to tolerate the physical stress so that an overtraining phenomenon does not develop (365). The development of periodized training cycles uses variations in training frequency to alter the exercise stimulus and provide for recovery and enhance the exercise stimulus.

Using a training program of 3 days a week for all situations and sports is not ideal. It is the individual needs and goals that determine the amount of exercise required to increase a particular physiologic or performance variable. Progression in frequency is also a key component to resistance training. Frequency of training will vary, depending on phase of the training cycle, fitness level of the individual, goals of the program, and training history. Careful choices need to be made regarding the rest between training days. These choices are based on the planned progress toward specific training goals and the tolerance of the individual to the program changes made. If excessive soreness is present the morning after a training session, this may indicate that the exercise stress is too
demanding. If this is the case, the workout loads, sets, and/or rest periods between sets and training frequency need to be evaluated and adjusted.

The configuration of these variables results in the exercise stimulus for a particular workout. Workouts must be altered to meet changing training goals and to provide training variation. Within this paradigm for the description of resistance exercise workouts, careful control of various components can be gained in manipulating variables to create new and optimal training programs (366). Because so many different combinations of these variables are possible, an almost unlimited number of workouts can be developed. Understanding the influence and importance of each of the program variables in achieving a specific training goal is vital to creating the optimal exercise stimulus.

Each program must be designed to meet the individual’s needs and training goals with recognition of the individual’s initial fitness level. It is important to remember that evaluation of a fitness level (e.g., 1-RM strength test) is typically not done until it is known that the individual can tolerate the test demands so that the data generated are meaningful (367). One of the most serious mistakes made in designing a resistance workout is placing too much stress on the individual before it can be tolerated.

FLEXIBILITY

Importance
It is important to maintain a range of motion adequate to perform one’s desired activities. Lost range of motion can interfere with such functional activities as ambulation, self-care, or attendant care. Severe restrictions in range of motion may even produce complications such as skin breakdown. Even relatively small reductions in range may result in biomechanical accommodations that place abnormal stress on tissues elsewhere in the body that can induce secondary problems.

Although it is clear that restoration and maintenance of a functional range of motion is desirable, the benefits from greater flexibility are not clear. Proponents of stretching have claimed numerous positive effects, including the prevention of musculoskeletal injuries and improved performance in sports, reduced postexercise muscle soreness, and improved general well-being. Nevertheless, the objective support for these claims is limited (368). Pre-exercise stretching has been demonstrated to have no effect on the development of DOMS (369). It has also recently been demonstrated that running economy may even be better among those who have some lower-extremity tightness (370,371). This has been suggested to be related to the ability to better use elastic storage and recovery and minimize the need for muscle-stabilizing activity.

Recent studies have shown that when strength and power performance are important, static and proprioceptive neuromuscular facilitation (PNF) stretching should not be done immediately prior to the event or training activity since performance will be impaired (372,373). Under such circumstances, static stretching for improvements in flexibility should be done at the end of an exercise session as part of a cool down or at another time separate from the strength and power exercises.

Factors Affecting Range of Motion
A number of factors can limit joint range of motion, including tightness of soft-tissue structures such as muscle, tendon, ligament, and joint capsule. Involuntary muscle contraction in the form of spasm can also restrict range. The bony contour of the joint is important in determining the full range of motion. When there is abnormal bone growth around a joint, range can be restricted. In addition, intra-articular loose bodies (e.g., bone or cartilage) and excessive fluid can restrict joint range of motion.

Range of motion varies widely among individuals. Regular activities using a full range of motion will help maintain range, but the maintenance of range of motion is specific to the joints that are used. For instance, an individual can have normal range in one joint but have severely restricted range in another. When connective tissue is not stretched, the collagen component gradually shortens. As a result, the periarticular collagen and the connective tissue of the muscle shorten. Furthermore, immobilization of a muscle in a shortened position also causes a decrease in the muscle length through a decrease in the number of sarcomeres in the muscle (374).

Age and sex also seem to affect range of motion. Women tend to have greater range of motion than men, and young people usually have greater range than the elderly. Tissue temperature is another factor affecting range of motion, with warm tissue having greater distensibility than cool tissue (375–377).

Techniques to Assess Range of Motion
Range of motion may be assessed in various ways. Angle measurements can be made with a goniometer, electrogoniometer, or flexometer (358). Flexibility for some movements is also commonly assessed through distance measurements between specific reproducible reference points. One example of this technique would be the assessment of temporomandibular joint motion through measurement of the distance between the upper and the lower incisors.

Methods to Improve Range of Motion

Techniques
Restoration of joint range of motion and soft-tissue extensibility can be achieved through the use of several different stretching techniques and modalities. The three general categories of stretching techniques include ballistic, static, and PNF procedures.

Ballistic stretching is characterized by repetitious bouncing movements, where the momentum of a moving body segment is used to generate forces producing a rapid stretch. Fast rates of stretching, as is the case with ballistic stretching, are not advisable, particularly during the early phases of rehabilitation. When fast stretch rates are used, greater tensions are developed,
and more energy will be absorbed within the muscle-tendon unit for a given length of stretch (378,379). As a result, there is a greater risk for injury with this type of stretching. Furthermore, ballistic stretching does not appear to be as effective in enhancing range as other methods of stretching (380,381).

Static stretching involves a slowly applied stretch that is held for several seconds. Proponents of this technique believe that the muscle stretch reflex is minimized through a slow, progressive stretch (382). Static stretching is generally easy to perform, can be done voluntarily or received passively by the individual, and has little associated risk of injury. Although the optimal time to hold a static stretch and number of stretches that should be performed is not known, experimental studies on animals suggest that most stress relaxation takes place during the initial 12 to 18 seconds and that there is little alteration in muscle-tendon unit lengthening after the fourth stretch (379).

Several stretching techniques use the principles of PNF that were developed by Kabat (383) to aid in the rehabilitation of injured World War II veterans. The concept of these stretching techniques is to enhance relaxation of the muscle to be stretched through reciprocal inhibition and the stretch reflex. The two most common PNF stretching techniques are the contract-relax technique and the agonist contract-relax technique. With the former, the muscle is passively stretched, then contracted for 6 to 8 seconds, and then relaxed and passively stretched further to an increased pain-free range. This process is repeated three to six times. Theoretically, the prestretch contraction of the muscle results in inhibition through Golgi tendon organ reflexes.

The agonist contract-relax technique is identical to the contract-relax technique except that the stretch is accompanied by a submaximal contraction of the opposing muscle to the one being stretched. This voluntary contraction of the opposing muscle theoretically results in reciprocal inhibition of the stretched muscle.

When able to participate, the patient can provide valuable feedback to ensure that appropriate positioning is used so that the desired tissues are being stretched. The patient should also understand that some discomfort may be required for adequate stretching, but prolonged poststretching pain is indicative of an overzealous approach. It also seems to be of considerable importance for the patient and the therapist to understand that the most rapid gains in range of motion will be achieved through regular stretching. In some cases, the performance of stretching exercises several times each day is desirable, but increases in range of motion can be achieved through three to five sessions per week (380,381).

**Modalities**

Some modalities may be used to enhance range of motion. Because the distensibility of warm tissue is greater than that of cool tissue (375–377), heat is commonly used before stretching. Ultrasound is the best therapeutic modality for heating deep-lying tissues (384) and can be used effectively as a prestretching treatment when there are no contraindications for its use. One should also take advantage of the elevation in tissue temperature that can be achieved through exercise. For example, the quadriceps intramuscular temperature can be elevated by 2°C after 10 minutes of cycling at a moderate intensity (Fig. 61-28).

Simultaneous stretching and brief application of fluorimethane spray (stretch and spray) or passage of ice along the course of the muscle being stretched has also been advocated as a means to enhance the stretch of muscles (385). When an active trigger point is present, stretching may be enhanced by injection of a small amount of local anesthetic into the trigger point before stretching. The reader is referred elsewhere (385) for a description of the techniques for stretch and spray and trigger point injections.

**PROPRIOCEPTION**

**Importance**

Proprioceptive organs provide the central nervous system with information about the position and movement of body parts. Recent work has demonstrated the importance of proprioceptive sensation in the rehabilitation after injuries, and in understanding the etiologies and predisposing factors to injuries and reinjuries.
Decreased joint proprioception has been observed in patients with osteoarthritis, rheumatoid arthritis, and Charcot’s disease and is believed to influence the progressive joint deterioration with these disorders (386,387). The risk of falling may also relate to proprioception as measures of postural sway have been demonstrated to relate to the falling risk (388–391).

Much of the work demonstrating the importance of proprioception in rehabilitation comes from the sports medicine literature. The importance of improving proprioception has been demonstrated by the finding of a significantly greater risk of sustaining an ankle injury during soccer among players with abnormal proprioceptive testing compared with those players with normal values (392). The successful return to sport after a ligament injury may be even more dependent on proprioception than on ligament tension (393).

Factors Affecting Proprioception
Sensation of the position and movement of a joint arise from afferent information originating from muscles, joint capsules, ligaments, and skin. As a result, injuries or disease to these structures may affect proprioception. For instance, reduced proprioception has been shown to be associated with knee (394) and ankle (395–397) ligament injuries, rheumatoid arthritis (398), and osteoarthritis (386). The importance of muscle receptors in proprioception has been demonstrated through the finding of impaired reproduction of joint angle after fatiguing contractions of the associated muscles (399).

Disease of the neurologic system may also affect proprioception. Diabetics with cutaneous sensory neuropathy have been shown to have a significant loss of ankle joint proprioception (400). Decreases in proprioception have also been found with increasing age (386,398,401,402) and may be part of the normal aging process.

Techniques to Assess Proprioception
Methods for accurate and objective assessment of joint movement and position perception have not been adequately developed for widespread clinical use. The most common clinical method of evaluating proprioception is a nonquantitative test of joint position perception consisting of the subject verbally describing the joint position after the examiner passively moves the segment into one of two or more predetermined positions (403). Another nonquantitative testing technique used in the clinical setting involves testing the ability of a person to return a joint to a predetermined joint position.

In the research setting, a number of techniques have been used in the evaluation of proprioception. These research techniques include using an isokinetic dynamometer to quantify angles during reproduction of a predetermined joint position (404,405), matching of joint position with the side contralateral to the one being passively moved or positioned (406–408), and measurement of joint movement perception threshold (400,404,409). Although not exclusively a measure of proprioception, other techniques that may provide related information about balance and coordination include measures of postural sway (410,411) and responses to perturbations that create sway (412).

Techniques to Improve Proprioception
Proprioceptive exercises are performed with the goal of reducing the proprioceptive deficits that may have resulted through injury or disease. There is some evidence for the effectiveness of proprioceptive exercises in that it has been shown that the proprioceptive deficits present after ligament injuries can be reduced (413–415).

Classic lower-extremity proprioceptive exercises have used the tilt (teeter or wobble) board. Unidirectional boards may be used initially, followed by multidirectional boards. More functionally specific proprioceptive exercises can also be developed. These activities may include backward and sideways walking and running, or other agility drills.

It should also be noted that elastic bandaging (386) has been shown to improve position sense in those with impaired proprioception. This may be achieved through enhancement of the activity of the skin proprioceptors.

THE EXERCISE PRESCRIPTION
General Considerations
Appropriate exercise prescriptions should systematically and individually recommend physical activity to develop fitness, maintain health, and/or treat specific conditions. Basic components of an aerobic exercise prescription include exercise mode, intensity, duration, frequency, and rate of progression. Exercise sessions should begin with appropriate warm-ups and end with a cool-down; these periods of reduced intensity improve performance and likely decrease risk of injury.

Although the specific goals of an exercise program depend on the individual, basic exercise goals include (a) counteracting the detrimental effects of sedentary living or reduced activity (e.g., from disease or injury), and (b) optimizing functional capacity within the physical limitations of present medical conditions. Additionally, exercise training programs can provide valuable clinical information for the ongoing treatment of a patient. Using these basic goals as a foundation, an individual’s specific health history, risk factors, behavioral characteristics, personal goals, and exercise preferences will shape and detail the exercise prescription.

Risk Stratification
Exercise prescription for middle-aged and older people must consider the risk of cardiac complications associated with exercise. Well-controlled exercise presents very little risk, with a low incidence of sudden death (approximately one death per year for every 15,000 to 18,000 individuals). In people older than 35 to 40 years of age, incidents usually occur in the presence of coronary artery disease and vigorous exercise (416,417). Despite this minor risk, the benefits of exercise conditioning far outweigh the remote possibility of sudden cardiac complications (96,417–419). Furthermore, regular exercise likely protects against sudden death or myocardial infarction from strenuous physical activities (61,66,96,418).
A clinical identification of coronary risk factors crucially guides program development. Table 61-7 displays the coronary artery disease risk factor thresholds. The ACSM considers the following risk categories (99):

- **Low-risk** individuals are asymptomatic younger adults (men <45 years of age, women <55 years of age) who meet or exceed one risk factor threshold or less. These people can safely begin exercise programs without further assessment (Table 61-8).
- **Moderate-risk** individuals are older individuals (men ≥55 years of age, women ≥55 years of age) or those who meet the threshold for two or more risk factors. These people require further assessment and a medical examination only if they intend to initiate an exercise program containing vigorous activities (i.e., >60% $\dot{V}O_2_{max}$).
- **High-risk** individuals possess one or more signs/symptoms from Table 61-9 or have known cardiovascular (cardiac, peripheral vascular, or cerebrovascular), pulmonary (chronic obstructive pulmonary disease, asthma, interstitial lung disease, or cystic fibrosis), or metabolic disease (DM, thyroid disorders, renal, or liver disease). These people should complete a thorough medical examination and exercise test before starting any exercise program.

Cardiac rehabilitation programs safely and effectively initiate exercise for individuals recently suffering cardiac events (420). Surveys indicate low morbidity and mortality rates associated with supervised exercise programs (421–425). Although

### TABLE 61.7  **Coronary Artery Disease Risk Factor Thresholds**

<table>
<thead>
<tr>
<th>Risk Factors</th>
<th>Defining Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Positive</strong></td>
<td></td>
</tr>
<tr>
<td>Family history</td>
<td>Myocardial infarction, coronary revascularization, or sudden death before 55 years of age in father or other male first-degree relative (i.e., brother or son), or before 65 years of age in mother or other female first-degree relative (i.e., sister or daughter)</td>
</tr>
<tr>
<td>Cigarette smoking</td>
<td>Current cigarette smoker or those who quit within the previous 6 mo</td>
</tr>
<tr>
<td>Hypertension</td>
<td>Systolic blood pressure ≥140 mm Hg or diastolic ≥90 mm Hg, confirmed by measurements on at least two separate occasions, or on antihypertensive medication</td>
</tr>
<tr>
<td>Dyslipidemia</td>
<td>Low-density lipoprotein (LDL) cholesterol &gt;130 mg·dL$^{-1}$ (3.4 mmol·L$^{-1}$) or HDL cholesterol &lt;40 mg·dL$^{-1}$ (1.03 mmol·L$^{-1}$), or on lipid-lowering medication. If total serum cholesterol is all that is available use &gt;200 mg·dL$^{-1}$ (5.2 mmol·L$^{-1}$) rather than LDL &gt;130 mg·dL$^{-1}$</td>
</tr>
<tr>
<td>Impaired fasting glucose</td>
<td>Fasting blood glucose ≥100 mg·dL$^{-1}$ (5.6 mmol·L$^{-1}$) confirmed by measurements on at least two separate occasions</td>
</tr>
<tr>
<td>Obesity$^d$</td>
<td>Body mass index ≥30 kg·m$^{-2}$, or waist girth &gt;102 cm for men and &gt;88 cm for women, or waist/hip ratio: ≥0.95 for men and ≥0.86 for women</td>
</tr>
<tr>
<td>Sedentary lifestyle</td>
<td>Persons not participating in a regular exercise program or not meeting the minimal physical activity Recommendations$^b$ from the U.S. Surgeon General’s report</td>
</tr>
<tr>
<td><strong>Negative</strong></td>
<td></td>
</tr>
<tr>
<td>High serum HDL cholesterol$^c$</td>
<td>&gt;60 mg·dL$^{-1}$ (1.6 mmol·L$^{-1}$)</td>
</tr>
</tbody>
</table>

$^d$Professional opinions vary regarding the most appropriate markers and thresholds for obesity and therefore, allied health professionals should use clinical judgment when evaluating this risk factor.

$^c$Accumulating 30 min or more of moderate physical activity on most days of the week.

$^b$It is common to sum risk factors in making clinical judgments. If HDL is high, subtract one risk factor from the sum of positive risk factors, because high HDL decreases the risk of coronary artery disease.


### TABLE 61.8  **Recommendations for Medical Examination and Clinical Exercise Testing Prior to Exercise Participation**

<table>
<thead>
<tr>
<th>Risk Level</th>
<th>Moderate Exercise</th>
<th>Vigorous Exercise</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low Risk$^a$</td>
<td>Not necessary</td>
<td>Not necessary</td>
</tr>
<tr>
<td>Moderate Risk$^a$</td>
<td>Not necessary</td>
<td>Recommended</td>
</tr>
<tr>
<td>High Risk$^a$</td>
<td>Recommended</td>
<td>Recommended</td>
</tr>
</tbody>
</table>

$^a$See text for description of risk categories.

$^c$Absolute moderate exercise is defined as activities that are ~3 to 6 METs or the equivalent of brisk walking at 3–4 mph for most healthy adults. Nevertheless, a pace of 3–4 mph might be considered to be “hard” to “very hard” by some sedentary, older persons. Moderate exercise may alternatively be defined as an intensity well within the individual’s capacity, one that can be comfortably sustained for a prolonged period of time (~45 min), which has a gradual initiation and progression, and is generally noncompetitive. If an individual’s exercise capacity is known, relatively moderate exercise may be defined by the range of 40% to 60% of maximal oxygen uptake.

$^b$Vigorous exercise is defined as activities of ≥6 METs. Vigorous exercise may alternatively be defined as exercise intense enough to represent a substantial cardiorespiratory challenge. If an individual’s exercise capacity is known, vigorous exercise may be defined as an intensity of >60% of maximal oxygen uptake.

injury and illness precautions

All exercises carry some risk of injury, but individualizing the exercise prescription minimizes the chance of causing or exacerbating health problems. Characteristics of the activity (e.g., weight bearing vs. non–weight bearing, intensity, volume, and progression) and the exerciser (e.g., biomechanical abnormalities or previous injury) influence injury potential. Persons predisposed to musculoskeletal injuries (due to obesity, previous injury, or biomechanical abnormality) should avoid weight-bearing activities to reduce injury risk. Appropriate footwear and exercise progression also minimize the chance of injury.

Exercise also exerts a noticeable influence on immune function and decreased risk of infection compared to sedentary counterparts (189,191,192). The stress imposed by extreme exercise training (increased volume, intensity, or both) likely predisposes the exerciser to infection and illness, making the relationship between training intensity and risk of illness a "J"-shaped curve (188,190). How exercise affects the progression of existing infectious diseases like viral respiratory infections is poorly understood. As myocarditis can develop when individuals exercise in the presence of such illnesses, individuals should avoid strenuous exercise when suffering from a fever or myalgia (429,430).

Environmental Factors

Hot and/or humid conditions increase the physiological challenge and health risk associated with exercise. Combining exercise and environmental heat stress produces a competition for blood flow between active muscle (to deliver oxygen and nutrients) and cutaneous vascular beds (to dissipate heat), reducing exercise performance and/or thermoregulatory capacity (431,432). Decreases in total body water (i.e., hypohydration) exacerbate this situation (433–435). The exertional heat illnesses (heat cramps, heat syncope, heat exhaustion, and heat stroke) occur more frequently in competitive events than recreational exercise, as the drive to maximize performance can override normal signals to decrease exercise intensity. Wise race organizers minimize these potential health hazards by warning participants about environmental conditions, scheduling events during the least stressful times of day, offering appropriate aid stations on the course, and if necessary, cancelling the event (436).

During exercise in the heat, people should (a) monitor physiological markers of stress (heart rate and rating of perceived effort typically rise relative to heat load (77,80)), (b) monitor mental state (confusion, disorientation, and/or irrational behavior strongly suggests the presence of a serious heat illness (437)), and (c) attempt to minimize hypohydration by matching sweat loss with fluid intake (438,439). Additionally, anyone planning to compete in a hot environment should acclimatize before competition. Daily exercise in the heat for 7 to 10 days significantly improves performance and thermoregulation. These adaptations occur secondary to improved sweating (earlier onset and increased capacity), decreased loss of electrolytes, and increased blood volume. As a consequence, acclimatization reduces physiologic strain during exercise in the heat as evidenced by lower heart rates and core temperatures at similar exercise intensities (77). For further details, see the ACSM position stand describing heat illnesses during exercise (437).

In extremely cold environments, exercisers must protect against hypothermia and cold injuries to exposed skin (frostbite and frostnip). Cold injuries occur more infrequently than heat illnesses because exercise produces large metabolic heat loads, defending core temperature against all but the coldest temperatures. In most conditions, appropriately layered clothing and protection of sensitive areas (e.g., hands, feet, and face) allow exercise to continue comfortably. After repeated exposures to

### TABLE 61.9 Major Signs or Symptoms Suggestive of Cardiovascular, Pulmonary, or Metabolic Disease

<table>
<thead>
<tr>
<th>Sign/Symptom</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pain, discomfort (or other anginal equivalent) in the chest, neck, jaw, arms, or other areas that may result from ischemia</td>
<td></td>
</tr>
<tr>
<td>Shortness of breath at rest or with mild exertion</td>
<td></td>
</tr>
<tr>
<td>Dizziness or syncope</td>
<td></td>
</tr>
<tr>
<td>Orthopnea or paroxysmal nocturnal dyspnea</td>
<td></td>
</tr>
<tr>
<td>Ankle edema</td>
<td></td>
</tr>
<tr>
<td>Palpitations or tachycardia</td>
<td></td>
</tr>
<tr>
<td>Intermittent claudication</td>
<td></td>
</tr>
<tr>
<td>Known heart murmur</td>
<td></td>
</tr>
<tr>
<td>Unusual fatigue or shortness of breath with usual activities</td>
<td></td>
</tr>
</tbody>
</table>

a cold environment, the body does exhibit adaptations that enhance heat production or minimize heat loss; however, cold acclimatization causes far less dramatic physiological changes than heat acclimatization (434).

Exposure to moderate and high terrestrial altitudes also challenges exercise. Altitude exposure reduces work capacity because the lower barometric pressure directly translates to less available oxygen. This effect becomes noticeable at approximately 1,500 m; each additional 1,000 m of ascent further reduces exercise capacity by 6% to 10%. Available information suggests cardiac patients can safely exercise to at least moderate altitudes (440–444). Because altitude exposure limits exercise capacity, the absolute work intensity that elicits a prescribed index of myocardial work (i.e., rate-pressure product) decreases as altitude increases, suggesting heart rate and perceived effort provide more appropriate indices of exercise intensity than absolute work units at altitude. Similar to other environmental stressors, chronic exposure to altitude motivates physiological changes that benefit exercise performance (primarily increased oxygen-carrying capacity of the blood) and reduce the incidence of altitude-related illnesses (e.g., acute mountain sickness).

Air pollution can also detrimentally affect exercise performance. Several air pollutants (e.g., particulate matter, carbon monoxide, sulfur dioxide, nitrogen oxides, and ozone) negatively impact work tolerance or cardiovascular responses to exercise (445,446). Further, exercise-induced increases in pulmonary ventilation expose the exerciser to larger doses of pollutants than experienced at rest. As strong evidence now suggests a link between air pollution and cardiovascular disease (447,448), individuals should avoid exercising (a) in highly polluted areas or near traffic congestion, and (b) in environments where ozone reaches a critical threshold. These recommendations especially apply to those with coronary artery or pulmonary disease.

Other environmental stressors also impact certain exercises. For instance, unique terrain or wind might substantially increase physiologic demands (449,450) and require a reduction in absolute work rates. Additionally, exercise prescription for potentially unstable individuals must always consider the risk of falling.

**Equipment**

Many specialized devices exist to develop strength, aerobic fitness, flexibility, and proprioception. Although clinics frequently house this potentially advantageous equipment, patients must possess all necessary equipment to complete their exercise programs outside the clinical setting. Fortunately, small modifications to most exercises reduce or completely eliminate equipment needs. For example, strength training could employ body weight, heavy household items, or elastic tubing for the resistance. Walking and running efficiently develop aerobic fitness (211) and require no equipment other than appropriate footwear. Additionally, some individuals might have access to community resources (e.g., swimming pool) or own specific exercise equipment. Regardless of the specific circumstances, prescribing exercise that patients can readily complete increases compliance (see the following section) and program effectiveness.

**Compliance**

The extent of the challenge in getting people to exercise is evident when considering that approximately one in four adults in the United States report no leisure-time activity, and less than half meet physical activity recommendations (96,100). Most teenagers do not participate in vigorous activities, and about 50% do not participate in physical education classes at school (96,96). Sedentary living occurs more commonly among Americans with less education and economic or social disadvantages (96,451). To achieve the exercise objectives established in the Healthy People 2010 Program, Americans must considerably change their physical activity participation. Although evidence suggests an improved public awareness of the benefits of regular physical activity, this awareness has not yet translated into substantial improvements in physical activity among sedentary individuals (452–455). The recent release of the Surgeon General’s report on physical activity emphasizes the importance of increasing physical activity among all Americans (456).

When using exercise as a treatment, it is important for the patient to understand and accept that they are not a passive participant as is commonly the case in the classic medical model. This significant difference challenges many individuals, and exacerbates the compliance issue. Instilling a sense of self-involvement, an understanding of exercise fundamentals, and a basic idea how exercise may alleviate the patient’s condition drastically improves compliance and exercise program success.

Patients who misunderstand the differences between exercises designed to develop strength, flexibility, and aerobic fitness are more likely to question the value of the exercise and less likely to complete the exercise correctly. For example, patients must comprehend that strength increases when muscles voluntarily contract against resistance and aerobic fitness improves when movement raises heart rate and ventilation. Regardless of the exercise, clinicians should (a) inform patients that success likely requires some discomfort, and (b) teach patients to recognize the difference between signs of exercise stress (e.g., increased heart rate and muscle soreness) and symptoms of overexertion or injury. Patients must especially comprehend these concepts during the initial few days of exercise program, when novel symptoms or DOMS occurs most frequently and discouragingly. Postexercise cryotherapy and reduced initial exercise intensities limit painful symptoms and DOMS; this attenuation of discomfort likely improves adherence when starting an exercise program. After these first few days, patients must also appreciate that adaptations resulting from an exercise program take time (usually a minimum of several weeks) and only occur with a commitment to a regular and consistent routine. The patient’s personality, the rehabilitation setting, and the clinician-patient relationship all influence adherence rates (457). Of these, patient motivation most strongly influences success rate. Education, goal setting, and regular follow-up
(with appropriate exercise modification and progression) play important roles in developing and maintaining motivation.

Other than these general recommendations, three specific techniques maximize compliance. First, adherence improves when the program consists of activities the individual enjoys. Personal preferences relative to exercise mode, intensity, frequency, and duration should be considered (96). A second technique involves encouraging people to build physical activity into their normal routines and/or to participate in sport activities instead of (or in addition to) more traditional aerobic modes such as running or cycling. Increasing the exercise stress of typical home activities (e.g., using a walk-behind mower rather than a rider mower, walking stairs rather than riding the elevator) conveniently and inexpensively increases physical activity (96,458). When combined with a more traditional exercise program, home and recreational activities can add welcome diversity and significantly enhance adherence. Finally, attitudinal and behavioral factors within different population subgroups influence motivation for physical activity (96,459–461); individualizing exercise prescription based on ethnic background, socioeconomic status, gender, age, and disease also improves compliance. Individuals largely resist physical activities they consider inappropriate (e.g., because of age, gender, or mode), so practitioners should tailor the exercise prescription to clearly describe the value of the exercise or completely avoid undesirable exercises.

In summary, exercise compliance occurs most frequently when the individual perceives a benefit, enjoys (or minimally accepts) the activity, feels safe and competent, and recognizes the benefit of performing daily activities in ways that increase the physiological stress. Further benefits to compliance occur when exercise conveniently fits into the daily schedule and costs little (96). Much work remains to develop strategies for optimizing exercise compliance, especially among low physically active subpopulations such as the socioeconomically deprived (96,453,454).

**EXERCISE PRESCRIPTION IN SPECIAL POPULATIONS**

**Considerations in Chronic Disease States**

Exercise effectively assists in managing numerous medical and disabling conditions. Many individuals treated in the rehabilitation setting with exercise also suffer concomitant medical conditions or disabilities that affect the ability to perform and response to specific exercises. Accordingly, effective program design requires special considerations for people with disease or disability. In general, the exercise program must not interfere with the standard medical treatment and must vary according to the presence or severity of disease. The following text discusses important considerations for medical conditions and disabilities commonly encountered in rehabilitation settings. Several other specific chapters also discuss relevant disease states.

**Cardiac Disease**

Cardiac disease occurs commonly among individuals in the rehabilitation setting. Although cardiovascular disease generally does not preclude exercise, those prescribing exercise must recognize the presence and extent of disease to design and initiate a safe program. Risk stratification (discussed earlier) significantly assists this process.

General guidelines for aerobic exercise apply to patients with stable cardiac disease. Patients with ischemic changes, angina, or arrhythmias during exercise should exercise at an intensity 10 to 15 beats per minute below the ischemic, angina, or dysrhythmic threshold (99). Individuals at high risk for exercise-induced cardiac emergencies (Table 61-10) require a more cautious application of exercise intensity and professional supervision, potentially including electrocardiographic and blood pressure monitoring. This population should complete gradual and prolonged cool-downs to reduce the risk of arrhythmias and postexercise hypotension.

Most low- to moderate-risk patients with stable cardiac disease can also safely perform and benefit from resistance exercise (462,463). Exclusion criteria precluding resistance training include congestive heart failure, severe valvular disease, uncontrolled hypertension, uncontrolled dysrhythmias, and other unstable symptoms (99). When monitoring exercise intensity during resistance exercise, rate-pressure product likely better indicates ischemic threshold than heart rate because resistance exercises frequently generate large blood pressures. Accurate determination of the rate-pressure product requires the measurement of blood pressure during muscular actions, as blood pressure decreases rapidly during relaxation.

**Diabetes**

Problems associated with diabetes (e.g., cardiovascular disease, peripheral neuropathy, peripheral vascular disease, autonomic

### TABLE 61.10 Characteristics of Cardiac Patients Associated with High Risk

<table>
<thead>
<tr>
<th>Characteristics obtained at rest</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Resting ejection fraction &lt;40%</td>
<td></td>
</tr>
<tr>
<td>History of cardiac arrest</td>
<td></td>
</tr>
<tr>
<td>Complex dysrhythmias</td>
<td></td>
</tr>
<tr>
<td>Complicated myocardial infarction or revascularization procedures</td>
<td></td>
</tr>
<tr>
<td>Presence of congestive heart failure</td>
<td></td>
</tr>
<tr>
<td>Presence of signs or symptoms of postevent/postprocedure ischemia</td>
<td></td>
</tr>
<tr>
<td>Presence of clinical depression</td>
<td></td>
</tr>
<tr>
<td>Characteristics obtained during exercise or recovery</td>
<td></td>
</tr>
<tr>
<td>Presence of complex ventricular dysrhythmias</td>
<td></td>
</tr>
<tr>
<td>Presence of angina or other significant symptoms (e.g., unusual shortness of breath, light headedness, or dizziness), especially at low levels of exertion (&lt;5 METs)</td>
<td></td>
</tr>
<tr>
<td>High level of silent ischemia (ST-segment depression ≥2 mm from baseline)</td>
<td></td>
</tr>
<tr>
<td>Presence of abnormal hemodynamics (i.e., chronotropic incompetence or flat or decreasing systolic blood pressure with increasing workload or severe postexercise hypotension)</td>
<td></td>
</tr>
</tbody>
</table>

dysfunction, renal disease, and retinopathy) can significantly complicate the development of an exercise program. These complicating factors suggest all diabetics should complete a thorough medical examination before initiating an exercise program (99).

Because exercise increases glucose uptake independent of insulin, exercising diabetics frequently experience hypoglycemia (99). Excessive volumes or absorption rates of insulin cause hypoglycemia, and this may typically occur when patients inject short-acting insulin near active muscles. Type 1 (insulin-dependent or juvenile-onset) diabetics must therefore adequately regulate their insulin before initiating an exercise program. Table 61-11 lists important recommendations for all diabetics that minimize the risk of hypoglycemia.

Although general guidelines for aerobic exercise apply to most diabetics, several symptoms and conditions associated with diabetes pose unique challenges. The prevalence of peripheral neuropathy affecting the feet suggests diabetics should pay special attention to proper footwear, foot hygiene, and weight-bearing exercise (especially obese diabetics). Impaired sweating and polyuria-induced hypohydration limit thermoregulatory capacity, so diabetics should consider their exercise environment. Diabetics using β-blocking agents may not experience hypoglycemic symptoms and/or angina. Autonomic neuropathy and chronotropic insufficiency complicate the relationship between heart rate and exercise intensity, suggesting that diabetics suffering from these conditions should monitor and establish exercise intensity with perceived exertion. Finally, advanced retinopathy precludes exercise modes that cause excessive jarring or marked increases in blood pressure.

### TABLE 61.11

<table>
<thead>
<tr>
<th>Recommended Precautions for Diabetics to Minimize the Risk of Hypoglycemic Reactions Associated with Exercise</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monitor blood glucose prior to and following exercise, especially when beginning or modifying the program.</td>
</tr>
<tr>
<td>Understand that a late-onset hypoglycemia can occur up to 48 h postexercise, especially when beginning or modifying the program.</td>
</tr>
<tr>
<td>Know the signs and symptoms of hypoglycemia (e.g., drowsy, tremors, etc.).</td>
</tr>
<tr>
<td>Adjust carbohydrate intake or insulin injections prior to exercise based on blood glucose and exercise intensity. 20–30 grams of additional carbohydrate should be ingested if pre-exercise blood glucose is &lt;100 mg·dL⁻¹.</td>
</tr>
<tr>
<td>Avoid injecting insulin into exercising limbs; an abdominal injection site is preferred.</td>
</tr>
<tr>
<td>When exercising late in the evening, an increased consumption of carbohydrates might be required to minimize the risk of nocturnal hypoglycemia.</td>
</tr>
<tr>
<td>Postexercise hypoglycemia might occur in the hours following basic (not intense) resistance exercise for patients on insulin or oral hypoglycemic agents.</td>
</tr>
</tbody>
</table>


**Obesity**

Body mass varies according to only two factors: calories consumed (via nutritional intake) and calories expended (via basal metabolism, exercise, and thermic effect of food). Body mass decreases when caloric expenditure exceeds caloric consumption; a mild caloric restriction combined with regular exercise best accomplishes this goal (146,464). Additionally, increases in physical activity reduce the risk of associated comorbidities independent of body mass loss, making exercise important even if weight fails to change.

Overweight individuals generally lead sedentary lives and maintain negative connotations about exercise. A thorough exploration of these individuals’ exercise histories crucially assists in combating psychological barriers and developing exercise compliance. The significant biomechanical stresses extra weight places on the body also predispose obese individuals to orthopedic injuries; exercises that minimize joint stress (e.g., walking, cycling, rowing, stair climbing, and exercise in water) attenuate these concerns.

**Arthritis**

Exercise programs for individuals with arthritis vary according to disease state; when disease activity peaks, minimizing activity might avoid tissue damage. Programs for arthritis generally rely more heavily on shorter, more frequent exercise sessions, non–weightbearing or low-impact activities (e.g., swimming and cycling), and flexibility exercises to motivate positive adaptations while concurrently minimizing joint stress (99,465). Resistance training also optimizes strength, muscular endurance, and joint stability in this population.

**Peripheral Arterial Disease**

Significant leg pain during physical activity (claudication) characterizes peripheral arterial disease, a condition typically associated with hypertension, hyperlipidemia, coronary artery disease, cerebrovascular disease, and diabetes. These frequent comorbidities generally place individuals with peripheral arterial disease at high risk for medical complications during exercise, suggesting these patients complete exercise testing before initiating a structured, medically supervised exercise program. Because treadmill walking and cycling cause significant pain, arm ergometry might most effectively achieve adequate myocardial stress during a descriptive test.

Of all exercise modes, walking best induces functional changes in those with peripheral arterial disease. Rather than continuous exercise, most experts (99,466) recommend intermittent exercise sessions; these repeated shorter exercise bouts induce symptoms that subside during brief, intervening rest periods. This method applies a suitable exercise volume while still accommodating claudication pain. Initial sessions of 35 minutes should progress to 50 to 60 minutes with the eventual goal of walking continuously for a similar duration. As functional capacity improves, increases in exercise intensity will induce beneficial central cardiovascular adaptations.
Spinal Cord Injury

Spinal cord injury affects exercise capacity by altering the amount of functional muscle mass, compromising autonomic nervous control of cardioacceleration, redistributing blood flow, and limiting thermoregulation. Each issue affects exercise capacity, potential training adaptations, and safety during exercise. Despite these physiologic limitations, spinal cord-injured people safely participate in many activities (e.g., long-distance wheelchair propulsion, swimming, kayaking, and skiing) and commonly complete exercise programs of wheelchair propulsion or arm cranking ergometry.

Depending on the characteristics of the lesion, spinal cord injury can limit functional muscle mass to the upper body. As the most effective aerobic exercise rhythmically recruits large muscle groups, this reduction in available muscle mass limits training-induced aerobic demands and therefore alters potential adaptations. Paraplegics can exercise intensely and adapt to exercise training (as demonstrated by the clearly positive effect of upper body endurance training on peak and work capacity during exercise), but most adaptations occur peripherally (206,208). In other words, spinal cord injury might severely limit the adaptive potential of central cardiovascular structure and function. Recent research suggests several techniques, including lower-body compression (467–469), functional electrical stimulation of the paralyzed lower limbs (470,471), or supine body position concomitant with upper-body exercise (472), might enhance venous return and cardiac output, thereby increasing the opportunity for central training adaptations.

Complete spinal cord lesions also affect the autonomic nervous system. Loss of sympathetic cardiac innervation from lesions above the sixth thoracic vertebra can limit maximal heart rate to 110 to 130 beats per minute (473). Lesions at the cervical and thoracic levels can impair control over regional blood flow during exercise, causing venous blood pooling in the legs and consequently reducing cardiac preload. As a result, stroke volume and cardiac output at a given oxygen uptake tend to decline in those with spinal cord injury (474,475). Furthermore, the loss of sympathetic nervous control over vasomotor and sudomotor responses of the insensitive skin impairs thermoregulation (476).

Individuals with spinal cord injury should complete exercise testing before initiating an exercise program for two reasons in addition to screening for coronary artery disease. Testing also helps identify cardiovascular issues that might otherwise remain undetected because some quadriplegics lack classical symptoms of angina pectoris (as most cardiac visceral afferents enter the spinal cord at the upper thoracic levels). Additionally, exercise testing can identify and assist in the treatment of exercise-induced hypotension. Assuming the prescription accounts for these potential autonomic, cardiovascular, and thermoregulatory conditions, the intensity and duration guidelines for cardiovascular training in the apparently healthy appear reasonable for the spinal cord injured (473). Overall risk of serious problems from participating in appropriately structured exercise programs appears minimal for the spinal cord injured.

Postpoliomyelitis Syndrome

Although many professionals feared those who suffered postpolio sequelae might experience overuse-induced muscle weakness with exercise, most evidence suggests these individuals can safely exercise without adverse effects. Individuals with postpolio sequelae generally respond to endurance (477–480) and resistance (481) training like healthy adults. Nevertheless, these individuals require careful individual monitoring and periodic muscle function tests (479).

Multiple Sclerosis

Most concerns regarding exercise safety in multiple sclerosis patients center around potential adverse effects from the autonomic dysfunction that frequently accompanies the disease and exacerbations of the disease via exercise-induced thermal stress. Although relatively little research examines these questions, a recent review article (482) concluded that endurance and resistance exercise training beneficially affect moderately impaired multiple sclerosis patients. Encouraging evidence from individual research articles supports these conclusions: individuals with multiple sclerosis can perform endurance exercise at intensities above the anaerobic threshold without developing significant neurologic symptoms (483), experience training-induced cardiovascular adaptations (484), and gain significant strength via resistance training (485). Until more research clarifies how people with multiple sclerosis respond to exercise, these patients (especially those suffering more than moderate impairments) deserve a conservative exercise prescription.

Cancer

The medical community increasingly accepts exercise as an important component of rehabilitation for cancer patients. Besides counteracting the effects of inactivity and improving psychological status (486), evidence suggests moderate exercise improves immune function (188–192). Although general guidelines for aerobic exercise training generally meet their demands (487), cancer patients should typically exercise at lower intensities. Patients with bone malignancies (particularly of the spine, pelvis, femur, and ribs) require non–weight-bearing modalities. Individuals suffering associated conditions that increase the risk of bruising, fractures, or falls should perform resistance exercise on machines rather than with free weights to minimize the risk of injury.

Besides the disease itself, the treatments for cancer challenge exercise capacity and prescription. Cancer treatments can induce cytotoxicity, immunosuppression, bleeding disorders, and anemia. Further, direct or peripheral effects of cancer or its treatments inhibit exercise performance by causing (a) direct damage to cardiac and pulmonary tissue, (b) difficulty in maintaining adequate nutrition, hydration, and electrolyte balance, and (c) fatigue and infection.

Considerations in Selected Groups within the Able-Bodied Population

Pregnant women, children, and the elderly possess unique physical, physiological, and behavioral characteristics that merit consideration in the design of exercise programs.
TABLE 61.12  Contraindications for Exercising During Pregnancy

<table>
<thead>
<tr>
<th>Relative contraindications</th>
<th>Absolute contraindications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Severe anemia</td>
<td>Hemodynamically significant heart disease</td>
</tr>
<tr>
<td>Unevaluated maternal cardiac arrhythmia</td>
<td>Restrictive lung disease</td>
</tr>
<tr>
<td>Chronic bronchitis</td>
<td>Incompetent cervix/cerclage</td>
</tr>
<tr>
<td>Poorly controlled type I diabetes</td>
<td>Multiple gestation at risk for premature labor</td>
</tr>
<tr>
<td>Extreme morbid obesity</td>
<td>Persistent second- or third-trimester bleeding</td>
</tr>
<tr>
<td>Extreme underweight</td>
<td>Placenta previa after 26 wk of gestation</td>
</tr>
<tr>
<td>History of extremely sedentary lifestyle</td>
<td>Premature labor during the current pregnancy</td>
</tr>
<tr>
<td>Intrauterine growth restriction in current pregnancy</td>
<td>Ruptured membranes</td>
</tr>
<tr>
<td>Poorly controlled hypertension</td>
<td>Preeclampsia/pregnancy-induced hypertension</td>
</tr>
</tbody>
</table>


**Pregnancy**

Previous concerns about exercise during pregnancy related to the competition between mother and fetus for oxygen and substrates, thermoregulatory issues, and uterine contractions. Currently, however, no evidence indicates that healthy pregnant women should limit exercise intensity for fear of adverse effects (99).

The American College of Obstetricians and Gynecologists has established contraindications for exercise during pregnancy (Table 61-12) and exercise recommendations for women without risk factors for adverse maternal or perinatal outcomes (Table 61-13). Equipped with an understanding of the signs and symptoms for discontinuing exercise (Table 61-14), pregnant women can safely exercise during pregnancy within these recommendations. Women who exercised regularly prior to becoming pregnant can continue their training with appropriate modifications; previously sedentary women or those managing medical complications should obtain physician approval before initiating a new exercise program during pregnancy (488).

**Children and Adolescents**

Decreasing levels of physical activity have contributed to the recent rise in childhood and adolescent obesity in the United States. To reverse this trend in inactivity, children and adolescents must experience positive physical activity/exercise opportunities. In developing physical activity programs for this group, consider their maturity, skill levels, medical conditions, and prior physical activity. Some general guidelines include (a) providing a variety of activities that exercise all major large muscles; (b) focusing on active, creative, enjoyable play in very young children; (c) encouraging children 5 to 12 years old to accumulate a minimum of 60 minutes of age-appropriate moderate or vigorous, intermittent activity on most or all days of the week; (d) encouraging older children to participate in an additional 20 to 30 minutes of higher-intensity activity at least three times per week; and (e) gradually progressing exercise volume (99). In addition to immediate health benefits, positive exercise experiences during youth translate into increased adulthood activity.

Safety represents a primary concern in designing exercise programs for children, as they are anatomically, physiologically,
TABLE 61.14 Reasons for Pregnant Women to Discontinue Exercise and Consult a Physician

1. Any signs of bloody discharge from the vagina
2. Any "gush" of fluid from the vagina (premature rupture of membranes)
3. Persistent contractions (>6 to 8/h) that may suggest onset of premature labor
4. Unexplained abdominal pain
5. Absence of or decreased fetal movement
6. Sudden swelling of the ankles, hands, and face
7. Persistent, severe headaches and/or visual disturbance
8. Unexplained spells of faintness or dizziness
9. Swelling, pain, and redness in the calf of one leg (phlebitis)
10. Elevation of pulse rate or blood pressure that persists after exercise
11. Excessive fatigue, palpatations, chest pain
12. Insufficient weight gain (<1.0 kg/mo during last two trimesters)
13. Dyspnea prior to exertion
14. Muscle weakness


and psychologically immature (99). This immaturity magnifies the risk of injury for any activity because the child’s body already combats the stress of growth and development. Additional concerns include an increased head-to-body size ratio, vulnerability of growing connective tissue, and underdeveloped motor skills (489). Using appropriate equipment, matching competition to maturation or skill, providing adequate skill preparation, and liberalizing rules all reduce the possibility of injury.

Additionally, children possess a less robust thermoregulatory capacity than adults: children begin sweating at a higher core temperature, sweat less at a given temperature, and acclimatize more slowly to the heat (490). An increased ratio of surface area to body mass also alters a child’s interaction with his or her environment. Limiting strenuous prolonged exercise in extreme environmental conditions, promoting good hydration practices, and encouraging appropriate clothing minimize the potential for thermoregulatory issues (99).

Like adults, children respond to resistance training by significantly increasing strength (99,229). Unlike adults, most strength gains result from neural development rather than frank muscle hypertrophy. Safe resistance training programs for children include close adult supervision, avoidance of maximal resistance loads, strong emphasis on proper lifting technique, and low-resistance high-volume workouts (99).

Elderly

Although basic exercise prescription models generally suffice for older adults, the elderly typically present with disease processes, orthopedic issues, and/or unique goals that require manipulations to the exercise program. By definition, the ACSM considers older adults at least of moderate risk, so even otherwise healthy elderly individuals should obtain medical clearance before starting an exercise program (see Table 61-9).

To limit the physiological stress of physical activity in new exercisers, older adults can divide their aerobic exercise sessions into short bouts; low impact modes (walking, cycling, and swimming) also minimize joint stress. In assigning exercise intensities, the variability in maximal heart rate in individuals greater than 65 years old suggests direct measurements of peak heart rate produce more valid and reliable prescriptions than age-predicted estimations of maximal heart rate (99).

Resistance training can considerably benefit the elderly (491). Given the role of muscular strength and endurance in activities of daily life (e.g., gripping objects), maintaining independence, and reducing the risk of falls, resistance training deserves special effort and attention for older adults. Although typical guidelines for resistance training apply to the elderly, older adults require uninterrupted supervision during the initiation of a resistance training program (99). The elderly also require exercises to maintain a functional range of motion. Improved flexibility likely decreases the risk of musculoskeletal injury and assists in the prevention of falls by improving balance and agility. Standard guidelines for improving flexibility apply to older adults (99).

REFERENCES


40. Fitts RH. Cellular mechanisms of muscle fatigue. Physiol Rev. 1974;64:49–94.


CHAPTER 61 | THERAPEUTIC EXERCISE


286. Spreuwenberg LP, Kraemer WJ, Spiering BA, et al. Influence of exer-
268. Mazzetti SA, Kraemer WJ, Volek JS, et al. The influence of direct super-
266. Lee A, Craig BW, Lucas J, et al. The effect of endurance training,
269. Kraemer WJ. Detraining the “bulked-up” athlete: prospects for lifetime
283. Jones DA, Rutherford OM. Human muscle strength training: the
278. Ebbeling CB, Clarkson PM. Exercise-induced muscle damage and adap-
272. Kraemer WJ, Fleck SJ. Optimizing Strength Training—Designing
270. Santana JC.
280. Ryschon TW, Fowler MD, Wysong RE, et al. Efficiency of human skel-
279. Ikai M, Fukunaga T. A study on training effect on strength per unit
275. Komi PV, Kaneko M, Aura O. EMG activity of leg extensor muscles
274. Knapik JJ, Mawdsley RH, Ramos MU. Angular specificity and test
276. Mazzetti SA, Kraemer WJ, Spiering BA, et al. Influence of exer-
287. Kraemer WJ, Fleck SJ. Optimizing Strength Training—Designing
277. Santana JC. Functional Training: Breaking the Bonds of Traditionalism,
281. Knapik JJ, Mawdsley RH, Ramos MU. Angular specificity and test
273. Jones DA, Rutherford OM. Human muscle strength training: the
effects of three different regimes and the nature of the resultant changes.
289. Goetghuch LA, Loebel CC, Nindl BC, et al. Hormonal responses to multi-
293. McDonagh MJN, Davies CMT. Adaptive response of mammalian skel-
294. McGee D, Jesse TC, Stone MH, et al. Leg and hip endurance adap-
295. Kraemer WJ. A series of studies: the physiological basis for strength
299. Landin D, Nelson AG. Early phase strength development: a four-
302. Häkkinen K. Factors influencing trainability of muscular strength dur-
304. Häkkinen K, Komis PV, Aalen M, et al. EMG, muscle fibre and force production characteristics during a 1 year training period in elite weight-


INTRODUCTION AND BRIEF HISTORY

Historical Applications

Throughout all recorded history, the sick and suffering have resorted to springs, baths, and pools for their soothing, healing, and powerful effects. “Taking the waters,” soaking in baths and pools, and resting at places called spas played an important social and spiritual role in the river valley civilizations of Mesopotamia, Egypt, India, and China. Ritual bathing pools were widely used for individual and social renewal and healing. Healing water rituals also appeared in ancient Greek, Hebrew, Roman, Christian, and Islamic cultures. During the Middle Ages in Europe, the beginnings of formal resorts formed for the purposes of healing began to emerge. These became the progenitors of the current group of European spas. Dr. Sidney Licht, a physiatrist and founding member of the American Society of Medical Hydrology and Climatology, defined a spa as a “place where mineral containing waters flow from the ground naturally, or to which it is pumped or conducted, and is there used for therapeutic purposes.”

In the decades before the American Civil War, major health reforms swept the nation, and numerous hydropathic establishments, institutes, and medical colleges were built for the practice of water cure cold water bathing. The American Medical Association Committee on Sanitaria and Springs published its first national report on sanitaria in 1880. At the turn of the century, medical men such as Simon Baruch, John Harvey Kellogg, and Guy Hinsdale regularly conducted clinical hydrotherapeutic experiments and prescribed aquatic therapeutic exercises that “carefully regulated dosage, character, frequency, and duration for remedying bodily deformities and restoring muscle function.” During the 1930s, the Simon Baruch Research Institute of Balneology was established, and the facility began the printing of their scientific bulletin, The Publications of Saratoga Spa. At Hot Springs, Arkansas, a warm swimming pool was installed for special underwater physical therapy exercises and pool therapy treatments for chronic arthritic patients.

By 1937, Dr. Charles Leroy Lowman published his Technique of Underwater Gymnastics: A Study in Practical Application, in which he detailed pool therapy methods of specific underwater exercises that “carefully regulated dosage, character, frequency, and duration for remedying bodily deformities and restoring muscle function.” During the 1950s, the National Foundation for Infantile Paralysis supported the corrective swimming pools and hydrogymnastics of Charles L. Lowman, and the therapeutic use of pools and tanks for the treatment of poliomyelitis. In 1962, Drs. Sidney Licht, Herman Flax, Sigmund Foster, William Erdman, Lucille Eising, J. Wayne McFarland, Jens Henriksen, and Richard Gubner organized the ASMH, which traditionally continues to meet at the annual meeting of the American Academy of Physical Medicine and Rehabilitation.
Current Trends and Uses
With the end of the polio epidemic and the rise of newer and more exciting technology in rehabilitative therapeutics, the use of the aquatic environment in rehabilitation waned. Medical hydrology became a smaller and smaller component of physiatric and physical therapy training. Principles of medical hydrology and aquatic rehabilitation that were formerly an important part of a therapist’s training program were deemphasized. Since the mid-1980s, the increase in health-care costs, constraints on physical therapy utilization under managed care, and an emerging public understanding of the value of aquatic rehabilitation and better training within the rehabilitative professions have begun to drive a broader utilization of aquatic therapies. Fortunately, basic science research in the biologic effects of immersion accelerated during the late 1960s for two fortuitous reasons. Water was recognized as a wonderful surrogate for the weightlessness of space, and as we prepared to place humans outside gravity, it became essential to forecast the effects of space flight. At the same time, Murray Epstein and other endocrinologists realized that aquatic immersion was a benign means of simulating central volume expansion to better understand volume homeostasis. The end result of these research efforts has been that aquatic therapy has a considerable foundation of high-quality basic science research.

THE PHYSICS OF WATER AND ITS RELATIONSHIP TO AQUATIC REHABILITATION

Density and Specific Gravity
Density is defined as mass per unit volume and is given the Greek letter rho (ρ). The relationship of ρ to mass and volume is characterized by the following formula:

\[ ρ = \frac{m}{V} \]

where \( m \) is the mass of a substance whose volume is \( V \). Density is measured in the international system by kg/m³ and occasionally as g/cm³. Density is a temperature-dependent variable, although much less so for solids and liquids than for gases. In addition to density, substances are defined by their specific gravity, the ratio of the density of that substance to the density of water. By definition, water has a specific gravity equal to 1.00 at 4°C. Salt water density varies considerably, from ocean to ocean, up to the Dead Sea, with a density greater than 1.16 (11). (The Black Sea has the lowest ocean density nearly approximating 1, due to significant fresh water inflow, but this has created significant environmental concerns.) Although the human body is mostly water, the body’s density is slightly less than water and averages a specific gravity of 0.974, with males averaging higher density than females. Lean body mass, which includes bone, muscle, connective tissue, and organs, has a typical density near 1.10, whereas fat mass, which includes both essential body fat plus fat in excess of essential needs, has a density of about 0.90 (12). Consequently, the human body displaces a volume of water weighing slightly more than the body, forcing the body upward by a force equal to the volume of the water displaced. The buoyancy differences between lean and fat body mass make aquatic activity particularly useful for severely obese individuals, for whom land-based exercise may produce joint pain.

Hydrostatic Pressure
Pressure is defined as force per unit area, where the force \( F \) by convention is understood to act perpendicularly to the surface area \( A \). This relationship is

\[ P = \frac{F}{A} \]

The standard international unit of pressure is called a pascal (Pa), after the French scientist Blaise Pascal, and is measured in Newtons (N)/m², dynes/cm², kg/m², and pounds per square inch (PSI). Fluids have been experimentally found to exert pressure in all directions. The pressure exerted on a theoretical point surrounded by a fluid is equal from all directions. Obviously, if unequal pressure were being exerted, the point would move until the pressures were equalized on it. Pressure is directly proportional to both the liquid density and the immersion depth, when the fluid is incompressible, as water is at the depths used in therapeutic environments. Because pressure responds not only to the fluid depth but also to any force exerted on its surface, the pressure of the Earth’s atmosphere is an important contributor to the total force from immersion. Water exerts a pressure of 22.4 mm Hg for every foot of water depth, which translates to 0.73 mm Hg/cm or slightly under 2 mm Hg/in. of H₂O depth. Thus, a body immersed to a depth of 48 in. is subjected to a distal pressure equal to 88.9 mm Hg, far greater than venous or lymphatic pressures. This external compressive force significantly aids the resolution of edema in an injured body part or where postsurgical edema resolution is a therapeutic goal.

Buoyancy
Buoyancy causes immersed objects to have less apparent weight than the same object on land. Buoyancy, a force opposite to gravity, acts on the object with a force generated by the volume of H₂O displaced. This principle was discovered by Archimedes (287 to 212 B.C.) and is the reason that water can be used to advantage in the management of medical problems, such as arthritis or severe obesity requiring weight off-loading. A human with a specific gravity of 0.97 will reach floating equilibrium when 97% of his or her volume is submerged.

Because the force of buoyancy is an upward force, there are important consequences in the therapeutic aquatic environment. The center of gravity is a point at which all force moments are in equilibrium. For an average male adult human being standing in the “anatomic” position, this point is located slightly posterior to the midsagittal plane and at about the level of the second sacral vertebra, because the human body is nonuniform with respect to density and varies depending upon body build within a radius of 1.9 to 9 cm (13). The lungs...
obviously are less dense than the lower limbs, for example. The center of buoyancy is defined as the center of all buoyancy force moments summating on each body segment. Typically, the human center of buoyancy is in the midst of the chest. The difference between the center of gravity (a downward force) and the center of buoyancy (an upward force) may generate rotational torque. This rotational torque may produce instability for some patients and should be carefully monitored, preferably by a therapist in the pool with the patient.

**Water in Motion**

**Flow Characteristics**

When water moves smoothly, in layers moving at the same speed, the water is defined as in laminar or streamline flow. When water moves more rapidly, even minor oscillations create uneven flow, and parallel paths are knocked out of alignment, resulting in turbulent flow. Within the mass of water, flow patterns arise that run dramatically out of parallel. These paths are called eddy currents. An example of the latter are the eddy currents that form in the blood stream behind artery walls encrusted with cholesterol plaque. Turbulent flow absorbs energy at a much greater rate than streamline flow, and the internal friction within the fluid determines the rate of energy absorption. This internal friction is called viscosity. The major determinants of water motion are viscosity, turbulence, and speed. Flow rates decrease when turbulence occurs, largely due to the significant nonlinear increase in internal friction in the fluid. The onset of turbulent flow obviously is a function of fluid velocity, but it is also related to fluid density, viscosity, and enclosure radius. The transition from laminar flow to turbulent flow often occurs abruptly with increasing velocity.

**Viscosity and Drag**

Viscosity refers to the magnitude of internal friction specific to the fluid. As layers of fluid molecules are set into motion, molecular attraction creates resistance to movement, and is detected as friction. Energy must be exerted to create movement, and as in the first law of thermodynamics, energy is never lost but rather is transformed and stored as potential or kinetic energy. Some energy is transformed into heat, some into kinetic energy, and some may be stored as potential energy by increasing surface tension. Fluids are in part defined by individual viscosity, expressed quantitatively as the coefficient of viscosity (Table 62-1). The greater the coefficient, the more viscous the fluid and the greater the force required to create movement within the fluid. This force is proportionate to the number of molecules of fluid set in motion and the velocity of movement. The SI unit of measurement of viscosity is called a poise, after the French scientist J. L. Poiselle (1799 to 1869), who studied the physics of blood circulation. Because velocity is described by distance/time, viscosity is the first time-dependent property. An equation that expresses this relationship must define the volume of the fluid in motion, measured as the area \( A \), depth \( l \), and velocity \( v \) of the motion.

\[
F = \eta A \frac{v}{l}
\]

**Specific Heat**

Water is used therapeutically in all three of its thermal states: solid, liquid, and gas. A major reason for its usefulness lies in the physics of aquatic thermodynamics. All substances on earth possess energy stored as heat. This energy is measured in a quantity called a calorie (cal). A calorie is defined as the heat required to raise the temperature of 1 g of water by 1°C (e.g., from 14.5°C to 15.5°C). The energy required to raise temperature is defined in kilocalories, the amount of energy required to raise 1 kg of water 1°C, and this unit by convention is termed a Calorie (with a capital C) (Cal). This is the unit in which food energy content is measured. The British system measures heat energy in British Thermal Units (BTU), the amount of energy required to raise 1 lb of water 1°F (1 BTU = 0.252 Cal). A mass of water possesses a measurable amount of stored energy in the form of heat. Energy stored may be released through a change to a lower temperature, or additional energy may raise the water temperature. The formula defining the quantity of energy required or released is

\[
Q = mc\Delta T
\]
where \( m \) equals the mass of water, \( c \) equals the specific heat capacity of the fluid, and \( \Delta T \) equals the change in temperature. The work required to produce this energy is called the mechanical equivalent of heat and is measured in joules (J). One Calorie is equivalent to \( 4.18 \times 10^3 \) J. A body immersed in a mass of water becomes a dynamic system. If the temperature of the water exceeds the temperature of the submersed body, the system equilibrates to a different level, with the submersed body warming through transference of heat energy from the water, and the water cooling through loss of heat energy to the body. By the first law of thermodynamics, the total heat (and thus energy) content of the system remains the same. Energy applied to this system increases the kinetic energy of some of the molecules, and when high kinetic energy molecules collide with lower kinetic energy molecules, they transfer some of their energy, increasing and equilibrating the total energy of the system. Water is defined as having a specific heat capacity equal to 1.00. Air, in contrast, has a far lower specific heat capacity = 0.001. Thus, water retains heat 1,000 times more than an equal volume of air.

**Thermal Energy Transfer**

The therapeutic utility of water is greatly dependent on both its ability to retain heat and its ability to transfer heat energy. Exchange of energy in the form of heat occurs in three ways: conduction, convection, and radiation. Conduction may be thought of as occurring through molecular collisions that take place over a small distance. For example, a hydrocollator pack transfers heat by conduction. Substances vary widely in their ability to conduct heat. Convection requires the mass movement of large numbers of molecules over a large distance (i.e., fluid flow). For example, a whirlpool bath transfers heat by convection. Heat transfer across a gradient is measured by the amount of heat in calories transferred per second across an imaginary membrane. Liquids and gases are generally poor conductors but good convectors. Water is an efficient conductor of heat and transfers heat 25 times faster than air. Radiation transfers heat through the transmission of electromagnetic waves. For example, a heating lamp transfers heat by radiation. The rate of radiant energy transfer from a body is proportional to the fourth power of its temperature in degrees Kelvin. It is also proportional to surface area, to the emissivity of the material, and to the second power of the distance between the energy-radiating and energy-absorbing bodies. The thermal conductive property of water, in combination with water’s high specific heat, makes its use in rehabilitation versatile, as it retains heat or cold while transferring it easily to the immersed body part.

**BIOLOGIC ASPECTS OF AQUATIC REHABILITATION**

**Circulatory System Effects**

Water begins to exert external pressure on the body immediately on immersion. Intrinsic pressure within the venous and lymphatic side of the circulation is much lower than pressure on the arterial side of the system. Venous and lymphatic pressures vary, depending on the part of the body and its vertical relationship to the heart, but are in part controlled by the system of valves within both systems, which prevents backflow. These one-way valves act to divide the large vertical column fluid into many short columns with little vertical height. This allows much lower hydrostatic pressure gradients inside the vessel wall, so that the maximum venous pressure is 30 mm Hg peripherally, decreasing steadily as blood travels toward the right atrium, which has a negative pressure of −2 to −4 mm Hg. The role of these valves in maintaining a low-pressure system is critical, as can be observed when they fail, creating venous varicosities due to the lack of sufficient vessel wall strength to support the increased height of the fluid column. Consequently, venous and lymphatic returns are sensitive to external pressure changes, including compression from surrounding muscles and from external water pressure. During water immersion, hydrostatic pressure displaces blood upward through this one-way system, first into the thighs, then into the abdominal cavity vessels, and finally into the great vessels of the chest cavity and into the heart. Central venous pressure begins to increase with immersion to the xiphoid and increases until the body is completely immersed. Right atrial distension occurs, and pressure increases by 14 to 18 mm Hg during immersion to the neck, going from about −2 to −4 mm Hg to +14 to +17 mm Hg (15,16). The transmural pressure gradient of the right atrium increases significantly, measured by Arborelius and others at 13 mm Hg, going from 2 to 15 mm Hg. Extra systoles may result, especially early into immersion (15).

Pulmonary blood flow increases with increased central blood volume and pressure. Mean pulmonary artery wedge pressure increases from 5 mm Hg on land to 22 mm Hg during immersion to the neck (16). Most of the increased pulmonary blood volume is distributed in the larger vessels of the pulmonary vascular bed and only a small percentage (5% or less) at the capillary level. This is validated by the fact that the diffusion capacity of the lungs changes very little.

Central blood volume increased by 0.7 L in Arborelius’ classic study (15). This represents a 60% increase in central volume, with one third of this volume taken up by the heart and the remainder by the great vessels of the lungs. Cardiac volume increases 27% to 30% with immersion to the neck (17), but the heart is not a static receptacle. The healthy cardiac response to increased volume (stretch) is to increase the force of contraction. As the myocardium stretches, an improved actin/myosin filament relationship is produced, enhancing the myocardial efficiency (Starling’s law) (18). Mean stroke volume increases 35% on average with immersion to the neck from a resting baseline of about 71 mL/beat to about 100 mL/beat, which is close to the exercise maximum for a sedentary, deconditioned individual on land (19,20). There is both an increase in end-diastolic volume and a decrease in end-systolic volume (15). Stroke volume is one of the major determinants of the increase in cardiac output seen with training because heart rate response ranges remain relatively fixed (19).

Most of the changes are temperature dependent, with cardiac output increasing progressively with increasing water
temperatures. Weston found cardiac output to increase by 30% at 33°C rising to 121% at 39°C (17). There is considerable individual variance in the many studies assessing this phenomenon.

As cardiac filling and stroke volume increase with deeper immersion from symphysis to xiphoid, the heart rate typically decreases (16,21). This decrease is variable, with the amount of decrease dependent on water temperature. Typically, at average pool temperatures, the rate decreases by 12% to 15% (17). There is a significant relationship between water temperature and heart rate. At 25°C, the heart rate decreases by approximately 12 to 15 beats/minute (17,22), whereas at thermoneutral temperatures, the rate of decrease is less than 15%, and in warm water, the rate generally increases significantly, contributing to the major increase in cardiac output at high temperatures. The reduction variability is believed to be related to decreased peripheral resistance at higher temperatures and increased vagal effects (23). Ongoing work in our lab has corroborated the heart rate effects, with healthy college-aged students showing a reduction of heart rate in cool (30°C) water, an increase to about baseline in neutral (36°C), and elevation averaging about 10% in warm water (39°C). (Becker BE, Hildenbrand K, Sanders JP, et al. Biophysiological Effects of Warm Water Immersion. *International Journal of Aquatic Research & Education*. 2009;3(1):24−37.) At the same time, by monitoring autonomic nervous system activity, we found elevation of sympathetic activity in cool water, substantially reducing in warm water, with vagal power increasing through the cool to warm immersion sequence. (Becker BE, Hildenbrand K, Sanders JP, et al. Age-Dependent Autonomic Changes Following Immersion in Cool, Neutral, and Warm Water Temperatures. *International Journal of Aquatic Research and Education*. 2010;4(2):127−146.)

The most efficient way for the heart to deliver more blood during exercise is to increase stroke volume. Maximal myocardial oxygen consumption efficiency (peak heart muscle efficiency) occurs when stroke volume increases because heart rate acceleration is a less efficient means of increasing output (18,24). Thus, as cardiovascular conditioning occurs, cardiac output increases are achieved through smaller increases in heart rate but greater stroke volumes per beat. This is the reason that conditioned athletes are able to maintain lower pulse rates for a given cardiac output than those who are deconditioned (18,24).

Several studies have validated the use of aquatic environments in cardiovascular rehabilitation after infarct and ischemic cardiomyopathy by actively rehabilitating heart patients in an aquatic environment (25). Tei and Tanaka found that a single brief immersion in a hot water (41°C) bath decreased both pulmonary wedge pressure and right atrial pressure by nearly 30%. Over a period of 1 month of daily therapy, patients showed nearly a 30% increase in ejection fraction, significantly improving by one and sometimes two New York Heart Association classifications (26−30). Recent work on the use of aquatic exercise in heart disease has shown significant benefit in both mild to moderate heart failure and in general cardiac rehab (31−33). Severe congestive heart failure may still be a contraindication to aquatic exercise, however (34).

Because the ultimate purpose of the heart as an organ is to pump blood, its ultimate measure of performance is the amount of blood pumped per unit time. Cardiac output is the product of stroke volume times pulse rate per unit time. Submersion to the neck increases cardiac output by more than 30% (15). Output increases by about 1,500 mL/minute, of which 50% is directed to increased muscle blood flow (15). Normal resting cardiac output is approximately 5 L/minute. Maximum output in a conditioned athlete is about 40 L/minute, which is equivalent to 205 mL/beat times 195 beats/minute. Maximum output at exercise for a sedentary individual on land is approximately 20 L/minute, equivalent to 105 mL/beat times 195 beats/minute (24). Because immersion to the neck produces a cardiac stroke volume of about 100 mL/beat, a resting pulse of 86 beats/minute produces a cardiac output of 8.6 L/minute and is already producing increased cardiac work. The increase in cardiac output appears to be somewhat age dependent, with younger subjects demonstrating greater increases (59%) than older subjects (22%) (35). The increase is also highly temperature dependent, varying directly with temperature increase from 30% at 33°C to 121% at 39°C (17). Recent research has shown that conditioned athletes demonstrate an even greater increase in cardiac output than untrained control subjects during immersed exercise and that this increase was sustained for longer periods than in the control group (36). Therefore, the myth that water exercise is not aerobically efficient is faulty; it may be an ideal cardiovascular conditioning medium. There is an emerging body of research on water exercise–induced cardiac output, but significant work needs to be done to delineate the effects of age, gender, temperature, and conditioning, as well as to explain the significant individual response variations. The cascade of cardiovascular responses to immersion is shown in Figure 62-1.

During immersion to the neck, systemic vascular resistance decreases by 30% (15). Diminished sympathetic vasomotor tone decreasing by 30% from 17 to 12 mm Hg at thermoneutral temperatures (37). Total peripheral resistance decreases during the first hour of immersion and remains low for several hours thereafter. This decrease is related to temperature, with higher temperatures producing greater reductions. This resistance drop decreases end-diastolic pressures. Immersed systolic pressures increase with increasing workload, as they do on land, but these increases appear to be reduced in magnitude when compared with equivalent land-based work (38). Venous pressures also decrease during immersion because less vascular tone is required to support the system.

The effect of immersion on blood pressure has been quite extensively studied. A consistent finding has been the striking individual variation, but trends have emerged that are useful. Very short-term immersion (10 minutes) in thermoneutral temperatures has been found to slightly increase both systolic and diastolic pressures, perhaps as part of the thermal accommodation process (38). By contrast, a pilot study in our institution found that mean arterial pressures demonstrated a 15% to 25% decrease at 5 minutes of immersion. Other studies conducted in carefully controlled environments have found no
effects or actual decreases in pressures during longer immersion periods more typical of therapeutic sessions (39). In an important study for aquatic rehabilitation, Coruzzi et al. (40) found that longer immersion produced significant decreases in mean arterial pressure, with sodium-sensitive hypertensive patients showing even greater decreases (18 to 20 mm Hg) than normotensive patients, and sodium-insensitive patients showing smaller decreases (5 to 14 mm Hg). No studies have demonstrated consistent sustained increases in systolic pressure with prolonged immersion, although several have found no significant decrease. Based on a substantial body of research, the therapeutic pool appears to be both a safe and potentially therapeutic environment for both normotensive and hypertensive patients, with both sodium-sensitive and sodium-insensitive hypertensive individuals demonstrating decreases in pressure during therapeutically customary periods of immersion (41).

In 1989, Gleim and Nicholas (38) found that oxygen consumption (VO₂) was three times greater at a given running speed (53 m/minute) in water than on land. Therefore, during water walking and running, only one-half to one-third the speed was required to achieve the same metabolic intensity as on land (38). It is important to note that the relationship of heart rate to VO₂ parallels the relationship during land-based exercise, even with accounting for the heart rate decrease in water. Consequently, metabolic intensity in water may be predicted as on land from monitoring heart rate.

**Pulmonary System Effects**

The pulmonary system is profoundly affected by immersion of the body to the level of the thorax. Part of the effect is due to the shift of blood into the chest cavity, and part is due to compression of the chest wall itself by water. The combined effect is to alter pulmonary function, increase the work of breathing, and change respiratory dynamics.

Functional residual capacity reduces to about 54% of the normal value with immersion to the xiphoid (42). Most of this loss is due to reduction in expiratory reserve volume (ERV), which decreases by 75% at this level of immersion (20). The change in this volume may be readily experienced at poolside: while sitting on the edge of the pool exhale normally, and then expel the rest of the reserve volume forcibly. Take a normal breath and then enter the water to neck level and perform the same
experiment. The difference is very perceptible. Little air remains to exhale at the end point of relaxed exhalation. ERV is reduced to 11% of vital capacity, equal to breathing at a negative pressure of −20.5 cm H₂O (42,43). There is some loss of residual volume, which decreases by 15% (42). Vital capacity decreases by about 6% to 9% when comparing neck submersion to controls submerged to the xiphoid (20,42). About 50% to 60% of this vital capacity reduction is due to increased thoracic blood volume, and 40% to 50% is due to hydrostatic forces counteracting the inspiratory musculature (20,42). Pressure on the rib cage shrinks the rib cage circumference by approximately 10% during submersion (18,20). The reduction in vital capacity does appear to fluctuate somewhat with temperature, with cooler water immersion (25°C) producing a greater reduction and warm water immersion (40°C) a smaller reduction (44). Figure 62-2 depicts the changes in pulmonary function during immersion.

The ability of the alveolar membrane to exchange gases is called diffusion capacity. Diffusion capacity of the lungs is reduced slightly, as is partial pressure of oxygen (PO₂) as the lung beds become distended with blood shifted from the extremities and abdomen. Total intrapulmonary pressure shifts to the right by 16 cm H₂O (45). This causes airway resistance to the movement of air to increase by 58% or more, resulting from reduced lung volume (20). Expiratory flow rates are reduced, increasing the time to move air into and out of the lungs. Chest wall compliance is reduced as a result of the pressure of water on the chest wall, increasing pleural pressure from a −1 to +1 mm Hg (15).

The combined effect of all these changes is to increase the total work of breathing. The total work of breathing for a tidal volume of 1 L increases by 60% during submersion to neck. Three quarters of this effort is attributable to an increase in elastic work (redistribution of blood from the thorax), and the rest is due to dynamic work (hydrostatic force on the thorax) (20,46). Thus, for an athlete accustomed to land-based conditioning exercises, a program of water-based exercise results in a significant workload challenge to the respiratory apparatus. In the authors’ experience, this challenge can raise the efficiency of the respiratory system and improve athletic performance if the time spent in water conditioning is sufficient to achieve respiratory apparatus strength gains. A study in our lab completed in 2006 and pending publication at the time of this writing found that students completing a semester-long land-based aerobics class three times per week did not show an increase in respiratory endurance, although an aquatic aerobics group of students did show statistical improvement in inspiratory endurance, even though both groups showed equivalent improvement in VO₂max, and increase in lean body mass (47). Corroboration of this experience has been found in two studies done on elite cyclists following an inspiratory muscle-strengthening program (48,49). These studies found athletic performance improvement as well as measurable strength and

<table>
<thead>
<tr>
<th>Head-out Water Immersion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Central Blood Volume</td>
</tr>
<tr>
<td>Chest Wall Pressure</td>
</tr>
<tr>
<td>Abdominal compression</td>
</tr>
<tr>
<td>Pulmonary vessel fill</td>
</tr>
<tr>
<td>Chest circumference</td>
</tr>
<tr>
<td>Diaphragm height</td>
</tr>
<tr>
<td>Diffusion capacity</td>
</tr>
<tr>
<td>Airway resistance</td>
</tr>
<tr>
<td>Lung volume &amp; VC</td>
</tr>
<tr>
<td>efficiency</td>
</tr>
<tr>
<td>↓ PO₂</td>
</tr>
<tr>
<td>Expiratory flow rate</td>
</tr>
<tr>
<td>Pulmonary compliance</td>
</tr>
<tr>
<td>Work of breathing 60%</td>
</tr>
</tbody>
</table>
endurance gains in respiratory function. A series of studies on patients with emphysema found that a treatment program involving breathing out into water for 30 minutes/day, 5 days/week for 8 weeks increased the FEV1, percentage of forced vital capacity, and PaO₂ and that clinical improvement followed these laboratory results (50–53). Inspiratory muscle strengthening has also been shown to improve both ventilatory function and perceived difficulty of breathing in tetraplegia (54). Aquatic exercise has been used successfully in both young and old asthmatic patients (33,55–57). (Becker B, Hildenbrand K, Freson T, et al. Development of an Aquatic Exercise Training Protocol for the Asthmatic Population. International Journal of Aquatic Research and Education. 2010;4(3):278–299.)

Musculoskeletal Effects
Joint Effects
As the body gradually immerses, water is displaced, creating a progressive off-loading of the immersed joints. With neck immersion, only about 15 lb of compressive force (the approximate weight of the head) is exerted on the spine, hips, and knees. A person immersed to the symphysis pubis has effectively off-loaded 40% of body weight, and when further immersed to the umbilicus, approximately 50%. Xiphoid immersion produces 60% or more off-loading, depending on whether the arms are overhead or beside the trunk (58). A body suspended or floating in water essentially counterbalances the downward effects of gravity by the upward force of buoyancy. This effect may be of great therapeutic utility. For example, a fractured pelvis may not become mechanically stable under full-body immersion for a period of many weeks, but with water immersion, gravitational forces may be partially or completely offset so that only muscle torque forces are present on the fracture site(s), allowing “active-assisted” range-of-motion activities, gentle strength building, and even gait training.

The effects of buoyancy and water resistance make possible high levels of energy expenditure with relatively little movement and strain on lower-extremity joints. Off-loading occurs as a function of immersion but the water depth chosen may be adjusted for the amount of loading desired (see Fig. 62-1). The spine is especially well protected during aquatic exercise programs, thus facilitating early rehabilitation.

Shallow-water vertical exercises generally approximate closed-chain exercise, but with reduced joint loading because of the partial buoyancy counterforce. Deep-water exercises more generally approximate an open-chain system, as do horizontal exercises such as swimming. Paddles and other resistive equipment tend to close the kinetic chain. Thus, the therapist can vary the amount of open- versus closed-chain joint loading by varying the activity and resistive equipment used.

The ground counteracts the force exerted against the floor by the walking body. This force is termed ground reaction force and may easily be measured through a force plate. It has been found to differ substantially during walking in chest-deep water. Force plate tracings of the pressure generated during a gait cycle on dry land compared with chest-deep water walking are substantially reduced in magnitude by more than 50% and are generated more slowly. Moreover, the forces are transmitted over a longer time interval during water walking (59). Clinically, this means that less joint compression is produced, and impact strain is diminished.

Water immersion causes significant effects on blood circulation through muscle tissue. These effects are caused by the compressive effects of immersion as well as by the reflex regulation of blood vessel tone. To resist blood pooling during dry conditions, sympathetic vasoconstriction tightens the resistance vessels of skeletal muscle. Immersion pressure removes the biological need for vasoconstriction, thus increasing muscle blood flow. During immersion, it is likely that most if not all of the increased cardiac output is redistributed to skin and muscle rather than to the splanchnic beds (37). Resting muscle blood flow has been found to more than double during immersion to the neck, and with this increased perfusion, muscle tissue washout was found to increase 130% above dry land clearance (60,61). Thus, oxygen delivery is significantly increased during immersion, as is the removal of muscle metabolic waste products. Hydrostatic forces add an additional circulatory drive to remove edema, muscle lactate, and other metabolic end products.

Renal and Endocrine Effects
Because aquatic immersion produces central volume expansion in a pharmacologically and physiologically noninvasive manner, aquatic immersion serves as an excellent model for volume homeostasis. Aquatic immersion creates many effects on renal blood flow, on the renal regulatory systems, and on the endocrine systems. These effects have been studied extensively in both the American and international literatures. Epstein (37), one of the most skilled and prolific researchers in studying immersion effects on human, published an exhaustive summary of these effects in 1992. The flow of blood to the kidneys increases immediately on immersion. This causes an increase in creatinine clearance, a measure of renal efficiency, initially on immersion (37). Renal sympathetic nerve activity decreases because of the vagal response caused by left atrial distension, as discussed earlier in this chapter, and this decrease in sympathetic nerve activity increases renal tubular sodium transport (62). Calculated renal vascular resistance decreases by about one third (37). Renal venous pressure increases almost twofold (37). Sodium excretion increases tenfold in individuals with normal total-body sodium, and this sodium excretion is accompanied by free water, creating a major part of the diuretic effect of immersion. This increase in sodium excretion is a time-dependent phenomenon. Sodium excretion also increases as a function of depth, as a result of the shifting of circulating central blood volume (37). Potassium excretion also increases with immersion (39). The renal effects of immersion are shown in Figure 62-3.

Renal function is largely regulated by the hormones renin, aldosterone, and antidiuretic hormone (ADH); by the dopadopamine system; and by the atrial natriuretic peptide (ANP) system. All these hormones are greatly affected by immersion (Fig. 62-4). Aldosterone controls Na⁺ reabsorption in the distal renal tubule and accounts for most of the Na⁺ loss with immersion (37).
ADH release is significantly suppressed with immersion by 50% or more and is the other major contributor to diuresis. Another factor important in sodium regulation is ANP, which has both sodium excretion-facilitating and diuretic activity. ANP relaxes vascular smooth muscle and inhibits production of aldosterone. Immersion produces a prompt and continuing increase in ANP (39). Renal prostaglandin E secretion increases steadily through the first 2 hours of immersion and then decreases gently over the next 3 hours. Plasma renin activity is reduced by 33% to 50% at 2 hours of immersion to the neck (37). Overall, immersion-induced central volume expansion causes an increase in urinary output accompanied by significant sodium and potassium excretion, beginning almost immediately on immersion and steadily increasing over several hours of immersion. This may
cause a decrease in blood pressure, and the resultant decrease may be sustained over a period of several hours. Immersion was historically one of the few effective ways of treating congestive heart failure noninvasively before the discovery of digitalis.

Accompanying the renal hormone effects are changes in the autonomic nervous system neurotransmitters, collectively called catecholamines, which act to regulate vascular resistance, cardiac rate, and force. The most important of these are epinephrine, norepinephrine, and dopamine. Catecholamine levels begin changing immediately on immersion (Fig. 62-5) (62,63).

**REHABILITATIVE APPLICATIONS**

**Arthritis and Related Disorders**

The value of the aquatic environment has a longer history in the management of arthritic diseases than in almost any other disease group. The losses that accompany chronic joint disease are many: loss of strength, loss of joint mobility and stability, and ultimately loss of functional capacity. It has been noted that rheumatoid patients as a group have lower than expected aerobic capacity and physical performance, with overall muscle strength 60% below that of age-matched control subjects. These deficits respond promptly to active rehabilitation, with well-tailored strengthening and endurance programs achieving gains in physical performance levels in as brief a time as 6 weeks (64). Long-term exercise regimens in rheumatoid patients over many years have been well tolerated, with resultant improvement in functional and other outcome measures (65). A well-researched Swedish study on the multidimensional value of dynamic therapeutic exercise in rheumatoid arthritis was published in 1993 (66). A recent study of the cost-benefit in persons with osteoarthritis demonstrated that individuals enrolled in an Arthritis Foundation aquatics course significantly reduced perceived disability related to arthritis and improved perceived quality of life specific to physical health, although once the participation, transportation, and time costs of participation were calculated, the program did not show dramatic cost-effectiveness over standard care in reducing health-care expenses (67).

Because patients with arthritis have been shown to have decreased endurance, these individuals should participate in some form of aerobic exercise to enhance their overall fitness. Studies have demonstrated the benefits of aerobic exercise for many conditions, including fibromyalgia pain (33,68–74), rheumatoid arthritis (33,65,66,75–79) lupus (80), and osteoarthritis (33,64,81–85). Low-impact exercise has been shown to be more efficacious than medications in the self-management of osteoarthritis (86). Because the safest medium to reduce impact is the pool, the incorporation of aquatic exercise and/or swimming provides an advantageous alternative to manage arthritic symptoms. This argument has been substantiated in two studies of patient groups, with nonacute rheumatoid and osteoarthritis participating in water exercise regimens. In a study of rheumatoid patients, Danneskiold-Samsoe et al. found markedly increased isometric and isokinetic muscle strength of the quadriceps after only moderate training in the pool. Other gains included an increase in aerobic capacity, freedom of movement, and a higher degree of independence in activities of daily living (76). Postural sway, a factor associated with fall risk, has been shown to decrease in patients
undergoing a 6-week aquatic exercise intervention (64). Bunning et al. concluded that pool therapy was efficacious and achieved high compliance for those with osteoarthritis and that aquatic exercise should be the cornerstone of active rehabilitation for severe arthritis. Patients who participated in the study exhibited significant improvements in aerobic capacity and walk time and physical activity level and were less depressed than controls (82). More recent studies have corroborated these results (33, 64, 65, 68, 70, 71, 76–78, 81, 83, 87).

The Arthritis Foundation, in conjunction with the YMCA, has developed a series of aquatic exercise regimens and classes for groups and individuals and certifies instructors. Training materials are available through the Arthritis Foundation (http://www.arthritis.org/conditions/exercise/water-exercise.asp). The original program was developed in 1983, was revised in 1989 and 1996, and includes range-of-motion, strengthening, and endurance-building exercises. Swimming skills are unnecessary. The certification process is sufficient so that the medical practitioner may feel comfortable referring a patient to the program. The authors’ experience has been that patient adherence has been high, therapeutic value significant, and cost to the patient low. This has been corroborated elsewhere (64). The Arthritis Foundation has developed other programs, PACE (People with Arthritis Can Exercise), a community-based group program, and Joint Efforts, a gentle exercise program for sedentary older adults, and has also developed specific programs for juvenile arthritis and deep-water exercise.

Spine Rehabilitation

Accurate diagnosis of patients’ spinal injuries and observation of their initial responses to land-based or aquatic stabilization programs help determine further therapeutic exercise treatment. A transition from dry to wet exercise conditions minimizes certain dry land risks such as axial loading of the spine, falling onto a hard surface, and limits the chance of rapid uncontrolled motion from occurring. It also establishes a supportive training environment due to reduction in axial loads and better control of spinal and peripheral joint motion, seems to improve the psychological framework in which spine-pain patients function, provides a new therapeutic activity, decreases the risk of peripheral joint injury, and allows a return to prior activity. Moving from dry to wet environments also should be considered if patients cannot tolerate axial or gravitational loads, they require increased support in the presence of a strength or proprioceptive deficit (88), or they are at risk of a compression fracture due to decreased bone density (89). Remaining in a water-supported environment is appropriate if a dry environment exacerbates symptoms or if patients prefer water. Transition from a wet to a dry environment should occur if patients are doing well in the water but must return to land to meet functional training needs efficiently and attain their ultimate competitive goals (90).

The aquatic rehabilitation programs that will be reviewed are based on dynamic lumbar, thoracic, and cervical stabilization techniques that have been previously described for land programs (91–95). Dynamic land-based stabilization training is also referred to as segmental or core stabilization exercise. This type of exercise program seeks to improve the neuromuscular control, strength, and endurance of muscles in the trunk, pelvic floor, thorax and cervical spine, theoretically helping patients gain dynamic control of segmental spine forces; eliminate repetitive injury to motion segments (i.e., discs, zygapophyseal joints, and related structures); encourage healing of injured motion segments; and possibly alter the degenerative process. The underlying premise is that motion segments and supporting soft tissues react to minimize applied stresses and reduce risk of injury (91–93). There are, however, few prospective studies on patients with low back pain and less discussion of patient selection, dose response, and long-term outcome with this type of rehabilitation construct (94, 95). The goals of aquatic stabilization exercise and swimming programs incorporate these elements but take into account the unique properties of water so that risk of spine injury is reduced. Aquatic stabilization programs help develop patients’ flexibility, strength, and body mechanics so that a smooth transition to aquatic stabilization swimming programs or other spine-stabilized activities may occur. Such programs can help first-time swimmers or patients who previously swam (96–99).

Graded elimination of gravitational forces through buoyancy allows patients to train with decreased yet variable axial loads and shear forces. In essence, water increases the safety margin of patient postural error by decreasing the compressive and shear forces on the spine. Gravitational forces may be further attenuated by using aquatic traction techniques for lumbar, thoracic, and cervical spine dysfunctions and pain (100). Velocity can be better controlled by water resistance, viscosity, buoyancy, and training devices. Buoyancy increases the range of training positions. The psychological outlook of athletes may be enhanced because rehabilitation occurs in their competitive environment. Many believe that pain attenuation takes place in the water because of the sensory overload generated by hydrostatic pressure, temperature, and turbulence (101).

Aquatic Spine Stabilization Techniques

Although the spine stabilization principles discussed for land programs also apply to aquatic programs, certain exercises that can be performed on land cannot be reproduced in water and vice versa. Aquatic programs can be designed for patients who cannot train on land or for those whose land training has reached a plateau. Eagleston (97) first described aquatic stabilization in 1989.

Eight core aquatic stabilization exercises with four levels of difficulty have been developed to provide graded training of stabilization skills (97). Programs must be customized to meet the needs of each patient’s unique spine pathology, related musculoskeletal dysfunctions, and comfort with the aquatic environment. Also, patients who have had joint replacements require particular care during positioning in the water because the replacements can change the center of buoyancy and may cause patients to sink due to high specific gravity (102). When a program is mastered, a more advanced program is provided. That said, there are numerous types of aquatic spine stabilization programs. All share at least one similar goal, to enhance dynamic stabilization. The aquatic spine specialist must be
aware of as many styles and types of aquatic stabilization programs as possible. They can then pick the most appropriate exercises from each program when creating an exercise regime that best meets their patient’s needs (94).

Eventually, if a patient wants to incorporate a swimming program, a series of transitional aquatic stabilization exercises are initiated. These help to establish a spine-stabilized swimming style that minimizes the risk of further spine injury and helps maximize swimming performance (103).

**Spine-Stabilized Swimming Programs**

Once a patient’s stabilization skills have progressed to the point when swimming is possible, a thorough analysis of stroke technique and its effect on spine motion is critical (103). The following overview focuses on lumbar spine injury and indicates the role that the cervical spine plays in the mechanics of lumbar aquatic motion.

Similar to gait analysis, analysis of stroke mechanics should be performed in an ordered, sequential manner so that all deficits and their relationships are carefully and fully scrutinized. Typically, the analysis begins at the head and progresses distally.

**Prone Swimming**

During prone swimming, the patient’s head should be midline. Breathing should occur by turning the head (i.e., rotating the head along the axial plane). There should be no craning (i.e., suboccipital cervical extension and rotation) or cervical extension and rotation (C2-C7). Body roll also contributes to proper breathing mechanics and is essential to minimize dysfunctional cervical positioning and subsequent pain. The cervical spine should be kept in the neutral position along the sagittal plane because excessive extension causes the legs and torso to drop in the water and excessive flexion can cause a struggle for air (103).

The upper-body arm position is evaluated by stroke phase. Freestyle is made up of three phases. The entry phase includes hand entry and hand submersion (ride). The pull phase incorporates insweep, outsweep, and finish components. The recovery phase includes the exit and arm swing. Several stroke defects can cause poor lumbar mechanics. If the arm abducts beyond 180 degrees, lateral lumbar flexion and rotation are produced. During the pull phase, decreased body rotation can cause lateral lumbar flexion and rotation that stress the lumbar motion segments, particularly the annular fibers surrounding the nucleus pulposus. Inadequate strength in the triceps during the finish phase results in low arm recovery, which in turn generates secondary lateral flexion and rotation through the lumbar spine. During recovery, inadequate body roll causes the neck to crane, which results in a struggle for air and accompanying lateral flexion and rotation through the lumbar spine (Table 62-2).

**TABLE 62.2 Freestyle Stroke Defects**

<table>
<thead>
<tr>
<th>Primary Peripheral Joint Stroke Defect</th>
<th>Secondary Effect</th>
<th>Spine Reaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Head high</td>
<td>Lower body sinks</td>
<td>Increased cervical extension</td>
</tr>
<tr>
<td>Head low</td>
<td>Upper body sinks</td>
<td>Increased lumbar flexion</td>
</tr>
<tr>
<td>Crane breathing</td>
<td>Lower body sinks</td>
<td>Increased lumbar flexion</td>
</tr>
<tr>
<td></td>
<td>Contralateral shoulder sinks</td>
<td>Increased cervical and suboccipital extension</td>
</tr>
<tr>
<td>Crossover hand entry</td>
<td>Lateral body movement</td>
<td>Increased lumbar lateral flexion and rotation</td>
</tr>
<tr>
<td>Wide hand entry</td>
<td>Contralateral shoulder roll</td>
<td>Increased cervical rotation</td>
</tr>
<tr>
<td>Inefficient pull power</td>
<td>Upper body sinks</td>
<td>Increased lumbar lateral flexion and rotation</td>
</tr>
<tr>
<td></td>
<td>Difficulty breathing</td>
<td>Increased cervical rotation</td>
</tr>
<tr>
<td>Increased hip flexion</td>
<td>Decreased kick propulsion</td>
<td>Increased lumbar extension</td>
</tr>
<tr>
<td></td>
<td>Lower body sinks</td>
<td>Increased lumbar lateral flexion and rotation</td>
</tr>
<tr>
<td>Crossover kick</td>
<td>Decreased kick propulsion</td>
<td>Increased cervical extension</td>
</tr>
<tr>
<td></td>
<td>Increased hip roll</td>
<td>Increased lumbar flexion</td>
</tr>
<tr>
<td>Increased knee flexion</td>
<td>Decreased kick propulsion</td>
<td>Increased lumbar extension</td>
</tr>
<tr>
<td></td>
<td>Lower body sinks</td>
<td>Increased compensatory lumbar lateral flexion and rotation</td>
</tr>
<tr>
<td>Increased ankle dorsiflexion</td>
<td>Decreased kick propulsion</td>
<td>Increased lumbar extension</td>
</tr>
<tr>
<td></td>
<td>Increased hip roll</td>
<td>Increased lumbar extension</td>
</tr>
<tr>
<td></td>
<td>Lower body sinks</td>
<td>Increased compensatory lumbar lateral flexion and rotation</td>
</tr>
</tbody>
</table>

Data from Aquatechnics Consulting Group, Inc., Aptos, CA.
The aquatic environment provides numerous advantages to assist rehabilitation of patients with spine pain. A series of aquatic stabilization exercises has been designed that incorporates the intrinsic properties of water and enhances rehabilitative efforts. When these exercises are mastered, injured patients can soon advance to spine-safe swimming or other high-level aquatic training activities (104). Swimming programs, in particular (105–107), require that close attention be directed to proper swim-stroke biomechanics and to the effect that abnormal mechanics may have on the spine. This attention ensures the most rapid rehabilitation of painful spinal disorders.

**Neurologic Disorders**

The advantages of the aquatic environment in the rehabilitation of neurologic disease have been noted for many centuries. Roosevelt’s rehabilitation at Warm Springs, Georgia, is common knowledge. The use of pool exercises for spinal injury rehabilitation is a mainstay at many rehabilitation centers, and Sir Ludwig Guttmann built a central focus for aquatics in the Spinal Injury Unit at Stoke-Mandeville. Throughout the United States, the Easter Seals organization has focused on the rehabilitation of children with cerebral palsy through pool-based programs. Programs have been designed for many neurologic diseases, and efficacy has been noted for multiple sclerosis (105–108), late poliomyelitis (109,110), and spinal cord injury (54,111,112).

From a technical standpoint, there is little difference in the aquatic techniques used in neurologic rehabilitation, with some disease-specific exceptions. Patients with multiple sclerosis may benefit greatly from aquatic programs, because of the ability of water to prevent core temperature increase during exercise, but the pool temperature should be cool to start, below thermoneutral. The authors have found ideal pool temperatures for patients with multiple sclerosis to be in the range of 25°C to 28°C (77°F to 82°F). There are literature reports of warmer temperatures being used with safety, but patient comfort should be an essential consideration (113). In contrast, spinal-injured patients lacking thermoregulatory capacity require much higher temperatures, in the range of 35°C to 37°C (95°F to 98°F). All neurologically impaired individuals obviously require close monitoring by poolside or in-water personnel. Patients with normal thermoregulatory ability and tolerance may be safely treated between 28°C and 35°C.

The techniques developed at Bad Ragaz and the methods developed by James McMillan, which have been called the Halliwick method because of their initial development and use at the Halliwick School for Crippled Girls in England, have been used extensively in the management of cerebral palsy rehabilitation. Their description is beyond the limitations of this chapter, but these methods are described thoroughly elsewhere (114,115). Essentially, the techniques use the properties of buoyancy, warmth, and careful body positioning to achieve reduction in muscle tone. The patient is floated through the use of buoyant rings to support arms, legs, and head as clinically warranted. These techniques are valuable in spinal injury rehabilitation as well and have been used in hemiplegia following stroke.

**Health Facility Pools**

Aqua facilities located within health-care facilities are generally warm-water pools, which usually cannot be adjusted for temperature, and are shallow, rarely exceeding a depth of 4.5 ft. They are usually in-ground pools. They often offer ramp access or may use slings to gain access for disabled individuals. The pools are usually small, because of constraints imposed by expensive space and construction. Although warm water is more comfortable for low-level activities and acute rehabilitation, it is not ideal for patients with multiple sclerosis, and it makes high-level activities exhausting. Pool size limits high-level activities as well, and often, group therapy must be done in small groups only. Staff is medically knowledgeable but often varies in aquatic skill. Facility-based pools are suitable for the acute rehabilitation of orthopedic populations, neurologic rehabilitation (with the exception of patients with multiple sclerosis), and arthritis rehabilitation. Common appropriate techniques include Bad Ragaz, Halliwick, aquatic spine stabilization, Red Cross arthritis, and low-level conditioning programs. Some health facilities also offer the Arthritis Foundation YMCA Aquatic Exercise Program, which has added adaptations for both deep-water and juvenile programs.

**Community Pools**

Community pools nearly always use cool water; vary in depth, often to 9 ft; usually offer stair or ladder access; and are in-ground designs. Varied depths make a wide range of programs possible, both horizontal and vertical, although the cool water precludes sedentary activities. Staff size is usually limited and without medical sophistication. These facilities are ideal for transitioning patients from the acute health facility pool, especially to group activities such as aquaerobics, low back classes, arthritis classes, or general conditioning. Because of access difficulty and cool water, they are often difficult for populations with severe physical limitations.

**Hot Tubs, Spas, and Therapeutic Tanks**

Hot tubs, spas, and therapeutic tanks usually feature hot water, are small and very shallow (rarely exceeding 3 ft in depth), and are often above-ground designs. Most offer added turbulence through jets and temperature adjustability through a narrow range. Although the high temperature is useful for acute treatments, patients quickly elevate their core temperatures even when not exercising. Hot tubs and spas are suitable for acute joint rehabilitation and relaxation; movements are limited. Staff is varied in skill level but often does not have a therapy background.

**Deep-Water Environments**

Deep-water environments, including therapeutic tanks, may be deeper and cooler and facilitate suspended activities for avoidance of weight bearing. These environments are very adaptable, permitting even high-level conditioning, including aqua-running, cross-country skiing, and ballet movements, if the temperature can be brought to an appropriate level.
Tethering may be used to stabilize the immersed individual during exercise. Swimming skill is not required, although tank use usually requires flotation and always requires immediate supervision. Deep-water facilities are ideal for athletic rehabilitation, rehabilitation of early fractures when weight bearing is precluded, and high-level conditioning activities. Staff tends to come from an athletic training background.

**Techniques in Current Use**

**Treatment Schools of Thought**

Treatment schools of thought include Bad Ragaz (116,117), Halliwick (114,115), and Watsu (118,119). All these methods of aquatic therapy are techniques consisting of a developmental and progressive set of structured exercises. The first two are primarily intended for neurologic impairment, and the latter is an offshoot of shiatsu massage technique and was not originally developed for medical therapy, although more recently it has been found useful in the management of cerebral palsy (120).

**Specialized Equipment**

**Flotation Devices**

A broad range of flotation devices have been developed for aquatic rehabilitation. For central trunk flotation, neoprene vests and foam waist belts are the most commonly used. Bad Ragaz techniques use foam rings that are placed around arms and legs or under the head. Kick boards, leg floats, vinyl foam flexible buoys, and combinations of the preceding are all important pieces of the aquatic rehabilitation armamentarium if a broad base of patients is to be treated.

**Resistive Devices**

As strengthening proceeds, the natural resistance of the water may be augmented through devices to increase the surface area of the moving part. Finned dumbbells, finned boots, kick boards, and flotation devices all may be used to add resistance to movement.

**Performance Measurement Tools**

Water is a more challenging environment for the therapist wishing to quantify performance. Waterproof heart rate monitors are useful, and as in the spine section, quantifying time, resistance, and movement freedom can add quantification (104,121). The authors have found standardized exercise log sheets to be useful. These sheets are completed during the treatment session and monitored during subsequent visits as well as in outpatient medical visits. If properly created, they may be used as a clinic progress note, providing support for reimbursement processes.

**SUMMARY**

The decline in utilization of the aquatic environment has been to the disadvantage of patients. This environment offers significant benefit, a wide margin of therapeutic safety, and cost-effectiveness to a great variety of clinical situations. In a time of scrutiny of health-care expenditures, it becomes critical to find safe, inexpensive treatment modalities for common problems. We must find methods that are suitable for self-management regimens, ideally across a large variety of clinical concerns and which may be easily learned by the patient. These methods should have the added advantages of high patient compliance and consistency. The aquatic environment offers a significant step toward these goals. Aquatic therapy is a scientifically grounded, useful approach to a broad range of rehabilitative problems from acute to chronic, and patients usually find it helpful and pleasurable. Although specific aquatherapeutic approaches are plentiful, many problems lend themselves to creative aquatic-based solutions as well. Successful rehabilitation may occur with a high safety margin and at low cost, especially when community pools are used, and professional extender personnel may be used for group programs, further decreasing cost while increasing regimen adherence (122).

**REFERENCES**


This chapter reviews the physical agents with an emphasis on their clinical use, scientific basis, and effectiveness. Although the properties of many agents overlap, discussion will begin with superficial heat and cold and then progress to hydrotherapy, the diathermies, and electrical therapies. The chapter will conclude with an analysis of some of the less established modalities, such as lasers, monochromatic light, vibration, low-intensity electrical stimulation, and extracorporeal shock wave therapy.

HEAT AND COLD

Heat and cold have potent effects on tissue. For example, temperatures above 42°C may be painful, and prolonged exposure to those above 45°C may cause injury (1). Temperatures below 13°C are also uncomfortable, and systemic temperatures below 28°C can cause death (1). In addition, metabolic and enzymatic processes are temperature dependent: an increase of 3°C increases collagenase activity severalfold (2). Heating the hands to 45°C reduces metacarpophalangeal joint stiffness 20%, whereas cooling them to 18°C increases stiffness by a similar amount (3). Temperature changes of a few degrees affect nerve conduction velocities, and changes of 5°C to 7°C alter blood flow (4–6) and collagen extensibility (7). In practice, most clinical treatments attempt to warm tissues to between 40°C and 45°C. As a point of reference, a gradual 20-minute immersion of the body in 22°C cold or 42°C warm baths result in core temperature changes of 0.3°C to 0.4°C (8). Cold therapy can produce intense local effects; ice treatment over an inflamed knee joint reduces skin temperatures by 16°C and intra-articular temperatures by 5°C or 6°C (9). Conversely, hot paraffin is reported to increase local skin temperature and knee intra-articular temperature by 7.5°C and 1.7°C, respectively (9). Although the heating modalities differ, most gain their effects by producing analgesia, hyperemia, changes in local or systemic temperatures, and reduced muscle tone. As a result, they share many of the indications (Table 63-1) and contraindications (Table 63-2) of heat in general (10). Cold’s main effects are analgesia, reduced perfusion, and muscle tone reduction. As a result, cold has indications (Table 63-3) and contraindications (Table 63-4) that are often surprisingly similar to heat.

Superficial Heat

Tissue can be heated or cooled by conduction, convection, or conversion. The first of these, conduction, requires the physical contact of two or more objects at different temperatures. A hot pack in contact with a patient’s back exemplifies conductive heating. The second, convection, also involves the transfer of energy between objects at different temperatures. In this case, however, one object, the medium, is moving relative to the other. Whirlpool baths are the most common convective heating modality. (Note that the medium in contact with the patient is in motion. As a result, the heating and cooling of the immersed body is more intense than would occur in a conductive situation with a stationary medium.) The last approach, conversion, involves the transformation of one form of energy to another. Heat lamps and ultrasound devices heat conversely, since they rely on transformation of radiant or sound energy to heat.

The physical properties of superficial heating and cooling modalities differ, but none is able to overcome the combination of skin tolerance, tissue thermal conductivity, and the body’s responses to produce temperature changes of more than a few degrees at depths of a few centimeters. The following sections review the characteristics of the most common superficial agents.

Hot Packs

Hot packs (e.g., hydrocollator packs) typically consist of segmented cloth bags filled with a silicon dioxide that, when exposed to moisture, absorbs many times its own weight of water. These packs are available in various sizes and are stored in water bath reservoirs at temperatures of 70°C to 80°C. When needed, the packs are removed from the reservoirs and excess water is allowed to drain off. The packs are then wrapped in an insulating cover or toweling and placed on the patient (Fig. 63-1).

Hot packs are one of the most commonly used heat modalities (11) due to their advantages of low cost, minimal maintenance, long life (i.e., packs last as long as 5 years; reservoirs, up to 30 years), patient acceptance, and ease of use. A variety of home-use packs are available and many can be heated in a microwave oven.

Hot packs have few risks that are not outlined in Table 63-2. However, scalding is a possibility, and it is important that excess water is drained off before use, that the insulating pad/toweling is not wet, and that the pack is placed over, not under, the patient to avoid the body’s weight expressing hot water out of the pack and wetting the insulation.

We often think of hot packs as the source of a relatively stable temperature. This is not necessarily true. Thus, although
hot packs maintain therapeutic temperatures throughout a treatment session, they cool significantly over the 20- to 30-minute period. At one time, this cooling tendency was used therapeutically. Kenny packs, for example, were thin wool cloths that were soaked in 60°C water and spun dry before use. After spinning, they contained little water and cooled rapidly. Consequently, they were replaced at 5- to 10-minute intervals and produced a cyclic heating pattern that was considered particularly effective for muscle pain and spasms of poliomyelitis.

Electric heating pads, hot water bottles, and circulating-water heating pads are alternatives to hot packs. Many of these do not cool spontaneously and not all have reliable timers or thermostats. Burns are possible and exposure may need to be limited in the elderly, people with diminished sensation, or those who may fall asleep during use.

**Heat Lamps**
Radiant heat, although less frequently used than in the past, is a versatile and easy way to warm superficial tissue. Specialized infrared (IR) sources may be used due to their convenience and durability. However, incandescent lights release most of their energy as heat, and clamp lamps using these bulbs can both be used as an inexpensive alternative.

Skin temperatures are controlled by adjusting the distance between the heat source and the patient. Point sources such as light bulbs heat according to the “inverse square law”; that is, energy density decreases inversely in proportion with the square of the distance \(1/r^2\). Elongated sources tend to follow a slower inverse distance \(1/r\) relationship. Heating is most intense when a source is perpendicular to the patient and decreases in proportion to the cosine of the angle between the beam and the perpendicular. In practice, heat sources are typically placed directly above, and about 40 to 50 cm from, the patient.

**Choice**
The physiologic effects of dry heat from a heat lamp differ little from those of moist heat of a hot pack (12). As a result, the choice depends on ease of use and preference. If the patient is in bed or cannot tolerate pressure, radiant heat may be a better option. Similarly, if a patient prefers moisture, the choice is also easy.

**Safety**
The precautions in Table 63-2 apply to the superficial heating agents (10). In addition, these modalities can burn the patient and, with chronic use, may produce a permanent brown skin discoloration (erythema ab igne).

---

### TABLE 63.1 General Indications for Therapeutic Heat

<table>
<thead>
<tr>
<th>Indications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pain</td>
</tr>
<tr>
<td>Muscle spasm</td>
</tr>
<tr>
<td>Contracture</td>
</tr>
<tr>
<td>Tension myalgia</td>
</tr>
<tr>
<td>Production of hyperemia</td>
</tr>
<tr>
<td>Acceleration of metabolic processes</td>
</tr>
<tr>
<td>Hematoma resolution</td>
</tr>
<tr>
<td>Bursitis</td>
</tr>
<tr>
<td>Tenosynovitis</td>
</tr>
<tr>
<td>Fibrosis</td>
</tr>
<tr>
<td>Fibromyalgia</td>
</tr>
<tr>
<td>Superficial thrombophlebitis</td>
</tr>
<tr>
<td>Induction of reflex vasodilation</td>
</tr>
</tbody>
</table>

### TABLE 63.2 General Contraindications and Precautions for Therapeutic Heat

<table>
<thead>
<tr>
<th>Contraindications and Precautions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acute inflammation, trauma, or hemorrhage</td>
</tr>
<tr>
<td>Bleeding disorders</td>
</tr>
<tr>
<td>Cutaneous insensitivity</td>
</tr>
<tr>
<td>Inability to communicate or respond to pain</td>
</tr>
<tr>
<td>Poor thermal regulation (e.g., from neuroleptics)</td>
</tr>
<tr>
<td>Malignancy</td>
</tr>
<tr>
<td>Edema</td>
</tr>
<tr>
<td>Ischemia</td>
</tr>
<tr>
<td>Atrophic skin</td>
</tr>
<tr>
<td>Scar tissue</td>
</tr>
<tr>
<td>Unstable angina or blood pressure</td>
</tr>
<tr>
<td>Decompensated heart failure within 6–8 weeks of a myocardial infarction (31)</td>
</tr>
</tbody>
</table>

### TABLE 63.3 General Indications for Therapeutic Cold

<table>
<thead>
<tr>
<th>Indications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acute musculoskeletal trauma</td>
</tr>
<tr>
<td>Edema</td>
</tr>
<tr>
<td>Hemorrhage</td>
</tr>
<tr>
<td>Analgesia</td>
</tr>
<tr>
<td>Pain</td>
</tr>
<tr>
<td>Spasticity</td>
</tr>
<tr>
<td>Adjunct in muscle reeducation</td>
</tr>
<tr>
<td>Reduction of local and systemic metabolic activity</td>
</tr>
</tbody>
</table>

### TABLE 63.4 General Precautions and Contraindications for Therapeutic Cold

<table>
<thead>
<tr>
<th>Precautions and Contraindications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ischemia</td>
</tr>
<tr>
<td>Cold intolerance</td>
</tr>
<tr>
<td>Raynaud’s phenomenon or disease</td>
</tr>
<tr>
<td>Severe cold pressor responses</td>
</tr>
<tr>
<td>Cold allergy</td>
</tr>
<tr>
<td>Inability to communicate or respond to pain</td>
</tr>
<tr>
<td>Poor thermal regulation</td>
</tr>
<tr>
<td>Cutaneous insensitivity</td>
</tr>
</tbody>
</table>
Effectiveness
As noted above, superficial heat and cold have clear and established effects of a variety of physiological, metabolic, and tissue effects. As such, their main use has been for the amelioration of discomfort and, to a lesser extent, the control of swelling and edema. Given their limited penetration, usage is usually restricted to the musculoskeletal system where there is support of benefits for its use as an adjunct to the treatment of a variety of musculoskeletal conditions such as low back pain (13,14), contractures (15), and rheumatoid arthritis (16).

Hydrotherapy
Hydrotherapy uses fluid to transfer thermal energy and mechanical forces to tissue. Whirlpool baths (Fig. 63-2) and Hubbard tanks (Fig. 63-3) use agitated water to produce convective heating, cooling, massage, and gentle debridement. Agitation, however, is not essential, and sitz baths, paraffin baths, and contrast baths all use a stationary medium. Immersion itself has effects. Subjects placed in water up to their chins will have transient elevations of their serum atrial natriuretic protein that are independent of temperature and result from increased venous return or right atrial loading (17).

This section discusses hydrotherapy as well as other alternatives, such as balneotherapy and fluidotherapy.

Whirlpool Baths and Hubbard Tanks
Tanks vary in size from small portable whirlpools designed to treat a single extremity to Hubbard tanks containing thousands of liters. Pools and tanks are not always essential; hand-held shower heads and small water jets are often used for local treatment such as the irrigation and debridement of deep wounds and burns.

Temperature choice depends on the amount of immersion, treatment goals, and the patient’s medical condition. Neutral temperatures of 33°C to 36°C are usually well tolerated on limited portions of the body, although for a healthy patient, temperatures of 42°C to as much as 45°C or 46°C are possible. Full-body immersion for 20 minutes can increase systemic temperatures by 0.3°C (8); Hubbard tank temperatures are usually limited to 39°C. Temperature selection should take into account the fact that for any given temperature, turbulent water heats and cools more vigorously than stationary water.

Whirlpools and Hubbard tanks are well suited for wound and burn treatment in which gentle agitation, heat, and solvent action are needed. Neutral temperatures (to somewhat warmer, depending on comfort) are chosen. Once the patient is immersed, agitation is increased to provide gentle debridement and aid in dressing removal.
“Sterile” tanks should be specified for burns and wounds. Although true sterility is not possible, a sufficient approximation is possible to make disease transmission unlikely. If wounds are large or if there is a significant exposure of internal tissue, sodium chloride may be added to the water (5 kg or more for a Hubbard tank) to improve comfort and lessen the risks of hemolysis and electrolyte imbalance. Additional agents such as potassium permanganate and gentle detergents may be added as desired.

Hydrotherapy is a common adjunct to the treatment of rheumatoid arthritis, “muscle spasm,” and joint stiffness after cast removal. Immobilized patients and those with wounds often find the concept frightening, and treatment should be reviewed before the first session with an emphasis on its comfort, transfers, and any use of plinths and hoists.

Hydrotherapy is resource intensive, expensive, and consumes large amounts of hot water and floor space. As a result, therapy clinics are tending to discard their larger units and to use the smallest baths possible.

### Contrast Baths

Contrast baths use two reservoirs, one is typically at 38°C to 40°C and the other significantly cooler at about 13°C to 16°C. The theory of their use is that the alternating exposure to heat and cold results in a reflex hyperemia and neurological desensitization. Treatment typically begins with the distal extremity placed in the warm bath for about 10 minutes and then proceeds to four cycles of alternating 1- to 4-minute cold and 4- to 6-minute warm soaks. If edema is an issue, a case can be made for ending with a cool, rather than a warm, soak.

Contrast baths may be most commonly used in treatment programs for rheumatoid arthritis and complex regional pain syndrome I (CRPS I). People with rheumatoid arthritis often benefit but may find simple warm-water soaks as effective and less difficult. Patients with CRPS I may prefer to begin with less extreme bath temperatures and also seem to benefit, but this improvement is difficult to separate from the effects of other desensitization and activity measures. There is limited evidence to suggest that contrast baths may also improve autonomic regulation and blood pressure in people with hypertension (18); research findings in other popular areas of application such as sports medicine are limited (19).

### Sitz Baths

Warm sitz baths are enshrined in the treatment of hemorrhoids, anorectal fistulas, and postpartum pain. While there has been criticism about the limitations of our knowledge of sitz bath effectiveness (20), the available research provides support for this practice. For example, sitting in water between 40°C and 50°C (with warmer temperatures perhaps more effective) lessens sphincter activity and anal pressures in normal subjects as well as those with hemorrhoids and anorectal fistulas (21). A recent trial following anal surgery found those receiving sitz baths obtained pain relief equivalent to that produced by oral analgesics (22,23). While research provides limited reports of side effects, there is, as with any heat and water modality, a risk of scalding and perineal burn at higher temperatures (24).

### Edema

The ancient Greeks and 18th-century physicians treated edema with water immersion, a practice mimicked today with compressive sleeves. Research supports this intuitively reasonable treatment; immersion increases natriuretic proteins (17) and renal water and salt loss in normal subjects as well as those with the nephrotic syndrome and cirrhosis (25). Because warmth produces a reactive vasodilation, neutral temperatures appear to be the most effective.

### Water-Based Exercise

Water-based exercise and therapeutic pools are a popular way to provide gentle, progressive exercise in a setting that permits limited weight bearing. The approach has shown benefits (26), but its value over land-based programs has been more difficult to establish. Although patients with rheumatoid and osteoarthritis often like the warmth and support of water-based exercise, benefits may not be as clear as they might seem. For example, a number of studies involving land- and water-based programs have found that water-based programs provide little or no benefits over their land-based counterparts for patients with osteoarthritis, rheumatoid arthritis, juvenile inflammatory arthritis, and anterior cruciate rehabilitation (27–29). Although many believe exercise is beneficial for patients with fibromyalgia, the benefits of water-over land-based approaches remain unclear (30–34).

### Balneotherapy

Balneotherapy (spa) therapy is now little more than a curiosity in North America. In Europe, however, acceptance remains strong and visits to spas are supported by some governments and insurance programs (35). Treatments involve a combination of physical therapy, water baths, mud treatments, mineral water consumption, and education that may take place in a resortlike atmosphere for periods as long as 3 weeks. The fundamental tenet of the approach is the belief that water containing dissolved gases (such as nitrogen and carbon dioxide), elements (e.g., calcium, magnesium, zinc, and cobalt), and compounds (e.g., hydrogen sulfide) has therapeutic effects.

Balneotherapy research has traditionally been hampered by poor quality and an environment in which a patient is undergoing multiple changes in his or her lifestyle and exercise patterns. Research quality is improving, however, and higher quality controlled studies are becoming available. Furthermore, although intact skin is relatively impervious, we know that blood concentrations of substances such as bromine and rubidium may be increased after bathing in patients with impaired skin barriers (36,37).

Although most U.S. physicians are skeptical, some controlled investigations support the use of balneotherapy for the inflammatory and degenerative arthritides. For example, although the differences are often small, patients with rheumatoid arthritis, ankylosing spondylitis, and osteoarthritis appear
to respond similarly to or somewhat better than they do to alternative treatments (35,36,38–46). There are many studies, but there is limited evidence that patients with nonradiating low back pain may do better following balneotherapy than a control treatment (47). In all, balneotherapy studies are often limited by poor size, incomplete blinding, and short follow-up. Nevertheless, the findings are intriguing and warrant honest consideration.

**Fluidotherapy**

Hydrotherapy usually uses water as the heat-exchanging medium, but substances such as pulverized corncobs and small beads “fluidized” by hot air jets may be substituted. Although these devices have been available since the 1970s, the benefits of this high-temperature, low-heat capacity approach over other approaches for improving joint mobility and sensory desensitization remain controversial (48).

**Safety**

The general precautions of heat (Table 63-2) apply to hydrotherapy. Drowning, cardiac disease, systemic hyperthermia, and disease transmission are also concerns, but they may be overemphasized. Hot water–associated seizures are rare but are known to occur (49).

Cardiac disease may not be an absolute contraindication to hyperthermia as was once thought. For example, even though sauna baths (80°C to 100°C) elevate body temperatures by 1°C or 2°C (50,51), research indicates that they may improve the function and quality of life of individuals with severe congestive heart failure (CHF) (52). Furthermore, Finnish heart attack survivors return to sauna bathing without apparent increased risk, and 15-minute soaks in 40°C hot tubs do not create ischemic electrocardiogram changes or alter systolic and diastolic blood pressures more than cardiac rehabilitation–associated stationary bicycle exercises (50,51,53).

Hydrotherapy-associated infections seem to be rare. However, there may be some reproductive consequences: neural tube defects may be increased (relative risk, 2.6 to 2.9) in the children of women who take sauna baths during early pregnancy (54), and sperm counts may be lowered after isolated or repeated sauna sessions (55).

**Paraffin Baths**

Paraffin baths are thermostatically controlled reservoirs filled with a 1:7 mixture of mineral oil and paraffin. Bath temperatures (45°C to 54°C) are higher than in most hydrotherapy but are well tolerated because of the low heat capacity of the mixture and the tendency for an insulating layer of wax to build up on the surface of the treated area.

Two paraffin treatment approaches predominate. Dipping is the most common and consists of the patient submerging the treated extremity in the bath ten times, with pauses between dips to permit a layer of paraffin to solidify. The treated area is then covered with a plastic sheet and placed in an insulating cover for about 20 minutes (Figs. 63-4 and 63-5). The paraffin is then stripped off and returned to the container.
Continuous immersion is an alternative approach in which the treated extremity is dipped six to ten times in the paraffin and then kept immersed for 20 to 30 minutes. Heating is more intense with this method but is still well tolerated because a layer of insulating solidified paraffin forms on the skin. Immersion produces the same initial maximum skin temperature as dipping; however, temperatures decrease less rapidly. At the end of a session, skin temperatures are about 41.5°C with subcutaneous and superficial intramuscular temperature increases (about 5°C and 3°C, respectively), which are higher and more persistent than what occurs with dipping (56). The dip-and-wrap method in conjunction with limb elevation is preferable if the patient is predisposed to edema.

Paraffin baths are often used to treat hand contractures associated with rheumatoid arthritis, scleroderma, burns, and injury. Dipping or immersion is the most common technique, but, at times, paraffin may be brushed on difficult-to-treat areas. Reports in the literature are scant, but these messy treatments seem frequently to be helpful and capable of producing significant temperature changes and improvements in joint mobility (9,48,57).

**Safety**

A thermometer should be kept in the reservoir, and paraffin temperature should be checked (typical values are about 48°C) before use to ensure that the paraffin is at the correct temperature and to avoid burns. (A film of solidified wax around the margins of a reservoir is a sign that the temperature is not dangerously elevated.) Small commercial units are available for home use. If the feet have poor circulation, a few insulating layers of paraffin may be brushed on before they are dipped or immersed.

Most clinicians do not heat acutely inflamed joints and tissues vigorously because even a few degrees of temperature elevation (as may occur with paraffin baths or hot soaks) increases intra-articular enzymatic activity (2). Although there has been controversy about its appropriateness, warmth clearly improves comfort, and most clinicians use superficial heat in the subacute situation.

**Diathermy**

There are three diathermy (literally “through heating”) agents: ultrasound, shortwave diathermy, and microwave diathermy. Ultrasound is the most commonly used of the three, but shortwave diathermy, despite a gradual decline, is still in use as is pulsed shortwave (which is claimed to be athermic). Microwave diathermy no longer has a place in routine therapy practice but continues to have specialized medical applications. All three will be reviewed but with a declining emphasis commensurate with their current use.

**Ultrasound**

Ultrasound is sound that occurs at frequencies above the 17,000- to 20,000-Hz limit of human hearing. As such, it shares the characteristics of sound in general: its waves consist of alternating compressions and rarefactions; it requires a medium for transmission; it transmits energy; and it can be focused, refracted, and reflected. Although arguments are made for a variety of frequencies, most therapeutic ultrasound treatment occurs between 0.8 and 3 MHz due to the practical considerations of focusing, penetration, and standardization.

**Biophysics**

Ultrasound has both thermal and nonthermal effects. Heat production, with its goals of hyperemia, enhanced soft-tissue extensibility, and lessen pain and muscle tone, is the best known. Nonthermal processes—which include cavitation, streaming, standing waves, mechanical deformation, and shock waves—may also be sought because of their ability to alter cell membrane permeability and function (58). The first of these nonthermal processes, cavitation, occurs when high-intensity ultrasound passes through a liquid and produces small bubbles that may either rhythmically oscillate in size (stable cavitation) or grow and abruptly collapse (unstable cavitation). In either case, large temperature and pressure changes may occur (59) and produce localized tissue distortion and injury. Pressure asymmetries produced by the presence of an ultrasonic beam can generate large shear forces that in turn may lead to media movement (streaming, microstreaming), tissue damage, or accelerated metabolic processes (58,60). Stopping waves are generated from the superposition of sound waves and produce fixed regions of high and low pressure at half-wavelength intervals (which for 1 MHz ultrasound [tissue velocity 1,500 m/s] is about 0.75 mm) (61). Graphic effects are possible; ultrasound exposure produces repetitive bands of red blood cells in chick embryo vessels (62).

Ultrasound penetration into tissue depends on a number of factors. Frequency is particularly important as penetration decreases by a factor of 6 as the frequency increases from 0.3 to 3.3 MHz (63). Orientation is also critical. For example, about 50% of a 0.87-MHz ultrasound beam penetrates 7 cm in a direction parallel to muscle fibers, but the same beam penetrates only 2 cm in a transverse direction (63). Tissue type is also significant. Fifty percent of an ultrasound beam penetrates several centimeters in muscle, only a few tenths of a millimeter in bone, and 7 to 8 cm in fat (63,64). In practice, 3-MHz ultrasound is used for superficial tissues such as those of the hand and temporomandibular joint, as most of its energy is absorbed within 1 to 2 cm of the skin’s surface. Lower frequency beams are used when deeper penetration is desired.

Ultrasound treatments are frequently delivered to anisotropic tissue, and it should be remembered that localized areas of temperature elevations of 5°C or more may occur at sound absorption discontinuities such as those that occur at bone–soft-tissue interfaces (63–65). Heat is lost from tissue as the result of conduction and cooling effects of the local blood flow.

**Equipment**

Ultrasound machines typically use ferromagnetic lead zirconate titanate (PZT) ceramics to convert electrical energy into sound. Machines are increasingly computerized, and in addition to
indicating treatment time, wavelength and waveform are now often provided with predesigned treatment programs. Additional capabilities such as waveform modulation and concurrent electrical stimulation are also common.

Ultrasound frequencies are relatively stable and usually remain within 5% of the manufacturer’s specifications (66,67). Machines, however, should be routinely calibrated as output powers and intensities (power output divided by the active area of the applicator) may vary by 20% or more during a session and as a unit ages (66,67).

**Technique**

There are two philosophies of ultrasound therapy. The most widely held is that ultrasound’s benefits are due to heating. This approach typically uses an unmodulated, continuous-wave (CW), or high-intensity pulsed beam with intensities of 0.5 to 2.5 W/cm². The second approach emphasizes ultrasound’s nonthermal properties. In this case, the beam is modulated to deliver brief pulses of high-intensity ultrasound separated by longer pauses of no power and as a result lower average energy intensity. Thus, heating is minimized and ultrasound’s nonthermal effects are emphasized.

Ultrasound is usually delivered by moving the applicator (Fig. 63-6) over the treated area in slow (1 to 2 cm/s), overlapping strokes. Treatments cover areas of about 100 cm² and last 5 to 10 minutes. Indirect ultrasound, while less common, is used to treat irregular surfaces such as the foot and ankle where it is difficult to keep the applicator in contact with the skin. In these situations, the body part is placed in a container filled with degassed water. The applicator is held a short distance away (0.5 to 3.0 cm) and moved without touching the skin. Power intensities may need to be higher due to transmission losses.

Coupling between the applicator and the skin is not a trivial issue. The treatment area should be cleansed before treatment, and a coupling agent is necessary. Degassed water (water that has been allowed to sit for several hours) is used for indirect ultrasound because the dissolved gases in water fresh from the tap form bubbles during treatment and attenuate the beam. Little practical difference exists between commercial gels and mineral oil for direct applications, with transmissivity similar to that for degassed water (68). However, mineral oil becomes watery and most people use the commercial products because of the convenience. Coupling agents should not be salt based (e.g., those used for EMG or ECG), as the salt may damage the applicator.

Phonophoresis is a variant of ultrasound in which biologically active substances are combined with the coupling medium in the hope that the ultrasound will force the active material into tissue. Although this technique has been in use since the 1960s, neither its effectiveness, penetration, optimal frequency, appropriate coupling mediums/active materials nor amount of material lost to the subcutaneous circulation is well established. Although claims of increased cortisol concentrations at depths of several centimeters after corticosteroid phonophoresis have been made, our own research as well as that of others (69,70) finds limited evidence for deep penetration. Clinical reports are mixed. Thus, although some clinical studies report phonophoresis with a variety of agents successful in terms of improved shoulder range of motion and pain following the treatment (71) as well as in the treatment of keloids and sarcoid nodules (72), other work involving an assortment of musculoskeletal conditions may find the approach no more effective than ultrasound alone (73).

**Indications**

**Musculoskeletal Conditions.** The research relating to musculoskeletal condition is surprisingly inconclusive (74). For example, a study of 63 patients with calcific shoulder tendonitis treated with pulsed ultrasound (0.89 MHz, 2.5 W/cm²) found that the treated patients had significantly larger improvements in their pain and decrease in calcium deposits relative to their sham controls at the end of an intensive 6-week treatment program. However, this difference had disappeared at follow-up 9 months later (75). Although ultrasound may be more beneficial than corticosteroid injection in the treatment of shoulder pain, other studies and reviews find it no more effective than placebo or nonsteroidal anti-inflammatory medications for a variety of conditions ranging from subacromial bursitis or lateral epicondylitis (76) to heel pain (77). More tellingly, perhaps, an evidence-based practice guideline panel came to the conclusion that although therapeutic ultrasound was effective in the treatment of calcific tendonitis of the shoulder, there was no compelling evidence of its clinical benefits for other sources of musculoskeletal pain (78). Although these studies raise legitimate issues about its effectiveness in such conditions, many physicians and therapists remain convinced that ultrasound is useful for the treatment of at least some musculoskeletal pain. In support of this, several reviews have found evidence of some benefits in the form of increased motion and lessened stiffness from the cautious use of ultrasound in

![FIGURE 63-6. Direct-contact ultrasound treatment of the elbow. Note the use of a folded towel to support the patient comfortably. Although not shown in the picture, a coupling agent is needed to acoustically couple the applicator and the skin.](image-url)
patients with rheumatoid arthritis (79–81). A recent small scale study of ultrasound treatment of myofascial trigger points found significant pain relief from conventional relative to low-intensity ultrasound stimulation (82).

**Contractures.** Ultrasound is effective in increasing the range of motion of the heel cords, parietal shoulders, and contracted hips (83,84). In fact, due to its penetration and focusing capabilities, it is the only agent that can significantly heat (by 8°C to 10°C) the hip joint (65). Hand and Dupuytren contractures also may benefit from ultrasound (85), although a small study of burns did not find treatment beneficial (86). Collagen and tendon extensibility increase as temperatures increase and decreases as tissue cools. As a result, stretching should begin during heating and should continue as the tissue cools and “sets.”

**Soft-Tissue Wounds and Inflammation.** Ultrasound treatment of wounds and inflammation is based on the belief that either its heat (by increasing blood flow, metabolic, or enzymatic activity) or nonthermal effects (perhaps by changing cell wall permeability) accelerate healing. Laboratory studies offer some support (87). Human studies, however, provide a more mixed picture with a recent review of decubitus ulcers concluding that the evidence was too weak to assess (88). Some feel that inflammation and swelling are indications for ultrasound. Others believe that the heating and membrane permeability changes associated with treatment prohibit its use in acutely inflamed conditions such as in recent trauma and active rheumatoid arthritis.

**Trauma.** Although ultrasound may aggravate tissue damage and swelling if used too soon after an injury, subacute hematomas (89) and postpartum perineal pain (90) may improve more rapidly with treatment. Ankle sprains are a common indication for ultrasound treatment. Even here, benefits are unclear with a systematic review involving more than 570 patients concluding that current evidence, at best, supported the presence of limited benefit (74). A number of clinical and animal studies find evidence that ultrasound at intensities on the order of 1 W/cm² may provide at least short-term benefits in the treatment of symptomatic carpal tunnel syndrome (91,92). Although these reports are intriguing, benefits are not always evident and remain controversial (93).

**Fractures.** Low-intensity ultrasound improves the repair of bony injuries (94). For example, 30 mW/cm² pulsed 1.5-MHz ultrasound accelerates the healing of closed and open grade 1 fractures (95,96). Although ultrasound is not typically used for this purpose, its use has been approved by the Food and Drug Administration (FDA) for the treatment of some fractures. In view of the fact that 5% to 10% of fractures heal slowly (97), there may be a wider applicability for this treatment in the future.

**Other Indications.** Postherpetic neuralgia is often resistant to conventional treatment. Treatment with both pulsed and continuous 1- to 1.5-MHz ultrasound has been evaluated. Unfortunately, the studies have again often been poorly controlled, and the results are unclear; some investigators have found improvement (98,99), whereas others have not (100). As ultrasound has thermal, and possibly nonthermal, effects on nerve conduction (101), further evaluation seems appropriate. Plantar warts, keloids, scars, and even chronic sinusitis are additional refractory conditions for which pulsed and conventional ultrasound has inconclusive benefits (102–104).

**Precautions and Contraindications**

Ultrasound produces intense heating and, under some conditions, potentially destructive nonthermal effects, such as shockwaves, cavitation, and media motion. The precautions in Table 63-2 should be heeded. In addition, fluid-filled cavities such as the eyes and gravid uterus are avoided due to the risks of cavitation and heat damage. The heart, brain, cervical ganglia, tumors, acute hemorrhage sites, ischemic areas, pacemakers, and infection sites should not be treated for obvious heat, neurophysiologic, and mechanical reasons. The spine should not be exposed to high intensities, and laminectomy sites, in particular, should be avoided. Ultrasound is used in children, although many are concerned about the effects of excessive energy and warmth on immature growth plates. As noted above, ultrasound is not used over acutely inflamed joints or tissues due to the fact that temperature elevations can increase enzymatic activity (105,106).

A common concern is that ultrasound treatment over metal implants in muscle or next to bone will elevate tissue temperatures to a higher level than would occur in their absence (107,108). The concept is certainly valid and although some studies found the effect smaller than might be expected, only a limited number of objects and geometries were studied. Although clear indications and contraindications do not exist, many simply avoid using ultrasound near metal.

**Shortwave Diathermy**

Shortwave diathermy (SWD) heats tissue with a combination of induced electrical currents and the vibration it imposes on the molecules of a tissue. Since SWD produces radio waves that can cause electrical interference, the U.S. Federal Communications Commission (FCC) has restricted its industrial, scientific, and medical use to 27.12, 13.56, and 40.68 MHz (wavelengths of 11, 22, and 7 m), respectively. Most SWD machines operate at 27.12 MHz.

**Biophysics**

Radio waves are absorbed as they pass through tissue, but actual penetration depths depend on the specifics of tissue as well as the frequency and characteristics of the applicator. Inductive applicators induce electrical (“eddy”) currents in tissue and, as a rule, produce the highest temperatures in water-rich, highly conductive tissues such as muscle. Capacitively coupled applicators, on the other hand, emphasize electric field heating. Maximum temperatures tend to occur in water-poor substances, such as fat, ligaments, tendons, and joint capsules (109). Significant heating is possible. Although effects depend on the dosage and application approach, SWD can increase subcutaneous fat temperatures by 15°C and
3- to 5-cm deep intramuscular temperatures by 4°C to 6°C (110,111).

SWD machines may produce pulsed as well as continuous wave (CW) output. CW diathermy is used when the treatment goal is heating. In contrast, pulsed shortwave diathermy, sometimes termed pulsed electromagnetic energy (PEMF), alternates brief periods of high power with longer periods of no power output. The average output power may be the same in both approaches, but, as is true with ultrasound, the pulsed approach may be chosen to emphasize nonthermal effects. Some investigators, for example, believe that certain pulse frequencies, even at low intensities, can increase blood flow and have resonant effects on cell function. Although nonthermal SWD phenomena (e.g., pearl chains) have been known for more than 40 years, their clinical benefits remain elusive.

**Technique**

A SWD machine is essentially a radio transmitter that is tuned in the same way that any transmitter is tuned. The patient is in the machine's field and is protected from injury by tuning the circuit (automatically in modern machines) for maximum coupling. Once coupling is maximized, movement can only reduce heating. Two treatment approaches are possible. In one, the shortwave diathermy machine transmits radio waves that induce electrical currents in tissues in the same manner as currents would be induced in a radio antenna (Fig. 63-7A). These eddy currents, in turn, are degraded into heat as they flow in the body by the tissue's electrical resistance. In the second, the tissue to be heated is placed between two plates that are connected to the SWD machine's output and serves as the dielectric of a capacitor (Fig. 63-7B). In this case, heating currents are produced as the radio wave polarity oscillates across the plates.

There are a variety of inductive applicators (see Fig. 63-7B). Drum applicators consist of coils encased in rigid containers that may be positioned with hinges around the body. Pad applicators are commercially available, semiflexible mats containing a coil that is connected to a shortwave diathermy machine and placed against the patient. Pads may have dimensions of 0.5 × 0.75 m and typically are used to treat broad areas such as the low back. Cable applicators consist of rubber-coated cables that were wrapped around an extremity or laid over the body. Cables were once common, but because of safety concerns and the need for careful placement, they have been replaced by drums and pads.

In the most common capacitive arrangement, the patient is placed between two platelike electrodes (see Fig. 63-7A). Rectal and vaginal applicators were once used in the treatment of pelvic inflammatory disease and chronic prostatitis, but due to improvements in antibiotics, are now seldom used or have been replaced by externally applied SWD or ultrasound. Recent research has shown variable benefits from SWD while no additional benefit for back pain over exercise and advice was found in large randomized controlled trials (RCTs), another investigator reported that treatment produced significant reductions in pain and synovial thickness in subjects with chronic knee synovitis (112,113).

Ultrasound can produce deep heating due to its easy use, and high availability is often substituted for SWD. This

---

**FIGURE 63-7.**

A: Shortwave diathermy treatment using a capacitive plate arrangement. Careful positioning is necessary. The patient should wear no jewelry and lie on a nonconductive table. B: Shortwave diathermy treatment using an inductive applicator. Note the capacitive plate to the side and the emergency cutoff switch in the patient’s hand. Again, metal is avoided, and the patient is carefully positioned on a nonconducting surface.
substitution may at times be inappropriate, as shortwave diathermy is less labor intensive and can warm larger areas than ultrasound.

**Microwave Diathermy**
Electromagnetic radiation can be well focused when the wavelength of the radiation, as is the case with microwave diathermy, is comparable with the dimensions of the antenna. The FCC-approved industrial, scientific, and medical frequencies for microwave diathermy are 915 and 2,456 MHz (33- and 12-cm wavelengths, respectively).

**Biophysics**
Microwaves do not penetrate tissue as deeply as shortwaves and ultrasound, but, like them, penetration decreases as frequency increases. Similarly, focusing becomes more difficult at longer wave lengths (110).

Microwaves are absorbed by water and should theoretically preferentially heat muscle. However, fat usually overlies muscle and absorbs a significant portion of the beam. As an example, at 915 MHz, subcutaneous fat temperatures may increase by 10°C to 12°C, whereas 3- to 4-cm-deep muscles will be heated only 3°C to 4°C (114).

**Technique**
Microwave diathermy was often used to heat superficial muscles and joints such as the shoulder and speed hematoma resolution (115). Today, its use is rare although there are interesting reports of its use as an adjunct to potentiate the effects of cancer chemotherapy and radiation treatment (116).

**Precautions and Contraindications**
The general contraindications listed in Table 63-2 apply to short- and microwave diathermy. Additional restrictions stem from their electromagnetic nature. For example, perspiration is conductive and, if present in a field, heats the skin. Metal produces localized heating in a field; patients should not wear jewelry, and treatment is given on a nonconductive table. Pacemakers, stimulators, surgical implants, contact lenses, and the menstruating or pregnant uterus should not be exposed to shortwave diathermy or microwave diathermy. Although many of these concerns seem almost academic, this is not always the case. For example, an inadvertent application of pulsed diathermy to the jaw led to a persistent vegetative state in a man with a deep brain stimulator (117). Some believe that small, metallic surgical clips or metallic intrauterine devices do not produce significant localized temperature elevations. This may be correct, but at least one study found that 1-cm wires positioned to mimic a surgical site produced temperature elevations during diathermy that were 3°C to 4°C greater than those that occurred in their absence (116). Many investigators follow a rule of “no metal” when using any diathermy. Microwaves selectively heat water, and its effects on the fetus are unknown. Treatment of edematous tissue, moist skin, the eyes, fluid-filled cavities, and blisters can produce unacceptable temperature elevations and cataracts.

---

**TABLE 63.5 Therapeutic and Environmental Magnetic Fields (T)**

<table>
<thead>
<tr>
<th>Field Description</th>
<th>T (Tesla)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Typical residential exposure (199)</td>
<td>0.4</td>
</tr>
<tr>
<td>Earth’s magnetic field</td>
<td>25–65</td>
</tr>
<tr>
<td>Automobile interior (221)</td>
<td>2</td>
</tr>
<tr>
<td>Subway passenger compartment (222)</td>
<td>16–64</td>
</tr>
<tr>
<td>Standard therapeutic magnet</td>
<td>250–5,000</td>
</tr>
<tr>
<td>Credit card erasure (223)</td>
<td>5,000</td>
</tr>
<tr>
<td>Pulsed electromagnetic field devices (224)</td>
<td>2,000–8,000</td>
</tr>
<tr>
<td>Bone growth stimulator</td>
<td>5,000</td>
</tr>
<tr>
<td>Exposure limits (continuous) (225)</td>
<td></td>
</tr>
<tr>
<td>Occupational</td>
<td>2 × 10^5</td>
</tr>
<tr>
<td>General public</td>
<td>4 × 10^4</td>
</tr>
<tr>
<td>MRI</td>
<td>1.5+ × 10^6</td>
</tr>
</tbody>
</table>

*1 Tesla = 10 Gauss.

Workplace safety is always an issue with electromagnetic devices. The situation, however, has been well addressed by the American National Standards Institute (ANSI) guidelines for workplace-associated exposure (118). As fields fall rapidly with distance and therapists work only intermittently with these agents, their exposures seem to fall well within these guidelines.

Epidemiologic studies seem to support the workplace safety of these electromagnetic diathermies. While it has been reported that miscarriage rates are increased or birth weights reduced in female therapists exposed to diathermy (119,120), there is no evidence of detrimental effects on the children of therapists working with these agents (119,121,122). Tables 63-5 and 63-6 display typical electromagnetic exposures and contraindications for use.

The use of any diathermy on an acutely inflamed joint is controversial. Thus, while patients have reported that treatment

---

**TABLE 63.6 General Contraindications and Precautions for Electromagnetic Agents**

<table>
<thead>
<tr>
<th>Contraindication</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jewelry</td>
</tr>
<tr>
<td>Moist tissue</td>
</tr>
<tr>
<td>Ischemia</td>
</tr>
<tr>
<td>Acute hemorrhage/injury/inflammation</td>
</tr>
<tr>
<td>Neoplasm</td>
</tr>
<tr>
<td>Implanted objects</td>
</tr>
<tr>
<td>Metal</td>
</tr>
<tr>
<td>Pacemakers</td>
</tr>
<tr>
<td>Stimulators</td>
</tr>
<tr>
<td>Pumps</td>
</tr>
<tr>
<td>Small clips/IUDs</td>
</tr>
<tr>
<td>Eyes</td>
</tr>
<tr>
<td>Gonads</td>
</tr>
<tr>
<td>Growth plates</td>
</tr>
<tr>
<td>Infection (e.g., joint)</td>
</tr>
</tbody>
</table>
improves comfort and some investigators feel that synovial

temperature elevations to 42°C or more denature enzymes and
produce a "thermal synovectomy" (123), most clinicians remain
skep
tical and avoid diathermy in this situation. As noted ear-
erlier, a similar, but less marked, controversy also surrounds
the use of superficial heating for these situations.

**Cold (Cryotherapy)**

Chilling a limited portion of the body produces local and
distant physiologic effects. If the cooling agent is ice, skin
temperature initially decreases rapidly and then more
approaches an equilibrium of about 12°C to 13°C in 10 min-
utes. Subcutaneous temperatures fall more slowly and decrease
by 3°C to 5°C at 10 minutes. Deeper, intramuscular tempera-
tures decrease the least and after 10 minutes have dropped a
degree or less (124). Chilling for longer periods generates more
pronounced cooling with intramuscular forearm temperature
decreases of 6°C to 16°C after 20 minutes to 3 hours of cool-
ing (6,9) and intra-articular knee temperatures decreasing 5°C
to 6°C over a 3-hour period of cooling with ice chips (9). It
should be remembered that although superficial tissues cool
more rapidly than their deeper counterparts, specifics vary,
depending on the nature of the cooling agent and the body's
physiologic responses.

Cooling produces an initial period of vasoconstriction
that may (125) or may not be followed by a reactive vaso-
dilation (125–127). Vasoconstriction becomes evident within
5 minutes of cooling, and after 25 minutes of packing a knee
in ice, soft tissue and bone blood flows are decreased by 30%
and 20%, respectively (9). Superficial cold also decreases met-
able activity and with prolonged application lessens muscle
tone, inhibits spasticity (128), increases gastrointestinal motil-
ity (129), slows nerve conduction, and produces analgesia
(130).

**Technique**

Ice packs, ice massage, iced compression wraps, and slushes
have high heat capacities and cool treated areas as, or more
rapidly than most alternatives such as gel packs or frozen peas
(see below) (131). Treatments tend to last 10 to 20 minutes.
Techniques are straightforward, but if ice packs are used, a
slightly damp, thin towel is placed between the pack and the
skin.

Iced whirlpools cool vigorously and usually are used for
10- to 20-minute periods. Although an athlete may be moti-
vated to tolerate them, the average patient finds temperatures
below 15°C uncomfortable. If the feet or hands are exposed
to a cold bath, neoprene booties, hospital slippers, or woolen
socks or gloves may increase tolerance.

Ice massage consists of rubbing a piece of ice (e.g., an
ice cube or water frozen in a small cup) over the painful area.
Analgesia can be achieved in 7 to 10 minutes, with most patients
reporting successive sensations of cold, burning, aching, and
numbness. A moisture barrier should be present between ice
and skin, and ice should slide smoothly over the treated area.
Although chemical and refrigerated agents may have tempera-
tures below 0°C and can produce frostbite, ice treatments of
healthy people for periods of less than 30 minutes do not seem
to cause injury.

Vapocoolant and liquid nitrogen sprays can abruptly
reduce skin temperature by 20°C (9) and are used for local skin
analgesia and the “spray-and-stretch” techniques. Prepackaged
chemical ice packs may consist of two compartments (e.g., one
filled with water, the other with ammonium nitrate), which,
when broken and the contents mixed, produce a cooling,
endothermic reaction. Although these packs are convenient
and pliable, they are small, expensive, and tend to cool poorly
(131). Alternatives such as refrigerated and pressurized water
pressure cuffs (Fig. 63-8) are also available. Twelve-ounce fro-
zen orange juice containers and plastic bags of frozen peas,
which conform to the body when struck on a hard surface, are
effective alternatives for home use.

**Indications**

**Trauma**

Many studies have found that chilling limits hypoxic damage,
lessens edema, speeds recovery, and reduces compartmen-
tal pressures after injury (132). Although there are concerns
that cooling does not reduce posttraumatic swelling, it does
lessen metabolic activity and blood flow (126). It has been
suggested that cooling may slow deterioration and improve
neovascularization of surgical grafts (133).

Rest, ice, compression, and elevation (RICE) are the ini-
itial steps in treating many musculoskeletal injuries. For ex-
ample, icing for 20 minutes per half hour to 30 minutes every
2 hours, for the first 6 to 24 hours, is a common ankle sprain
regimen. Although ice is the mainstay of acute soft-tissue injury
treatment, recent research indicates that the evidence support-
ing its use is more limited than previously thought (134). It is
interesting that postsurgical studies of the knee (135), cesarean

![Image 63-8](https://example.com/image.png)

**FIGURE 63-8.** Ice is the mainstay of cryotherapy. Nevertheless, more sophisticated approaches are possible. This machine uses a chilled, inflated cuff to apply pressure and cooling to a knee injury.
section (136), and oral surgery (137) have reported limited benefit from the use of cryotherapy despite its widespread use.

After the first 48 to 72 hours, the choice between cryotherapy and heat is typically a matter of preference and experience. Some prefer heat and will use it unless there is a worsening of edema or pain. Others prefer cooling and believe that a combination of icing for 10 to 20 minutes to reduce pain, along with active exercising, is the most effective way to speed recovery. In any event, heat and cold are only adjuncts to a mobilization and exercise program.

Chronic Pain
Patients with chronic musculoskeletal pain occasionally find ice extraordinarily effective. Here again, the results are mixed; while some investigators have found ice massage and transcutaneous electrical nerve stimulation (TENS) equally effective for chronic low back pain (138), a recent review found limited evidence of the effectiveness of cryotherapy in low back pain (139). Ultimately, patient tolerance will limit the applicability of ice therapy.

Spasticity
Ten to twenty minutes of vigorous cooling reduces the tone of spastic muscle and may improve the isolation of voluntary function (128). Although cooling muscles before therapy might be helpful, its utility must be balanced against the time required. It is interesting that systemic cooling appears to reduce fatigue, improve balance, and increase strength in patients with multiple sclerosis who are “cryosensitive” (140). It may also produce a mild increase of spasticity (141).

Precautions and Contraindications
The precautions shown in Table 63-4 should be heeded. Pressor responses aggravating cardiovascular disease should be considered, as should the effects of direct and consensual vasoconstriction on ischemic limbs and in people with Raynaud’s phenomenon. Cold hypersensitivity, urticaria, and even frostbite are possible. Patients with insensate areas and those who are unresponsive should not be treated with cryotherapy. Cryotherapy is uncomfortable, and it is important to explain its rationale before beginning treatment.

ULTRAVIOLET
Ultraviolet (UV) radiation has a history of use as a therapeutic agent that began in ancient times in the form of helio (sun) therapy. It regained prominence in the early 20th century as a treatment of dermal tuberculosis (scrofula) and, although supplanted by more effective agents, was employed in the treatment of open wounds as recently as the 1980s.

The biomedical literature divides the UV spectrum into three parts: UV-A (0.315 to 0.4 μm), UV-B (0.29 to 0.315 μm), and UV-C (0.2 to 0.29 μm). UV-A penetrates most deeply but has the least biologic activity. UV-B produces sunburn and skin erythema. UV-C is bactericidal and also produces erythema.

Treatment
Although the last 25 to 30 years have seen UV treatment almost disappear from the physical therapy clinic, it continues to be used in dermatology for the treatment of conditions such as psoriasis and atopic dermatitis (142). However, due to its past importance in therapy and its continued intermittent use, some discussion remains warranted.

UV exposure is limited by the nature of the tissue treated, source strength, and the separation of the tissue from the applicator. Exposures are quantified in terms of the time required to produce a minimal erythema (i.e., a minimal erythemal dose [MED]). One MED (usually requiring 5 to 30 seconds with a cold quartz lamp) is established by exposing the volar forearm to a source and determining the time required to produce a minimal erythema several hours later. For reference, 2.5 MEDs cause reddening and pain that persist for a few days; 5 MEDs produce local edema, pain, and desquamation; and 10 MEDs result in blistering (143).

In general, UV treatments begin with 1 or 2 MEDs and are kept to less than 5 MEDs to avoid tissue damage (143). Open wounds are treated directly with a lamp. Specially designed orificial probes may be used for fistulas, undermined wounds, or the oral cavity.

UV-C was once a common skin ulcer treatment as its bactericidal effects (on motile bacteria but not spores) are well recognized (144). UV may also accelerate wound healing and wound margin vascularization (143,145), but alternatives such as occlusive dressings may be more effective (18), easier to use and less likely to produce side effects. It may be that the rise of antibiotic-resistant bacteria will provide a reason for increased use in the future.

ELECTROTHERAPY
The ancient Greeks knew that amber rubbed with wool attracted light objects, and also electric eels, rays, and catfish could produce analgesia (146). This knowledge had little practical benefit until the rapid increase in the understanding of electricity and electromagnetism in the 18th and 19th centuries led to new search for healing effects. Early “medicinal” electricity applications were enthusiastic but haphazard. Electrostatic baths, spark treatments, and galvanically induced limb movement all had momentary popularity. Unfortunately, their utility was minimal and scientific interest waned until the last half of the 20th century (146).

Today, high-intensity electrical stimulation is used to strengthen muscles and to move paralyzed limbs. Less intense stimulation produces analgesia and delivers medications percutaneously. Stimulation at still lower intensities has gained FDA approval for fracture healing. Soft-tissue wounds, osteoporosis, and musculoskeletal pain represent additional potentially important but still investigational electrical stimulation applications. This section discusses these agents.
Transcutaneous Electrical Nerve Stimulation

The gate theory of pain postulated in 1965 that cells in the spinal cord substantia gelatinosa serve as gates to pain perception by inhibiting the passage of nociceptive information to the brain if sensory afferent signals are present (147). TENS provided sensory afferent signals and, after some successful trials, became widely accepted.

The mechanism of action of TENS remains controversial. Although research shows that stimulation reduces dorsal horn cell activity (148), the gate theory does not explain phenomena such as painless sensory neuropathy, analgesia persisting after stimulation, and delayed onset of analgesia. As a result, other explanations, such as those involving frequency-dependent effects and central nervous system endorphins, have been advanced by a number of investigators (149).

TENS units consist of a battery, one or more signal generators, and a set of electrodes. Units are small, usually programmable, and generate a variety of stimuli with currents of less than 100 mA, pulse rates ranging from a few to 200 Hz, and pulse widths of from ten to a few hundred microseconds. Asymmetric, biphasic waveforms are favored to improve comfort and to avoid the electrolytic and iontophoretic effects associated with unidirectional currents. Additional features such as “burst” modes and wave-train modulation are common but of unclear benefit.

Electrode positioning is more art than science. Placement over the painful area is usually the first choice, but locations over afferent nerves, nerve roots, acupuncture and trigger points, and auricular sites, as well as contralateral to the pain, are also possible. Although carbon-impregnated rubber electrodes are inexpensive, more expensive self-adhesive electrodes are far more convenient.

Stimulus parameter choice is also subjective. Many prefer to begin with low-amplitude, 40- to 80-Hz “conventional” settings and use the less comfortable high-intensity, 4- to 8-Hz alternative if the first trials are unsuccessful. Ideally, initial benefit is established in a few therapy sessions and an overnight trial. If TENS appears effective, the patient should rent an identical unit for a month or more. Response is difficult to predict, and parameter selection is ultimately based on trial and error. TENS units vary in price depending on their source and complexity. Some can be quite expensive and benefits often wane with time, purchase should be considered only if benefits persist for at least a few months.

Indications

TENS studies range in quality from well-designed, prospective, randomized controlled trials to case reports. Success rates vary from placebo levels to 95% and may be affected by the choice of stimulating parameters, electrode placement, condition treated, chronicity, and length of follow-up. Even when studies are controlled, it is not clear whether the controls should be sham TENS or an alternative treatment.

A number of studies in the 1970s and 1980s found that TENS lessened postoperative and early labor pain, decreased narcotic use, and shortened intensive care unit stays. These studies also tended to find TENS-induced pain relief equivalent to that achieved by the consumption of limited amounts of analgesics/narcotics and that success rates decreased as the duration of follow-up increased.

Acute Pain

Surgical pain is a convenient model for acute pain, and a number of more recent studies have found benefits in TENS use. Benefits are not universally found and be limited to a specific condition (150). Parameter choice appears important. Thus, two RCTs that used peri-incisional either low frequency (2 Hz), high frequency 100 Hz, or a combination of the two in large groups of women having gynecologic surgery found that TENS resulted in a 30% to 50% reduction in the use of analgesics with the 2- to 100-Hz combination appearing the most beneficial (150–152).

Musculoskeletal Pain

While the earlier reports were relatively encouraging, more recent research may put a harsher light on TENS benefits. For example, a 2001 evidence-based clinical practice guidelines panel concluded that TENS did not lessen acute or chronic low back pain (153,154). Similarly, a 2005 review of treatments for acute neck pain found insufficient evidence of any benefit (155). Findings for subacute low back pain may be more positive but are still equivocal (156).

It is intriguing that while the evidence of TENS effectiveness in spinal pain seems so limited, a stronger case appears possible for its use in osteoarthritis. For example, several reviews and panels find that TENS is beneficial in the treatment of osteoarthritis (particularly of the knee). There is also evidence that benefits require treatments of 40 minutes or more duration (157).

In summary, while some patients may use TENS for years at a time, support for its use remains mixed and may depend in ways we do not understand on factors such as chronicity, parameter choice (158), and the condition (159). For example, a frequent conclusion of even large systematic reviews is that while the majority of studies may find benefit, the power and quality of the trials are too low to permit a definitive determination. In the case of rheumatoid arthritis, for example, a 2004 review found evidence of effectiveness of TENS for pain relief (80), although these benefits may not be as clear for treatment of the hand (160).

Urologic and Gynecologic Issues

While musculoskeletal pain is the main focus for many TENS applications, a significant amount of attention has also been placed on other conditions such as gynecological pain. The findings are again mixed. For example, a meta-analysis of nine RCTs found support for the use of TENS in the treatment of dysmenorrhea (161), while another systematic review of eight reports involving more than 700 women found only weak support for the use of TENS during labor (162). Interestingly, an RCT crossover trial of 43 patients with detrusor instability found 20-Hz perianal TENS was as effective as 5 mg of...
oxybutynin three times a day in reducing the symptoms of urinary urgency (163). In a similar vein, a review of evidence in this area highlighted the potential role of TENS in childhood detrusor instability (164).

Other New and Potential Applications

**Traumatic Brain Injury**

Data in this area are sparse but may grow with time to support or dispute a few intriguing preliminary findings in this area. Dorsal column stimulators (DCS) have occasionally been used in an attempt to hasten the arousal of patients with severe traumatic brain injury. Investigators have begun extending the concept to stimulation of the median nerve with 40-Hz TENS on the theory that this approach is less invasive and the median nerve has a large cortical representation (165). However, effectiveness is not established, much less a comparison with an alternative treatment such as methylphenidate is limited. TENS has also been promoted as an effective noninvasive method of pain relief in patients with traumatic brain injury (166).

**Angina**

Many consider precordial TENS use contraindicated. However, TENS may reduce cardiac ischemia, perhaps by lessening sympathetic tone, and increase cardiac blood flow (167). In addition, TENS may lessen the number and duration of silent ischemic attacks, although its effect on the total number of attacks (silent and painful) may not be significant (168). Given the concerns about the risk of this approach, more research in monitored settings seems necessary despite the fact that a number of reviews have highlighted the potential role of TENS as an alternative to pharmacological management of the pain and symptoms of angina (169,170).

**Blood Perfusion**

The effect of TENS on cutaneous blood flow or temperatures in normal subjects is controversial (171,172). However, some evidence suggests that it may improve cutaneous perfusion, increase distal skin temperatures, and lessen pain in people with scleroderma and diabetic neuropathy (173).

**Spasticity**

Spasticity reduction following TENS use has been reported for patients with stroke, spinal cord injury, and multiple sclerosis (174). Effects are controversial, but a recent review of treatments for spasticity following spinal cord injury identified TENS as the only nonpharmacological intervention supported by high-quality evidence of effectiveness (175). Only a limited proportion of patients who try TENS benefit (176). Some clinicians feel that fewer than half the patients with chronic musculoskeletal conditions find TENS helpful initially and that benefits persist more than a month in fewer than half of these. Studies of long-term users find that about 75% of individuals who have purchased a unit may be still using it after 6 months to a year (177,178), but use may continue to decline to about 30% at 3 years (179). The best results in these chronic users seem to occur in musculoskeletal pain, neurogenic pain, and angina. People with psychogenic pain, central pain, autonomic dysfunction, and social distress may respond less well (179).

**Precautions and Contraindications**

TENS has few safety issues other than contact dermatitis and skin irritation, which usually respond to changes in electrode type and placement. High current densities, either due to the setting of the unit or to a partially detached electrode, are uncomfortable but are easily corrected. Cardiac pacemakers seem resistant to TENS, but, as noted above, concern about the possibility of ectopically generated rhythms seems to dictate avoidance over the pericardial area or in patients with pacemakers, electrical implants, and dysrhythmias. Treatment near the carotid sinus and epiglottis and over the low backs, abdomens, and lower extremities of pregnant women should be avoided. Psychogenic pain is particularly resistant to TENS, and use in this setting should be judicious. The preceding precautions may err on the side of caution as there is little documentation to support their validity.

**Iontophoresis**

Iontophoresis uses electrical fields to force electrically charged or polarized ions and molecules into tissue. Speed of movement is related to voltage and field strength, whereas the amount of material introduced into the tissue is proportional to the current. Penetration is dependent on a substance’s size and polarity and may be particularly intense at sweat glands and areas of skin breakdown (180). It may not always be necessary to have an “active ingredient.” Iontophoresis with tap water at current intensities of 10 A/cm² may increase local skin temperatures by 1°C and microcirculation severalfold (181).

An iontophoretic unit is simple and may be contained within a disposable skin patch. In general, it consists of a direct-current power source, two electrodes, and a pad moistened with a dilute (often 1%) solution of the desired (charged or polar) substance placed under the electrode of the same polarity. Currents are determined by multiplying the area of the active electrode by 0.1 to 0.5 mA/cm². The size of the inactive electrode is immaterial but is kept as large as convenient for patient comfort. Electroproportion uses brief periods of high voltages to increase cell membrane permeability in gene and cell studies. Its potential application to iontophoresis seems possible.

**Indications**

Tap water iontophoresis is reported to produce benefits of 4 to 5 weeks or more in 90% of patients with hyperhidrosis (180,182). The treatment approach varies with the location treated. In the case of the distal limbs, the hands or feet may be placed in a container with both the anode and cathode or with one extremity and one electrode in each of two containers. Currents vary with the approach but are often 10 to 30 mA. The mechanism of action is unclear but may result from the preferential flow of ions along the sweat ducts.
Sodium fluoride iontophoresis is reported to reduce tooth hypersensitivity (183). Iontophoresis has also been found to improve delivery of a variety of antibiotics such as gentamicin, penicillin, and cefoxitin into the poorly vascularized tissues such as the eye and burns and cartilage (184,185).

Iontophoretic treatment of postherpetic neuralgia with methylprednisolone and lidocaine as well as Achilles tendonitis with dexamethasone has been reported to produce prolonged benefits (186,187). A recent RCT found iontophoresis to provide equivalent short-term benefits to injection of corticosteroids for carpal tunnel syndrome; however, the latter provided superior effect in the longer term (188). In contrast, RCTs of steroidal iontophoresis in the treatment of plantar fasciitis and lateral epicondylitis as well as acetic acid in shoulder tendonitis found minimal to no benefits over conventional therapy or placebo (189–191).

**Safety**

Iontophoresis is a safe therapy. Complications occur but tend to involve allergic reactions or to be similar to those of TENS.

**Electric and Low-Intensity Electromagnetic Fields**

Bone and soft-tissue injuries produce electrical fields and currents that can alter cell orientation, proliferation, calcium concentrations, and motility (192). These fields have very low intensities (approximately 1 V/m rather than the 70,000 V/m [70 mV/m] associated with cell membrane potentials), and their mechanism of action may include a switch effect that alters cell permeability.

**Wound Healing**

Interest in electrically stimulated wound healing extends back to at least the early 1970s with reports that low-intensity direct current accelerated the healing of ischemic wounds (193). These early studies were marred by poor design, and while the approach was not widely accepted, the interest never disappeared. Today, protocols often similar to the earlier studies use low-frequency (10 to 200 Hz) and low-intensity (10 μA/cm²) electric currents and electromagnetic fields to speed the healing of soft-tissue wounds. Benefits remain difficult to establish, but there is some experimental support for the approach, and the well-accepted benefits of electrical stimulation on nonhealing bony fractures (194) has only served to maintain interest.

Wound studies are notoriously difficult to carry out, and unfortunately, this has proved true for those assessing the benefits of electrical stimulation. Thus, while TENS stimulation has been reported to improve healing and elevate distal limb temperatures in patients with diabetic neuropathy and scleroderma (195), the studies are usually limited by sample size and may or may not show stimulation accelerates healing. In fact, two recently updated Cochrane reviews concluded that the database was too limited to permit an assessment of benefits of electrical stimulation in the healing of either venous or pressure ulcers (196,197).

**Musculoskeletal Pain**

Musculoskeletal pain has proven to be a focus of electrotherapeutic research with a particular emphasis on pulsed EMF (PEMF). The results are again intriguing but less than convincing. For example, while controlled studies have reported that low-intensity pulsed shortwave diathermy (27.12 MHz, 1.5 mW/cm²) produced clinically and statistically significant improvements in neck pain or ankle sprain (198,199), a more recent investigation involving patients with subacromial impingement failed to find benefit from the addition of PEMF to their rehabilitation program (200). Reviews provide a similarly mixed picture. Again, sample size is an issue. Thus, while a 2002 Cochrane review that encompassed only three studies determined that electromagnetic fields were more effective than placebo in the treatment of osteoarthritis (201), a more recent 2006 review focusing on osteoarthritis of the knee that included five RCTs found no convincing evidence of benefit (202). Ultimately, acceptance of electrical and electromagnetic stimulation as a treatment option will depend on cost, comfort, convenience, and evidence of effectiveness.

**Interferential Current**

TENS units and muscle simulators are limited at times by the discomfort that strong stimulation at low (e.g., <80 Hz) frequencies produces in the skin. However, skin impedance decreases with frequency, and higher frequency waves can penetrate to deeper tissues without discomfort. Interferential current (IFC) devices (Fig. 63-9) take advantage of this fact by having two high-frequency (e.g., 2,000 to 4,000 Hz) sine waves differing by 20 to 100 Hz overlap each other in tissue. Beat frequencies equal to the sum and differences of the sine waves are produced with the difference in frequency within the therapeutic 20- to 80-Hz range.

**FIGURE 63-9.** IFC treatment. This treatment uses the interference pattern of two higher frequency sine waves to apply electrical stimulation. Although only three electrodes can be seen in the picture, four are required (two for each wave train).
IFC appears useful (particularly in a sweep or vector mode) when TENS effects are desired. Despite this, there is no clear demonstration of its superiority over other devices: for example, a recent trial on treatment of dysmenorrhea found no difference in the effectiveness of interferential or TENS stimulation (203). Evidence to date for musculoskeletal pain management has been mixed with positive (204,205) and negative findings (206) reported. Interestingly, studies using experimental models of pain in humans have consistently failed to show any significant effect with the use of interferential therapy (207–209).

Safety
These devices operate at low powers and are used for limited periods. Although there has been a report of a burn arising from treatment with IFC following knee arthroplasty (210), the devices appear safe, and the precautions necessary seem to be those outlined for TENS and electrical devices in general.

Epidemiologic studies have at times associated prolonged exposure to low-intensity EMFs (such as those associated with power lines, hair dryers, and cell phones) with increased rates of leukemia, miscarriage, and brain tumors (Table 63-5). These studies are of concern, but all have methodological shortcomings. A number of panels, including a 1996 National Research Council review panel, have met over the years and concluded that while extremely low-frequency (ELF) magnetic fields 0.4 T may be associated with increased rates of childhood leukemia, neither static magnetic fields nor ELF electric fields pose any known risk (211,212). This question has gained new importance with the advent of cell phones but at this point, risks seem minimal and, if present, subtle.

Alternative Therapies
Vibration
Vibration and tapping have had a long-term role in neuromuscular rehabilitation as a way to facilitate muscle recruitment (213). Traditionally, vibration was applied to a limited number of limb muscles with a handheld applicator. Research is limited but many clinicians believed that frequencies of about 150 Hz and amplitudes of 1.5 mm were particularly effective (213, 214) and might have some analgesic (215) or wound-healing effects (216).

Vibration, however, has recently been regaining attention. Much of this interest appears based on clinically observable improvements in recruitment that localized vibration produces in the muscles of patients with neurological conditions such as stroke and Parkinson’s disease. Gait and balance form a particularly strong area of investigation. In these cases, the whole body is typically vibrated with the patients standing on a platform while their body is subjected to small amplitude rhythmic displacements. Initial results in terms of improvement mobility and balance are mixed and require further assessment (217).

Light and Laser Therapy
Low-power lasers (and monochromatic light sources) have been used since the late 1960s to lessen the pain and speed the healing of a variety of neuropathic, inflammatory, and soft-tissue conditions. Initially, treatment involved short exposures to a variety of ≤1- to 5-mW lasers. With time, therapy practices have coalesced, and now, most treatments are performed with 30- to 150-mW 1R diode lasers (Fig. 63-10). Although treatments do not elevate tissue temperatures more than a few tenths of a degree, laboratory studies find irradiation stimulates collagen production, alters DNA synthesis, and may improve the function of damaged neurological tissue. Extension of these effects to humans has been more difficult.

Laser therapy received its first impetus in the 1960s from claims about its ability to heal chronic, nonhealing wounds (218). This claim has been surprisingly difficult to substantiate. Thus, while the underlying mechanisms have been well researched at cellular level, support for benefits has been progressively more difficult to establish at the animal and human levels. However, at least one well-controlled study of superficial wound healing found benefit of laser irradiation (219).

Use of this approach lagged in the United States due to the previous lack of FDA approval (220). The FDA reduced its requirements in 2002 to that of a single positive clinical study and has subsequently approved more than 25 devices for adjunctive use in the treatment of pain.

As is true for many modalities, confirmation of benefit is conflicting and confusing. Thus, several systematic reviews have provided evidence of effectiveness of these devices in the short-term alleviation of joint pain (221), knee osteoarthritis (222), and temporomandibular joint pain (223). Surprisingly, these benefits do not seem to extend to neck pain (224) or lateral epicondylitis (225) and in the shoulder may be limited to adhesive capsulitis (226). Similarly, while several trials have indicated potential benefits in the treatment of fibromyalgia (227) as well as part of a program for the management of chronic low back pain (228), a 2007 review of treatments...
for carpal tunnel syndrome found conflicting evidence for laser therapy (229).

**Static Electric and Magnetic Fields**
The potential health benefits of electric and magnetic fields have intrigued humans for thousands of years. Interest has fluctuated with time, but the magnetic devices advertised today differ little, other than in complexity, from the magnetic rings of the Greeks and therapies of 150 years ago.

There is support for the idea that these fields might have clinical utility. For example, low-intensity static and dynamic magnetic fields may alter balance, learning, pain perception, behavior, and anxiety (230,231). Again, however, findings are mixed, and other studies fail to find these effects (232,233).

**Indications**
Unfortunately, although there are a number of alluring reports, objective clinical support in either the human or the veterinary literature is limited. Thus, although a study may find magnetic fields beneficial in conditions such as postpolio pain, peripheral neuropathy, rheumatoid arthritis, and neck and shoulder pain (234,235), others may find no effect (236,237).

**Safety**
Safety does not seem to be a particular issue. As noted earlier, a number of panels have concluded that neither static electric nor magnetic fields pose any known risk (211,212). As a point of reference, the FDA has classified static magnetic fields of less than 4 T as nonsignificant risks for research subjects (238). Even hour-long exposures to 8-T field strengths, although they do produce temporary changes in ECG readings, do not alter heart or respiratory rates, blood pressure, cardiac output, or body temperature (239).

**Extracorporeal Shock Wave Therapy**
Although extracorporeal shock wave therapy (ESWT) has gained prominence over the last 15 years as a potential treatment of a variety of soft-tissue disorders (58), its potential for plantar fasciitis and calcific tendonitis seems particularly promising. A number of controlled studies have found ESWT effective in the treatment of calcaneal bone spurs (240) and calcific tendinitis of the shoulder (241–243), although results for lateral epicondylitis appear more equivocal (244,245). Better targeting of the treatment area as well as higher levels of stimulation may produce better results than generalized or low-intensity treatment, but whether ESWT has any unique benefits or is merely a new way to apply controlled trauma to reinitiate healing is yet to be established.

**MODALITY CHOICE AND PRESCRIPTION**
The physical agents are typically prescribed as part of a program that may also include massage, exercise, and education (Table 63-7). As this table shows, a well-written prescription involves the same elements that are common to all good writing: who (the patient), what (the agent), why (the diagnosis), where (the treatment area), when (the frequency and duration), and how (intensity, device settings).

Modality choice depends on balancing of the diagnosis, the characteristics of the agents, evidence of effectiveness, concurrent issues (e.g., level of cooperation, anticoagulation, preference), and treatment goals. General rules introduce some order into the situation. For example, acute (<24 to 48 hours) musculoskeletal conditions are usually treated with cooling. Hot packs, cool packs, hydrotherapy, shortwave diathermy, and some of the electrical therapies are commonly used to treat broader areas. More intense agents such as ice massage and ultrasound are more common if smaller regions are to be treated. The diathermies are frequently favored for deeper tissues, but the comfort-inducing effects of superficial agents may be as beneficial. In the end, choice involves blending a physiologic understanding of the agents, experience, preference, and equipment availability. As research continues, our knowledge will grow. Choices common today will undoubtedly appear quaint in the future.

**REFERENCES**


Manipulation, Massage, and Traction have been used in various forms for several thousand years throughout the world. Each includes treatments for the relief of painful conditions—sometimes with noticeable immediate effects—that are sought by an ever-increasing number of people; as an example, manual medicine procedures, including manipulation, are very much in demand by patients, of whom a reported 12 to 17 million Americans alone, who seek, receive (over 90 million manipulations alone) and are very satisfied with them on an annual basis (1,2). Despite consistent high demand for these treatments, evidence of the effectiveness of some of the techniques comprising them has been difficult to obtain and complete acceptance within the medical community has, to date, been mixed.

Some argue that patients who improve and report feeling better with these treatments may have done so anyway and that serious complications (which are rare and usually due to either an unqualified practitioner or the performance of an inappropriate maneuver) sometimes occur; that treatments are aimed at the pocketbook and not the pain; and that claims of success are sometimes unproven (3). The Agency for Health Care Policy Research (AHCPR) (4) recommendations that include manipulation, for example, as an option in acute low back pain management have done little to resolve these differences (5).

What has led to this apparently dichotomous thinking? One factor may be that some traction, manipulation, and massage techniques have an as yet uncertain mechanism of action. In addition, they are treatments that can be performed by nonphysicians. A treatment appears less scientific and perhaps less efficacious if it can be obtained from nonphysician, even nonlicensed, practitioners. A more acceptable rationale for rejecting these treatments may be the paucity of scientific proof of benefit.

The use of any treatment should be limited to those conditions that can reasonably be expected to improve with the physiologic effects of that treatment. The macroscopic effects of traction and massage have been partially elucidated and offer some degree of guidance regarding indications. This is not yet the case for manipulation, although evidence is growing. In the remainder of this chapter we will discuss what information is available regarding the use and efficacy of these treatments and attempt to provide the physiatrist with a basis for understanding their use and role in the treatment paradigm.

**MANIPULATION**

**Introduction**

Although manipulation has been widely practiced in almost all countries of the world since at least the time of Hippocrates (6), the subject often evokes strong emotional responses among health care professionals, particularly in the United States and Great Britain. Since the 1890s, when manipulation became a cornerstone of the therapeutic approach of the Osteopathic and Chiropractic Schools of Medicine, the subject has caused debate among practitioners.

Manual treatments are among the most common procedures utilized on a daily basis across the globe; they are performed by physicians, physical therapists, chiropractors, nurses, athletic trainers, naturopaths, massage therapists, lay practitioners, and others. Early references to manual medicine can be found in the writings of Hippocrates, as mentioned above, and in ancient Egyptian hieroglyphics and early Chinese medical literature (7). In the United States, the roots of manual medicine can be traced to Andrew Taylor Still, M.D., and his contemporary, Daniel David Palmer. The former started his American School of Osteopathy in Kirksville, Missouri, in 1892, the latter opened the first chiropractic college in Iowa in 1895 (8). Today, manipulative treatments continue to be developed and researched nationally and internationally by an ever-growing cadre of practitioners.

This section provides a brief review of manipulation, describes some of the forms of treatment, and discusses the use of manipulation and its prescription, with a focus on physiatry, and physiatric application of manual medicine. The techniques described will be neither osteopathic, nor chiropractic, per se, but rather a combination of terminology and treatments a physiatrist is most likely to encounter.

The purpose of this section is to introduce manipulation and discuss its applicability to physiatry and to present some general guidelines for appropriate prescription. The evidence base for these treatments will be referenced where applicable and available, but we will neither attempt to settle ongoing disputes concerning the efficacy of manipulation nor comprehensively cover all forms of manual treatment or diagnosis.

**Definition**

The definition of manipulation is controversial, but a consensus definition is that it is “the use of the hands in a patient’s management process using instructions and maneuvers to achieve...”
maximal painless movement of the musculoskeletal system and postural balance” (8). More specifically, spinal manipulation is a mechanical treatment applied to a specific vertebra or vertebral region, including the sacroiliac region and the rib cage, by a physician or therapist with the primary goal of restoration of lost vertebral motion. Nonthrust techniques (“mobilization” to a physical therapist) are encompassed by this narrower definition. Active exercise, self-applied forces, or self-induced motion, although occasionally effective in the restoration of vertebral mobility, are not included. Massage, which is applied only to soft tissue, and traction, which is nonspecific to a vertebral region, are excluded by this definition, although the reader will recognize that some overlap among manipulation, traction, and massage occurs. The interested reader will find a detailed history of manipulation in a number of sources (6,9,10).

While the majority of manipulative treatments performed in the United States are aimed at the cervical and lumbosacral regions to treat musculoskeletal symptoms, manipulation can be performed on any part or region of the body. It is generally directed at restoration of normal motion and elimination of pain secondary to disturbed biomechanics. It is an application of forces to the muscles, tendons, ligaments, joints and capsules, bones and cartilage of the vertebral column, or other tissues, with a primary aim to restore the body to a state of homeostasis (11). Although often taught in a regimented way, the application of manual treatment should not be formulaic or protocol driven, but, rather, customized to the individual patient and based on a careful and thorough history and physical examination. Assessment involves the examination of the restriction of motion and the asymmetry of the joints in a given area of the body, as well as that in contiguous and adjoining regions. Generally, one tests for areas of tissue texture changes, pain in or sensitivity of the area, or a change in the quality of motion within a body region.

From a biomechanical perspective, an area of structural hypermobility or hypomobility, such as in adjacent vertebrae or other skeletal or muscular structures, is most often a consequence of a problem elsewhere in the body. When a certain degree of motion is expected in one area and it is deficient, another region then typically compensates for this with its usual motion, plus at least a few degrees more. If, for example, a stroke patient with hemiplegia presents with low back pain, one can treat the restrictions and dysfunction found in the lumbar spine and, if present, in the pelvis as well. These treatments may give the patient substantial relief of lumbar pain, but the pain and dysfunction will soon return if there is an overlooked causative restriction elsewhere. Consider, as an example, the patient's lower limb weakness; the altered mechanics of the patient's gait alone will affect both sides and compromise both pelvic and lumbar motion. Manipulation alone would be inappropriate as the neurologic deficit is not amenable to this form of treatment. Using manipulation to treat the dysfunction of the lumbar spine and other areas, in combination with physical therapy, assistive devices, and other appropriate physiatric care, may optimize this patient's functional recovery and lessen the severity of the low back and pelvic symptoms.

The majority of the studies and review articles pertaining to manipulation have focused on the application of these treatments to the cervical and lumbar spine and on the use of thrusting techniques in particular. It is important to realize that there are many other types of manipulative techniques and which one is chosen should be based on a thorough examination of the patient and the musculoskeletal lesion to be treated. It is essential to find the cause of the illness or dysfunction if the most efficacious treatment or combination of treatments is to be chosen and applied.

Techniques of Manipulative Treatment

In the following introductions to the most common techniques of manipulation used in the United States, discussion is limited to those that the physiatrist is most likely to encounter early in the investigation of manipulative care.

The concept of direct and indirect treatment is ubiquitous in the literature regarding manipulative technique. As previously mentioned, an accurate and thorough examination and localization of musculoskeletal dysfunction are essential. Before initiating treatment, any restriction in range of motion, either gross or subtle, should be detected during passive motion testing. This restriction is commonly termed a barrier (this is fully explained below). Direct treatments engage the barrier and focus mechanical forces into the motion restriction to eliminate it. In contrast, indirect treatments avoid the barrier and instead focus treatment within the joint range that allows the most fluid motion in order to expand normal range by neuromechanical means.

Thrust Techniques

Other terms for thrust techniques (9,10) include impulse, which is of European origin (9), high velocity/low amplitude, adjustment, and spinal manipulative therapy (SMT), which are of chiropractic origin. In most approaches, the operator diagnoses the dysfunctional vertebral segment by identification of position or motion abnormalities or by related tissue changes, including tenderness to palpation or induced motion. He or she then rotates, side-bends, and flexes or extends the vertebral segments below it, whereby engaging or “locking” the facet joints so that further motion is limited to the vertebra in question. This vertebra is then passively moved to its limit of motion (i.e., its barrier) to remove all slack motion. A small force localized to the joint in question and in the direction of the restricted motion is applied to hold this position. Finally, a brief (9,10,12) controlled thrust is applied to the involved vertebra in the restricted direction, and a small motion, flexion, extension, rotation, or side bending, occurs as the vertebra traverses the barrier (Fig. 64-1). Often an audible “pop” or “click” is produced. The source of this sound has been attributed to cavitation (13) and nitrogen release in the synovial facet joint. Its presence infers that a mechanical process has occurred over a much shorter time than for inaudible events, such as reflex muscle contraction. This further suggests that audible thrusting procedures have a different mechanism of action and possibly different level of efficacy compared with
other manipulative procedures. It is arguable whether there is clinical relevance to this sound (9,10,11). All thrust techniques can potentially do harm if forces are not well localized and/or misdirected along axes which do not correspond to correction of the dysfunctional segment at the vertebra or vertebral region involved. Again, precision of localization and placing the joint or segment into its restrictive motion barrier so that the physician has maximal control of force applied and uses only the force sufficient to create or restore motion is desired to thereby complete the treatment (14). Ideally, one should use the minimal force over the shortest distance possible to achieve the desired results.

Articulatory Technique
The articulatory technique (9,10,15), also referred to as mobilization in Europe and low velocity/high amplitude in the United States, is very similar to oscillatory (16) techniques performed in Australia. The vertebral joint is passively moved within the reduced range defined by its resting position and the dysfunctional limit of motion. The extent of motion and its endpoint are varied, but ultimately the endpoint and dysfunctional barrier become the same, and the barrier is teased with repeated motion. Ideally, the quality or feel of the induced motion and the quantity of force and excursion are normalized by this procedure. Occasionally, a small amount of additional force takes the vertebra through its barrier or restriction, and this technique may then appear like a thrust. Other variables include frequency of repetition, duration of “hold time” at the extremes, and the rate, rhythm, and amplitude of oscillation (depending on underlying tissue texture) (16).

Positional Techniques
The underlying principle in counterstrain and functional technique is that hypomobility is caused by an inappropriately firing muscle, that is, active tissue, rather than a shortened passive tissue such as capsule, myofascial, or ligament. Thrust, articulation, and muscle energy techniques all employ forces that might be expected to lengthen shortened passive tissues, whereas the positional techniques attempt to change an inappropriate engram, or central pattern, of vertebral muscle behavior, similar to what is done in rehabilitation of head injury and tendon transplants.

Counterstrain
Counterstrain (9,16,17) is an indirect myofascial technique developed by Jones (18) that shares with functional technique (discussed later) the emphasis on relative positioning of a joint or body part as an essential component of the treatment. This technique involves placing the joint or body part into its position of maximal “ease” or comfort, relaxing myofascial or ligamentous soft tissues. This allows inappropriately shortened muscles to “reset” their spindles, which then are thought to normalize their proprioceptive input to the spinal cord. Generally, the restricting muscle is overly shortened by this positioning (counterstrain), and its antagonist muscle is gently overstretched (strained) in the process. The treatment position is found by palpatory pressure and minimizing the pain at an associated counterstrain “tender point” to less than 80% of its initial degree of tenderness (15,17,19). This position is then maintained and held for 90 seconds for vertebra and other points, except for ribs which are held for 120 seconds (19). On occasion, small “fine-tuning” passive repositioning movements with feedback from the patient may be needed. Because tenderness is a part of this feedback system, the patient must respond to the questions of the practitioner regarding their degree of pain or “tenderness” at the noted point(s). The patient is then slowly returned to the neutral position, usually in one plane of motion at a time, to prevent resumption of the inappropriate muscle firing. Structural reassessment is then repeated.

The tender points of counterstrain do not coincide with the trigger points of Travell (252) or the similarly named tender points associated with fibromyalgia. Those of
counterstrain are usually small, discrete, fibrotic areas thought to be manifestations of distant arthrodial or myofascial restriction (9); they are not paired but may be associated with the other features of fibromyalgia such as stiffness, fatigue, and sleep disorders. They are distributed throughout the body in reproducible locations depending on the nature and location of the associated biomechanical dysfunctions. They are similar to those found in fibromyalgia in that they cause local pain when palpated rather than in a distant referral pattern. Counterstrain is considered to be gentle, safe, effective, andatraumatic; because of this, it is considered to be useful in elderly and hospitalized patients, in children, and in patients with fear of pain. This technique is easy to perform and relatively forgiving. It is easily incorporated into a home exercise program as well.

Functional Techniques
Functional techniques (9,12,20), developed in the 1940s and 1950s, share with counterstrain a methodology oriented to reducing inappropriate afferent impulses from nociceptors and mechanoreceptors (and the resultant efferent α motor activity to the skeletal muscle) by placing the joint or body part into a position of maximum ease (11). Unlike counterstrain, the position is found and monitored by the practitioner sensing increased tissue tension to small, induced motions. The patient provides “respiratory assistance” as he/she inhales and exhales to varying degrees to facilitate the release of the involved soft tissues.

Practitioners feel that passive positioning to maximum ease allows the firing pattern of the aberrant muscles to reset such that they are normalized in a neutral position. This approach, unlike counterstrain, does not make use of tender points and is, in principle, more objective because only the practitioner’s palpation determines the positions of balance. Functional techniques (nontraumatic and repeatable) are useful in both acute and chronic conditions.

Muscle Energy
This nontrusting technique (9,10,21,22) is also called isometrics and, by some Europeans, like articulatory technique, mobilization. Similar to proprioceptive neuromuscular facilitation (PNF), it uses stretch reflexes and isometric contractions to relax the target muscle (23–25). The muscle energy technique requires the operator to position the patient into the motion restriction barrier and take up slack as in direct thrust techniques. Instead of an impulse against the restriction, however, the practitioner instructs the patient to exert a small isometric force away from the barrier. Simultaneously, the practitioner provides equal resistance, preventing active motion. The patient holds the isometric contraction 5 to 10 seconds. After the patient relaxes, the affected joint can passively move further beyond the original restriction to a new barrier. The contract-relax effort is often repeated several times, engaging successive barriers with sequentially improved motion.

Soft-Tissue Techniques
The soft-tissue approach uses mechanical stretch of the skin, muscle, and fascia to increase their motion. Lateral stretch (or “bow-stringing”), linear stretch, and deep inhibitory pressure are the most common procedures. This approach is useful in virtually all patients, often as a first step in treatment involving multiple manipulative approaches. Encouraging overall circulation and enhancing venous and lymphatic flow are central to these techniques (9). The overall purpose of soft-tissue technique is to relieve superficial muscle and fascial tension.

Myofascial Release
The myofascial release approach (9,26–28) to manipulation overlaps somewhat with massage and, conceivably, with traction. It can be directed at hypomobility of a vertebra, of a vertebral segment (as in spinal manipulation), or of the entire body, in which case it resembles some massage techniques. It can be indirect, whereby a restricted area is placed into a position of little resistance until relaxation occurs, or direct, in which case the affected area is placed against the restriction barrier with constant force until a fascial release occurs. Conceptually, in this approach, all the myofascia in the body are interconnected, and when one area is tight or restricted, diminished movement results locally and in more distant areas.

Practitioners continually palpate to assess tissue response and adjust the applied forces of stretch, traction, pressure, or twist they are using, until the affected tissues are felt to change, often called a “release.” The actual forces are combinations of those used in massage, soft tissue, muscle energy, and possibly craniosacral techniques. The mechanism of the release of the “tightness” can be biomechanical (such as viscoelastic strain) or neuroreflexive, but, when accomplished, fascial resistance to forces applied should be symmetric, and the tissues should be relatively mobile, that is, responsive to the force.

Joint Play
This concept addresses the immediate cause of the joint motion restriction. In an idealized joint, the bone moves in small precise ranges, independent of voluntary muscle contraction, over the stationary contours of opposing joint surfaces as motion in the voluntary direction occurs. But actual joints sometimes have noncongruent and sometimes diseased nonsmooth surfaces, resulting in movement in an additional perpendicular or tangential direction to the plane of motion; this motion is called joint play by manipulation practitioners and accessory motion (29) by many physical therapists. Joint play includes motions in a nonvoluntary direction as well. Thus, motion restriction in a direction perpendicular to that of voluntary motion will limit voluntary motion in the voluntary direction. This concept leads naturally to manipulation because the practitioner can “gap” (11) the restricted joint allowing a return to a symmetrical balanced seating of the joint, whereas voluntary or active motion cannot.
Craniosacral Manipulation
Craniosacral manipulative treatment (9,10,30,31) is an approach based on the concept of a primary respiratory mechanism, a cyclic, palpable, rhythmic wave of inherent motion, which corresponds to the Traube-Hering-Meyer wave (a rhythmic variation related to vasomotor tone), and is most easily appreciated in the cranial and sacral areas (32). It is not clear whether this wave is related to low-frequency waves noted in cerebrospinal fluid (CSF) monitoring (32). This primary mechanism is described to include the inherent mobility of the central nervous system (CNS), CSF fluctuation, the articular mobility of the cranial bones, involuntary motion between the ilia, of the sacrum, and the mobility of intercranial and interspinal membranes (30).

The practitioner palpates the head, sacrum, or anywhere on the body to feel pulsations of the postulated wave motion in the 8 to 12 per minute range and for symmetry, amplitude, regularity, and frequency of the wave. In infants and those with head injury, it is easy to accept the concept of motion, but in the normal adult, motion at the cranial sutures must be quite small, animal studies notwithstanding (33). Whether elastic distortions could give rise to these motions or to palpable changes in elastic compliance (34) is unclear. When abnormalities are found, the practitioner uses gentle pressure on the skull and sacral areas to restore the wave to its normal symmetry, rhythm, and amplitude.

At this time, this technique remains controversial, even among some practitioners of manipulation. However, its largest subset of potential patients, infants with birth defects and patients with head injury, intersects strongly with the patient base of many physiatrists, and its recent growth in popularity demands inclusion in any current discussion of manipulation. Importantly, there is some indication that in the head injury population, cranial manipulation can produce undesirable results (35).

Hypothesis of Mechanism of Action
Manipulation is essentially a mechanical intervention that alters either the mechanical behavior of passive soft tissues or the neural control of the active skeletal muscles that cross a vertebral joint. It is not surprising that almost all the various components of the conceptual framework of manipulative medicine have position or motion of the joint and the soft active and passive tissues that determine these positions or motions as their unifying core. But other than this essential portion, there is wide divergence among practitioners and researchers as to the nature of the dysfunctional problem, its etiology, and what the intervention does to treat it. It is helpful to recognize that various terms refer to the same entity; for instance, minor intervertebral derangement, osteopathic lesion, chiropractic subluxation, manipulable lesion, joint blockage, segmental dysfunction, and somatic dysfunction are all used to describe the same entity. Generally, some pathology, such as fractures, sprains, strains, tears, avulsions, tumors, and joint inflammation, may be excluded from the various hypotheses because it is generally recognized that manipulation would rarely be considered as primary therapy for those conditions. Fibrotic contractures, sprains, and strains may be exceptions because these may be amenable to manipulation, including, but not limited to, thrust (36).

There are several hypotheses aimed at explaining the cause of somatic dysfunction. Some of the more common of these will be briefly mentioned here, especially as they are more likely to be encountered in physiatric practice.

1. Barrier models have been described by several investigators to explain palpable findings (9,10,37,38) (Fig. 64-2). Normal joints possess an active range of motion and a larger passive range of motion. A barrier or motion restriction may result from abnormal muscle contraction, or capsular or ligamentous shortening across the vertebral segmental joints either in the direction of the intended motion or perpendicular to the joint surfaces that compress them. When a barrier forms between the neutral position and the usual limit, the patient cannot achieve normal range; the manipulator needs additional force to achieve the normal passive range. Asymmetry of the static force may lead to a new resting position for the vertebra, away from anatomic neutral. A variant of this approach considers strain of soft tissue as the basic pathologic condition (23); the resulting inflammation and edema then start a local vicious cycle resulting in pain and hypertonicity.

2. The facilitated segment model assumes that a vertebral body, chronically malpositioned by contracture or overly active muscle, floods the segmentally related area of the spinal cord with inappropriate nonfatiguing proprioceptive impulses. These in turn spill over and facilitate outgoing motor neurons and autonomies in the same vertebral segment of the cord (39). Thus, pathways are present for interaction between soma and viscera at related segments, and palpatory diagnosis of visceral disease and manipulative influence on the viscera are possible. This also allows abnormal segmental areas that were asymptomatic to develop symptoms from general illness, emotional stress, or distant disease. Physiatrists and other physicians familiar with reflex bladders and autonomic hyperreflexia will probably accept the concepts of somatovisceral reflexes and segmental spill-over, but may find the magnitude and importance of these effects in the spine-intact human to be unclear (40,41). Manipulation is thought to release the proprioceptive input by bringing the vertebra to its normal, symmetric position.

3. Numerous hypotheses have been offered by the chiropractic profession to explain the relationships between altered structure and resultant pathologic conditions. These are adequately reviewed by Haldeman (40,42–53). The deleterious neurophysiologic effects that result from nerve compression caused by vertebral derangement have been conjectured to be a major cause of both somatic and visceral complaints. This viewpoint, prominent since the
earliest foundations of chiropractic, is perhaps the one most associated with chiropractic by other professionals. However, other concepts, including neurodystrophic effects, viscerosomatic reflexes, and a proprioceptive insult phenomenon similar to the segmental facilitation model discussed earlier, also appear.

4. The concept of “ease and bind,” central to functional and other indirect techniques, uses a picture of vertebral motion around a symmetry point where motion in any direction encounters equal tissue resistance (Fig. 64-3). In a lesioned vertebra, the balance point is not the anatomic neutral, so that motion in one direction is met by more resistance (bind) than another (ease) (Fig. 64-4). The maintenance of the asymmetry arises in proprioceptive and γ motor spindle sensors, and treatment usually involves passive motion to and maintenance of the ease position so that the spindles can reset their output. The strain-counterstrain method suggests that the aberrant activity of the spindles can trigger reflex patterns and trigger points, which can be used to monitor the positional resetting of the spindle output (9, 20).

5. The concept of tight-loose refers to the tightness of the fascia around a stimulated agonist muscle and the accommodating looseness of the fascia around the antagonist muscle. This is a central theme in myofascial release approaches to manipulation. The “release” phenomenon is the practitioner’s sense of relaxation of the tight fascia when treatment is successful (9, 10).
The possible therapeutic effects of manipulation include the following:

1. Passive stretch or even microtrauma (36) to the soft non-contractile tissues (ligament, capsule, tendon, fascia) whose contracture is limiting the vertebral range of motion.

2. Relaxation of inappropriately contracted skeletal muscle whose shortening has limited vertebral joint range of motion. This can happen as a result of changes in the afferent branch of a reflex, by resetting of the muscle spindle equilibrium point or by a change in afferent activity at the vertebral joint’s range endpoint (40,41).

3. Flooding of the cord with proprioceptive or kinesthetic signal to close the gate on pain (54).

4. Endorphin release, although, as noted previously, the data are contradictory on this possibility (43).

5. Placebo effect may be present in manual technique (43). The concept of the therapeutic benefit of human contact, or touch, should not be discounted when discussing any manual treatment.

It should be clear that manipulation success is reported for many, if not most, patients treated with widely varying techniques by practitioners working with very different assumptions as to the cause of the problem and the therapeutic action of the interventions. It would seem that any procedure that induces muscular relaxation or vertebral motion has some chance of success. Because most of the postulated actions of manipulation are accomplished by most techniques, success itself for each technique provides little feedback as to the validity of the various concepts. All, or possibly none, of the hypotheses presented may ultimately be found valid. The placebo effect, which may be diminished or enhanced when the manipulation is relegated to nonphysicians, is impossible to objectively gauge or compare among the various approaches.

Choosing Manipulation

Risks and Benefits

Few risks are involved in manipulative treatment. Complications resulting from isometric or articulatory treatment have not been reported. There are reports of complications of thrust manipulation to the cervical spine, but the number of reported problems, which is in the hundreds (55), is actually quite small when compared to complication rates of spinal surgery or the use of NSAIDs.

Most recently, cervical manipulation has received attention in the lay press in connection with the development of permanent neurologic sequelae, sometimes with strokelike symptoms (56). This serious and most often reported side effect seems to be related to thrusting style techniques applied to the cervical spine, especially when this region is placed in a position of extension, and without localization to a single vertebral segment at the time of the thrust. It is believed that this type of manipulation, with the cervical spine maximally extended, affects the vertebral artery by causing embolization of plaque material lining the artery. Some basic precautions, including neurological screening examination, with assessment of vertebral-basilar symptoms, and persistent monitoring of the patient during treatment, as well as careful patient selection, can minimize the risk of this rare, but serious, adverse effect.
Osteopathic and chiropractic proponents of manipulative treatment believe in distant visceral benefits, beyond the systemic benefits produced by good structural and biomechanical efficiency and relief of pain (8). However, distant visceral complications temporally related to manipulation have been documented only twice in the past 100 years (57,58); therefore, the physiatrist may safely evaluate the benefits or detriments of manipulation solely on musculoskeletal results.

Benefits of manipulation have been definitively proven to be as effective as NSAIDs and traditional care in the treatment of acute back pain and consensus is growing because it may offer relief in the treatment of acute and chronic low back pain (4,59). Proponents anecdotally report excellent results in treating acute musculoskeletal problems and good results in treating chronic conditions. Empirically, these outcomes are comparable with those achieved with “conventional” modalities, which also often carry no proof of efficacy. The growing number of patients obtaining manipulative treatment indicates at least some degree of perceived benefit.

Few studies have investigated the use of manual medicine in combination with other “standard” treatments, in part because of the potential difficulty separating the additional benefit of manipulation from its counterpart treatment. Most recently, studies have investigated the use of “standard” physical therapy and exercise with and without manual treatments for lumbar spinal stenosis and the use of osteopathic manipulative treatment with acupuncture to treat acute low back pain (60,61). Each of these studies, though small in size, and differing in methodology, found favorable results with combination treatment involving manipulation. Perhaps this is where future research effort, particularly in the physiatric application of manipulation, should be focused.

Applicability to a Physiatrist’s Practice

The physiatrist should be able to identify, through a focused history and physical examination, that subpopulation of patients most likely to benefit from the incorporation of manipulation into their treatment plan. Although some manipulative techniques may have hospital applicability, for example, rib mobilization in patients with either pulmonary disease or complications from surgery, most patients with conditions appropriate for manipulation are in the outpatient setting. This group includes people with structural problems such as pelvic asymmetries, and vertebral rotations, and others whose diagnosis may rely on pulmonary expertise.

While there is a relative dearth of solid research on the use of manipulation, the bulk of what is available addresses what tend to be acute and chronic pain conditions, especially low back pain, treated in the outpatient setting. From these studies, the use of manipulation in acute conditions has been shown to be the most accepted indication (62). The use of manipulation for chronic spinal pain is not yet generally accepted (1,63). Studies (64) are difficult to perform because total eradication of chronic pain is rare, regardless of treatment method employed. Manipulative treatment in these cases should be directed at the restoration of normal spinal segmental motion, thereby reducing pain, with the hope that overall physical activity might increase (65). As with all treatments for chronic pain, whether there is demonstrable evidence of injury or not, the goal is to assist the patient to adapt, compensate, and continue to function well. An improved understanding of his or her condition can facilitate increased activity, social and vocational reintegration, and self-efficacy.

Objective evidence of the usefulness of manipulation for chronic pain is becoming more apparent. Randomized prospective clinical trials have shown improvement in pain complaints, increased paraspinal pain threshold levels, improved range of motion, and faster return to work (64,66–68). Patients with chronic back pain might be referred for a course of manipulation in conjunction with epidural injections, transcutaneous electrical nerve stimulation (TENS), psychological counseling, and other interventions.

In the inpatient setting, there is little published objective evidence for the use of manipulation. Much of what is written is anecdotal or case-study based. Historically, manual care was successfully used in the 1917 to 1918 influenza epidemic to assist respiratory function and resulted in significantly improved survival rates in persons treated with manipulation (69). Other small studies have been done with hospitalized pneumonia patients and other groups. Extrapolating from these data, it would seem that treating the thoracic region, rib cage, and diaphragm of severely deconditioned or bed-confined persons or those with spinal cord dysfunction might be beneficial and improve tolerance of routine activity or prescribed therapy. Patients with impaired mobility or who have undergone total joint replacement or other surgical procedure, especially if they have concurrent fluid accumulation in the limbs, might benefit from gentle manipulative techniques to assist in fluid mobilization until the patient’s own mobility could do so. Myofascial release techniques could assist in the recovery from trauma-induced issues. Perhaps persons with traumatic brain injury or children with developmental delay could benefit from craniosacral manipulation. Of course, all of these potential treatments come with the advice of careful patient, provider, and technique selection and to follow the “golden rule” of medicine: “First do no harm.”

Patient Selection

A general physiatric examination should be performed on every new patient. Any underlying impairment must be identified and treated, including fractures, disc herniations with neurologic signs and symptoms, and major sprains, strains, tears, hematomas, and joint injuries. A thorough physical examination, including assessment of gait, posture, muscle tone, muscle stretch reflexes, manual muscle testing, coordination, mental status, and cranial nerves, is strongly recommended prior to initiation of, or referral for, manipulation. Considerations for imaging or electromyography would be no different for the physician using forms of manipulative treatment than for any other type of intervention. There are specific contraindications for thrusting and articulatory manipulation that must be considered, which are discussed below.
Next, the physiatrist contemplating manipulative intervention performs a focused, detailed structural examination in areas suggested by symptoms or by the general examination, with the goal of finding treatable “manipulable lesions.” This evaluation involves careful observation and palpation as well as active and passive motion examination. Subsequent success with manipulation will depend on an accurate diagnosis of lost motion. The findings of palpation and segmental autonomic changes constitute the most important components of structural diagnosis.

The factors generally sought on vertebral or segmental levels are summarized in Table 64-1 (9,70). Passive vertebral motion is evaluated for range, symmetry, quality of motion, and force needed to achieve full range (10,71). Additionally, assessing forward, backward, or side bending and rotation of the vertebra (72) and tenderness elicited by local pressure on the interspinous ligament (70) are useful clues to dysfunctional conditions and loss of joint play. Subcutaneous tissue texture changes (e.g., edema, fibrosis) (72) noted by palpation and skin rolling (10) are signs of dysfunction, and some practitioners look carefully for segmental autonomic changes such as perspiration (10) and hyperemic response to light pressure. The ribs, occiput, and pelvis often need to be included in this part of the examination. Vertebrae identified by positive findings in this type of examination are considered dysfunctional and, if hypomobile, are candidates for manipulation unless contraindications exist.

Hypermobile vertebral segments are themselves not amenable to manipulative treatment (11) but as they may signal the presence of hypomobile segments elsewhere in the spine, should prompt a search for reduced motion elsewhere. If hypomobile segments are found, successful manipulation could conceivably resolve a distant hypermobility. This interconnect- edness of the vertebrae is one of the features that make dysfunctions of the spine difficult yet fascinating to study.

A structural examination of this type may add 5 to 10 minutes to an initial visit and less than 5 minutes to subsequent examinations. The physiatrist choosing to manipulate will of course need to do this evaluation with a relatively high degree of acumen and skill. Physiatrists referring a patient to another physician or a therapist for manipulation can probably stop with identification of a hypomobile or tender segment and leave some of the detailed motion analysis to the practitioner providing the manipulation. However, sufficient examination must be accomplished to identify the vertebral segment needing attention and to establish an endpoint of therapy that the physiatrist can verify.

### Modes of Practice

Availability of manipulative treatment depends on which mode of practice is chosen. Physiatrists who wish to use manipulation but not practice it themselves have two alternatives.

### Referral to a Physician

The referral of the patient for concurrent manipulative care to another practitioner works well in many instances (38), but potential problems may occur, such as oversight of the patient’s care. This impediment may be addressed through a specific referral that states the exact nature and the scope of treatment requested, that encourages discussion with the referring physi- cian, and that makes clear the intent of the physiatrist to resume the remainder of care, if that is the case. In addition, some physiatrists may be unable to assess the manipulator's competence, and some patients may resent a referral to another practitioner.

### Referral to a Licensed Nonphysician

The second option is for the physiatrist to diagnose the problem and write a prescription for manipulation to be performed by a licensed nonphysician, typically a physical therapist (6). This becomes a viable option as more nonphysicians acquire training in manipulation. In the United States, physical therapists traditionally have been limited to providing isometric and articulatory techniques, although counterstrain and myofascial release techniques are becoming more popular. Manipulation can be provided as part of a comprehensive physical therapy program. Even with such a referral pattern, however, the physi- cian should at least acquire basic palpatory diagnostic skills, which then allows one to write a detailed prescription for an anatomic area to be treated and motion to be restored, along with a suggested technique to be used and frequency and length of treatment. The physician can then monitor progress of the treatment objectively and determine the endpoint and possible side effects or failure of the manipulation. Sequential monitoring of each manipulative intervention can be left to the therapist, given the safety of isometric and articulatory techniques. The acquisition of minimal diagnostic palpatory skills may take 1 to 2 weeks. However, the physiatrist probably will use these skills frequently and maintain them easily. The time commitment and economic investment are far less for this option than in the provision of manipulative treatment itself.

Realistically, many physiatrists allow the practitioner to choose areas of treatment and techniques. This is effective in most cases but can lead to loss of physician direction. Caution must be used to avoid unnecessary continuation of treat- ment when pain relief does not correlate with biomechanical improvement. If the patient has little to no perceived benefit following four to six appropriate manipulative treatments, the patient should be reassessed by the physiatrist. If the biomechanical deficit persists, additional regions of the patient’s body should be reevaluated for their possible contributions to the somatic dysfunction. If the structural lesion is improv- ing, based on the evaluation of the prescribing physician, but the patient does not feel the treatments have been helpful,

### TABLE 64.1 Vertebral and Segmental Findings on Palpation

<table>
<thead>
<tr>
<th>Tenderness</th>
<th>Asymmetry of bony structure</th>
<th>Restricted range of motion</th>
<th>Tissue texture abnormalities</th>
</tr>
</thead>
</table>
additional sessions of manipulation should be approached with caution. Consideration of another type of treatment, or a different provider, should be entertained. Most physicians employing manipulation believe that restoration of the vertebral joint’s normal passive and active ranges of motion and resting equilibrium position is the endpoint of treatment (10). In this view, spinal manipulation is analogous to peripheral passive range of motion and end stretching. It is a natural extension of the physical therapy that physiatrists already prescribe for spine problems because it simply addresses individual joints rather than the spine as a whole.

**Acquiring Manipulation Skills**
There are three obstacles to the physiatrist performing manipulation:

1. Acquiring initial skills
2. Maintenance of skills
3. Economics of manipulation

Thrusting types of manipulation, because of the significant forces involved and potential for harm, are best learned on volunteers, not patients, under the close direction of a skilled practitioner. Estimates of the minimum initial learning time required vary from 6 to 8 weeks (7). Even the lower estimates place this type of training out of the reach of many physiatrists. Because of their inherent safety, isometric (muscle energy) and articulatory techniques sufficient for therapy initiation can be acquired by most in 1 to 2 weeks of formal course work. This is so because even an inappropriate isometric technique rarely causes detrimental effects. Thus, the physiatrist may begin working with patients early in training and develop improved techniques with time.

Postdoctoral programs such as those approved by the American Academy of Physical Medicine and Rehabilitation, the American Osteopathic Association, Colleges of Osteopathic Medicine, the American Manual Medicine Association, and other sources can provide a mechanism for the physiatrist to become acquainted with the skills necessary to begin manipulation. There are numerous other courses available that may also be of use to the physiatrist but would have to be investigated individually. There is growing interest among physiatric residents in acquiring this type of training (73), which has led to expanded training opportunities.

The actual minimum frequency of use needed to maintain competence or excellence varies considerably from one person to another. Clearly, however, the potential user should consider maintenance of skills before investing time and money in the acquisition of these skills.

Although any physician can study manipulation techniques, acquiring manipulation skill is demanding, requires practice, and not all physicians will derive satisfaction from providing it. Anecdotally, at least, manipulators who do not enjoy performing manipulation may not do it well.

**Indications for Manipulative Treatment**
Manipulative treatment, if there are no contraindications, is applicable in all musculoskeletal pain problems of the back, pelvis, neck, and limbs in which loss of motion or localized pain on induced motion may be a contributing factor. Some recent injuries (fractures, tears, sprains, strains, disc herniations with neurologic signs, tumors, anomalies, joint disease, inflammation, stenosis) will not respond to manipulation because of the resultant hypermobility, or because local conditions constitute contraindications. Again, all visceral or systemic pathologic conditions, which could manifest through viscerosomatic reflexes must be excluded or at least be under concurrent care.

Although mechanical back and neck pain is the most common indication, other musculoskeletal conditions responding to manipulative treatment have been described. Carpal tunnel syndrome and thoracic outlet syndrome have been shown to improve with myofascial stretch techniques (74–76). Techniques to optimize thoracic cage, clavicular, and scapular motion have benefit in many shoulder girdle problems (10,11). Common sports and repetitive use injuries can be treated with manipulation in conjunction with exercise to restore proper biomechanics (9,77).

**Adverse Effects**
Spinal manipulative treatment has a number of reported side effects (78,79). Perhaps the most common is a transient increase in discomfort for 6 to 72 hours after treatment (70). This should resolve to a level less than that of the pretreatment pain and should be less of a problem with each successive treatment. Minor autonomic effects, such as increased perspiration and early or increased menses, have been reported.

**Duration and Frequency of Treatment**
Appropriately applied treatment with the correct diagnosis generally produces objective improvement in 2 to 4 weeks (16). Lack of improvement in range of motion or function after 2 to 4 weeks would suggest the need to reevaluate the diagnosis or therapeutic plan. Duration of treatment is determined on an individual basis.

The optimal frequency of manipulative treatment is somewhat difficult to discuss. A search of the literature for a study to answer this question produced no results. Similarly, a review of several major textbooks on manipulation also failed to provide guidelines for treatment frequency. From experience, different practitioners of manipulation have their own preferences regarding style and frequency of treatment. The authors of this chapter feel that duration and frequency of treatment should be based on the condition being treated and the type of manipulation being used. There are undocumented, anecdotal reports of the development of hyperlaxity of ligaments resulting from the use of high velocity low amplitude (thrusting) techniques. The exact frequency and duration required to produce this undesirable effect is not clear. Soft-tissue techniques and indirect treatments, such as counterstrain, and other positional approaches, should allow the patient’s body adequate time to respond to each treatment before the same area is treated again, thereby preventing development of proposed reactive hyperlaxity of ligaments that is postulated with
more direct techniques. Another consideration to remember for areas of somatic dysfunction which seem to require frequent treatment to maintain the correction is to expand the physical examination to look for other areas of dysfunction which may be contributing to or causing the lesion in question, and appropriately treat this area first. This will no doubt be the topic for future studies on the use and application of manipulation.

**Contraindications**

Manipulative techniques differ as to their degree of invasiveness. Because of the higher forces involved, thrust is considered the most invasive, whereas articulation, isometrics, and functional approaches are considered less invasive. The more invasive the approach, the more likely it is to be contraindicated. The literature concerning contraindications for thrusting techniques has been reviewed comprehensively by Kleynhans and others (1,7,80,81); contraindications to articulation are discussed by several investigators (16,82–84).

Absolute contraindications for manipulation, especially nonthrusting techniques, are rare, and relative contraindications are few in number. Clearly, manipulation should only be performed for a hypomobile vertebral segment that might respond to manipulation; therefore, accurate diagnosis is important. Inadequate skill of the practitioner is a major contraindication for all types of manipulation.

Contraindications are summarized in Table 64-2.

Patients with concurrent cognitive or affective disorders may be poor candidates for thrusting interventions (84).

**TABLE 64.2** Contraindications for Manipulation

<table>
<thead>
<tr>
<th>General Contraindications for Manipulation Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vertebral malignancy</td>
</tr>
<tr>
<td>Infection or inflammation</td>
</tr>
<tr>
<td>Cauda equina syndrome</td>
</tr>
<tr>
<td>Myelopathy</td>
</tr>
<tr>
<td>Multiple adjacent radiculopathies</td>
</tr>
<tr>
<td>Joint instability (fractures, dislocations)</td>
</tr>
<tr>
<td>Cervical rheumatoid disease</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Contraindications for Thrust Technique</th>
</tr>
</thead>
<tbody>
<tr>
<td>Systemic anticoagulation</td>
</tr>
<tr>
<td>Severe uncontrolled diabetes</td>
</tr>
<tr>
<td>Atherosclerosis</td>
</tr>
<tr>
<td>Severe degenerative joint disease</td>
</tr>
<tr>
<td>Signs and symptoms of vertebrobasilar insufficiency</td>
</tr>
<tr>
<td>Spondyloarthropathies</td>
</tr>
<tr>
<td>Inactive rheumatoid disease</td>
</tr>
<tr>
<td>Congenital ligamentous laxity (Marfan’s, Ehlers-Danlos, Down’s, etc.)</td>
</tr>
<tr>
<td>Aseptic necrosis</td>
</tr>
<tr>
<td>Local aneurysm</td>
</tr>
<tr>
<td>Osteomalacia</td>
</tr>
<tr>
<td>Osteoporosis</td>
</tr>
<tr>
<td>Acute disc herniation</td>
</tr>
</tbody>
</table>

In addition, some investigators consider the absence of any pain-free direction of vertebral motion a contraindication to thrust. Pregnancy with known threat of miscarriage is an absolute contraindication to manipulation. Back pain and intervention in the normal low-risk pregnancy are discussed by Mantle (85). Because the disturbed biomechanics and hormonal ligan-mentous interactions of pregnancy are frequently implicated in back pain in pregnancy, this subject continues to stimulate discussion in the literature (86–90).

Objective radicular signs are a contraindication to articulation or thrust. The question of manipulation for clearly defined root pain without objective neurological findings is controversial, but most authors appear to avoid forceful techniques in favor of positional interventions (2,38,91–93). Isometrics in a nonlateral position appear to be safe, but there is no conclusive literature on this or any other contraindication to isometrics. Vertebral isometric techniques should not be used if the lesion would preclude passive range of motion, gentle stretching, or isometrics in a peripheral joint with the same condition. Catastrophic results do not occur in isometric techniques; some of the absolute contraindications to thrust and articulation, specifically osteoporosis, central cervical intervertebral disc herniation, rheumatoid disease distant from the cervical area, degenerative joint disease, and joint laxity may become relative contraindications, and patients with these problems probably can be cautiously treated with isometrics. Concurrent active myositis may contraindicate isometrics because of the patient’s active muscular contraction; however, for functional and counterstrain techniques, no contraindications have been documented, but it would seem reasonable to assume these may parallel restrictions against massage, discussed later in the chapter.

**Complications**

Clearly, complications may occur as a result of procedures that are done despite contraindications. However, the subject needs elucidation beyond that simple statement, best accomplished by discussing cervical procedures and thoracolumbar-sacral procedures separately.

**Cervical Manipulation Complications**

Several reviews of the literature of complications of cervical spine manipulation have appeared in recent years (12,94–100) including in children. Mild transient reactions to manipulation (78), particularly thrust manipulation, are not uncommon. In cervical manipulation, serious complications involving vascular insult in the vertebralbasilar artery system do occur, although rarely. Almost all are related to thrusting procedures, although this fact can be misleading because the greatest number of manipulations performed involve thrust. Myelopathy as a complication is even more rare (97).

Complicating the review process regarding vascular complications is the lack of available information in the literature specifying number of exposures, vertebral level, localization and direction of thrust, degree of cervical extension, and pertinent patient history. Many of the reported disastrous cases involved
relatively young patients who presumably had no preexisting vertebral vascular insufficiency and who had undergone prior, uncomplicated cervical manipulation. Haldeman et al. (95) found more spontaneous cases of vertebrobasilar artery dissection than cases associated with thrust manipulation.

The following appear to be reasonable deductions from the review literature regarding vascular complications in patients without contraindications to cervical manipulation:

1. The incidence of cervical complications from thrust manipulation is rare, with best estimates at one case per 1 to 1.5 million manipulations (2).
2. The risk to any given patient is unpredictable and random; provocative testing does not appear to be a sensitive or specific indicator of risk and, given the findings that head positioning alone can lead to insufficiency, may constitute a risk in itself (95).
3. The risk of nonthrust cervical manipulation would appear to be very small compared with that of thrust manipulation, but the paucity of documented cases makes even an estimation impossible.
4. It seems reasonable to avoid excessive cervical extension because this position may compromise the vertebral artery system.
5. It is prudent to avoid cervical manipulation in any patient with objective signs and symptoms referable to the vertebrobasilar system, cerebral spondylosis, or spinal instability.

**Thoracic and Lumbosacral Manipulation Complications**

Severe complications following thoracic and lumbosacral manipulation are very rare and appear to have been limited to thrust procedures. The most serious is the development of a cauda equina syndrome (101). Many of the cases in the older literature actually involved manipulation under anesthesia, which is rarely used today. The incidence of such severe complications, even when the potential for unreported cases is included, is estimated to be 1 in many millions; recent reviews (12,102–104) address the issue. There are no documented or anecdotal reports of serious complications arising from articular, functional, isometric, soft-tissue, myofascial release, or counterstrain techniques.

**Prescribing Manipulation**

The referral of a patient for manipulation requires a prescription. This can be, and perhaps is best done, through conversation with the provider, implying the physician has chosen a practitioner with appropriate credentials and training. The manipulative care will usually be included as part of a general ranging, strengthening, and educational program. What will be added to the prescription are the specific vertebral region to be manipulated, the suggested technique to be used, and any specific side effects or medical issues the provider must be aware of. For example, a patient with low back pain, for whom demonstrable pathology has been ruled out, with the palpable finding of hypomobility of T-12 relative to L-1, would have a prescription as follows:

**Patient age:** 

**Precautions:** Avoid thrusting techniques

**Diagnosis:** Unresolved mechanical low back pain without radiographic or neurologic findings

**Somatic dysfunction T-12**

**Rx:** Back stabilization program

**Back education program**

**Muscle energy procedures to T-12 to restore normal T-12 mobility relative to L-1**

**Home program**

**Frequency:** Three times per week

**Duration:** Three weeks or until T-12 motion is normalized

**Options:** May use soft-tissue massage before manipulation, may use superficial hot-cold modalities to supplement above. Please call to discuss other treatment options if indicated.

**Other:** Notify physician if patient complains of increased pain lasting more than 8 hours after last session. Notify physician when mobility in T-12 relative to L-1 is normalized.

Follow-up patient examination, in this case to confirm the restoration of T-12 mobility, occurs as with any other treatment program. If unsuccessful, reformulation of the diagnosis or of therapy would follow. The next step might be escalating the invasiveness of the manipulation to articulatory or thrusting techniques, if no contraindications exist. Conversely, if manipulation is producing discomfort lasting greater than 8 hours, consider switching to the less invasive counterstrain or functional techniques.

The manipulation technique used will be determined by the time course of the problem, the patient's age, general physical condition, presence of any contraindications, and expertise of the practitioner. The former considerations explain the need for a careful history, physical examination, segmental structural examination, and functional diagnosis, whereas the last justifies the choice of a particular practitioner for a problem. The provider should be allowed some input to the technique or approach, with agreement reached before therapy is initiated. A manipulable lesion must be identified for manipulation to be effective for any patient; this is particularly important for the subacute or chronic population that the physiatrist frequently encounters. Patients cannot localize muscular forces to a particular segment or joint in the spine; thus, self-mobilization of the vertebra is not feasible. The manipulative process will therefore always need to be performed by a practitioner. Also, it should be remembered that there is rarely complete eradication of pain in these patients. These factors lead to the possibility of dependence on manipulation, and the physiatrist should be aware of this possibility. Having well-defined biomechanical endpoints (i.e., normal range of motion and tissue compliance) will help objectify the endpoint to manual treatment. If pain relief alone is used as an endpoint, treatment could continue indefinitely. There is no well-established literature on this point, but experience with chronic pain patients suggests that manipulation, with accompanying physical therapy, should be
directed at obtaining an optimal biomechanical condition as quickly as possible, and thereafter used only infrequently to “tune up” the mechanics as needed. There is no evidence that prolonging care is in the patient’s best interest.

Research Relevant to Manipulation
There probably is no other medical modality that has had so many practitioners (50,000 doctors of osteopathy and chiropractic in the United States), consumers (90,000,000 chiropractic manipulations alone are performed annually in the United States), and critics with so little hard data to support or refute it.

The difficulties inherent to the study of manipulative interventions (105) are formidable and include the lack of a well-defined placebo treatment, of well-accepted outcome measures and agreement as to what procedures should be included or excluded, and of objective physical methods of quantifying the preexisting and posttreatment biomechanical state, the parameters of treatment, what procedure to use, and what duration of treatment should be investigated, to name a few. Even more confounding is the lack of basic understanding of the plethora of causes of musculoskeletal pain, the poorly understood relationships between pain generators and outcome, neurophysiologic and biomechanical changes, and the blurred boundaries between acute and chronic pain. The net result is that, despite encouraging trends, hard scientific proof of efficacy remains elusive.

Procedural Fundamentals
The basic mechanics of thrust manipulation have received much attention recently and are well summarized (106). Research in this arena continues to be done (107), but the other methods of manipulation have not been examined to determine their biomechanical and neurophysiologic bases of action, despite major advances in our understanding of the related areas of pain and mechanoreceptors (44), spinal cord physiology (44,108), disc pathology (45), and normal spinal biomechanics (46–50,109).

Two areas of research development are strikingly absent. One involves the lack of any firm association of somatic dysfunction (arthrodial and myofascial restrictions) to pain (36). Although somatic dysfunction is the entity treated by manipulation, the nearly universal reason for treatment is pain relief. However, the relationship between these two basic concepts is ignored even in the hypothesis stage of most concepts. The second area involves the absence of any instrument or device that can objectively measure somatic dysfunction in a patient.

Most physiatrists see patients who present with problems potentially treatable by manipulation and who have undergone manipulation or soon will, in some setting. At the very least, therefore, physiatrists should know the types of manipulation their patients might encounter, along with associated contraindications. Only some physiatrists will actively refer patients, usually those failing to improve with initial therapeutic interventions for manipulation, and few will fully incorporate manipulation into their own practice.

A.T. Still, M.D., the founder of osteopathic medicine taught that, through knowledge and facility with anatomy and physiology, physicians should understand how to apply manipulation principles to the body. For this reason, he never wrote a textbook of techniques, nor specific formulaic instructions, for treatment of a given body part or ailment. Since his time, many types of providers of manual care have published many such books and manuals, each of which applies its own perspective upon a concept and provides variations of one or more similar techniques and terminology. There is ongoing debate among manual medicine practitioners over which group was first and which techniques are the best. It is unlikely that this will be settled in the foreseeable future. It is probably best to rely on the experience and skill of an individual practitioner to provide appropriate manipulative treatments, rather than limit one’s referrals to only one type of provider.

Regardless of the type of manipulative treatment chosen, perhaps one of the most important things to keep in mind is that it is not a “stand alone” treatment. Physiatry is a medical specialty that has pioneered and embraced the “team approach” to patient care; the incorporation of manipulation as another part of the treatment armamentarium can fit well within its paradigm. One may choose pharmaceutical therapy for hypertension, spasticity, or other conditions; however, when it comes to physiatric application of manipulation, a combination of manual and other therapies may be the best approach.

Manipulation in the United States is now well over 100 years old; more research should be generated regarding its use. Ongoing concern about health care cost containment will demand answers to questions of its mechanism and efficacy, as in all areas of medical practice. An objective validation is daunting, given the present level of knowledge; although much of medical practice is not based on objective findings, efforts should continue to maximize current objective data. Until these questions are resolved, the physiatrist must evaluate manipulation based on incomplete information, as in other areas of physiatric practice (Table 64-3).

MASSAGE
Massage is one of the oldest treatment modalities in medicine (110,111). Throughout history, it has been woven into the cultural context of medicine. In this sense, it may be viewed as an essential component of traditional medicine, with the word “traditional” implying an old and very common usage. Massage may be classified into Western and Eastern variants. In the West, mentioned as early as the time of Hippocrates (111), its practice and popularity have waxed and waned over the centuries. The interested reader can find detailed accounts of the history in various sources (111–117). In recent years, its popularity has increased by a resurgence of interest in alternative therapies (114). Some claims for massage by its proponents may sound implausible, but there can be little doubt that it is considered efficacious for a wide variety of conditions
TABLE 64.3 Key Points for Manipulative Treatment

| The goal of manipulative treatment is restoration of lost joint motion and improved postural balance. Thrust, articulatory, counterstrain, functional, muscle energy, myofascial release, soft tissue, and craniosacral are common techniques used in manipulative treatment. Although the exact mechanism of action is unknown, manipulative treatment appears to alter the mechanical behavior of static musculoskeletal structures and the neural control of muscle. Indications for manipulative treatment include acute and chronic spinal pain of mechanical origin and overuse injuries of the extremities. Benefits are generally observable after 2 to 4 weeks of appropriate treatment. Complications are exceedingly rare, although serious vascular compromise has been reported with thrust technique in the cervical spine. Although methodologically sound studies are lacking, most research and consensus suggest a modest benefit of manipulative treatment in mechanical spinal pain. |

(115–118). There is a striking need for additional research regarding its mechanisms and efficacy.

Definition and Relation to Other Mechanical Therapies

Massage is the therapeutic manipulation of the soft tissues of the body, with a goal of normalization of those tissues. It is also defined as “hand motion practiced on the surface of the living body with a therapeutic goal,” (118) or, “a group of procedures which are usually done with the hands, such as friction, kneading, rolling, and percussion of the external tissues of the body in a variety of ways, either with a curative, palliative, or hygienic object in view” (119).

Manipulation and massage both involve similar hand positions and movements, and some techniques overlap significantly. Soft-tissue manipulation, performed to loosen and relax tissues before vertebral or more direct manipulation, is partially composed of massage techniques. Myofascial release (120) uses stretching forces similar to those used in massage. Myofascial release and craniosacral manipulation, discussed in the manipulation section of this chapter, are often available from massage therapists, and massage can easily be a preliminary treatment to manipulation. Regardless of the similarities, massage clearly targets the health of the soft tissues, whereas manipulation often targets vertebral or segmental motion; soft-tissue health is optimized secondarily with manipulation.

The difference between massage and traction is easier to delineate. Massage consists mostly of hand movements, some of which may be traction based. Traction as a therapy usually involves machinery but can also be applied manually; the latter would not use other massage movements. In addition, traction tends to effect changes in the spinal column, with soft tissues only secondarily modified. Like traction, massage effects tend to be nonspecific.

Massage can have mechanical, reflexive, neurologic, and psychological effects and can be used to reduce pain or adhesions, promote sedation, mobilize fluids, increase muscular relaxation, and cause vasodilatation (121,122).

Basic Western Massage Procedures

There is perhaps more confusion in the terminology of massage than in many other areas of manual medicine. Much of the basic terminology for Western massage was introduced by Pare (123) in the first half of the 20th century, but there is a tendency for textbooks and practitioners to use idiosyncratic variants. The essence of massage is the use of the hands to apply mechanical forces to the skeletal muscles and skin, although the intent may be to influence either more superficial or deeper tissues. The types of massage prevalent in the Western hemisphere are derived from the Swedish system. They are categorized by whether (a) the focus of pressure is moved by the hands gliding over the skin; (b) soft tissue is compressed between the hands or the fingers; (c) the skin and muscle are impacted with repetitive compressive blows by the hands; or (d) shearing stresses are created at tissue interfaces below the skin. The three former techniques retain their classic French names, effleurage, pétrissage, and tapotement, whereas the fourth is called deep friction massage.

Effleurage or Stroking Massage

The operator’s hands glide across the skin overlying the skeletal muscle being treated (112,113,115–117,120,124). Oil or powder is often incorporated to reduce the friction between the hands and the skin. Any part of the hand can be used, but many applications require the palm to be in contact with the skin. In all stroking techniques, hand-to-skin contact should be maintained throughout the stroke. If the compressive force on the muscle is kept relatively light, this is called superficial stroke massage; if relatively heavy, it is deep stroke massage. Presumably, a light stroke energizes cutaneous receptors and acts by some neuroreflexive or vascular reflexive mechanism, whereas a deep stroke is more capable of mechanically mobilizing fluids in the tissues beneath it, including the muscle, in the direction of the stroke. Therefore, deep stroking should be in the direction of venous or lymphatic flow, whereas light stroking can be in any direction.

Effleurage tends to be a technique common to all massage therapy because it is used to gain initial relaxation and patient confidence, to diagnose regions of spasm and tightness, and most importantly, to get the operator’s hands from one problem area to another, where other techniques may be employed. Effleurage has as its main mechanical effect the application of sequential pressure over contiguous soft tissues, with the result that fluids will be displaced ahead of the hands as the tissue is compressed.

Pétrissage or Kneading Massage

Kneading massage involves compression of the underlying skin and muscle either between the fingers and thumb of one hand or between the two hands of the practitioner. The
entrapped tissue is squeezed gently as the hands move in a circular motion perpendicular to the direction of compression (112,113,115,116). Some investigators call this “kneading” the tissue, whereas others separate kneading from pétrissage, depending on the direction of the compression and force of the motion. The pressure is such that the skin can move over the underlying tissue, but practitioners disagree on whether the operator’s hand should slip relative to the skin. Mobilization of fluid would be a result with this type of treatment, but pétrissage has as its main mechanical effect the compression and subsequent release of soft tissues, reactive blood flow, and neuromuscular response to that flow.

Tapotement or Percussion Massage

The operator strikes the soft tissue with repetitive blows using both hands rhythmically, gently, and rapidly (112,113,115,116). There is no rigid frequency requirement, but typically these motions occur about three times per second (117). There are numerous variants defined by the part of the hands that make the impact; using the back of the medial fingers is called hacking; the volar surface of all the fingers is called slapping; the hypothenar eminence is called beating; the pads or the tips of the fingers is called tapping. If the thumb and index finger do a light pinch on contact, it is called pinacement. A form more familiar to physicians occurs when the cupped palm of the hand is used to make the impact. This tends to make a clapping noise on each impact. The therapeutic effect is hypothesized to result from a compression of trapped air that occurs on impact. Because the frequency of sound is low, the waves can penetrate to deep structures, but whether this is the therapeutic effect is not clear. If done over the upper back, the lung tissue may absorb the waves; a variant of this technique, called clappatage, is used by respiratory therapists to loosen bronchial secretions.

The effect of tapotement is thought to be stimulatory. It tends to be used more often on healthy people by nonmedical practitioners in a nontherapeutic milieu. Nonetheless, it also is part of the armamentarium of medically trained practitioners.

Friction or Deep Friction Massage

Pressure of varying intensities is applied with the ball of the thumb or fingers to the skin and muscle of the patient. The digit is moved with constant pressure in small circular motions for several cycles before pressure is released and the application performed in a new location (112,113,115,116,125). There is no gliding, as in effleurage, and no compression between the fingers, as in pétrissage. The main mechanical effect of friction massage is probably the application of shear stresses to underlying tissue, particularly at the interface between two types of tissue (e.g., dermis-fascia, fascia-muscle, muscle-bone, scar tissue–bone). It is probably impossible to affect abnormalities within one tissue (e.g., muscle). The effect of the pressure is to keep superficial tissues under the thumb from shearing so that the shear and force are directed to a deeper interface. The higher the applied pressure, the thicker the plane of non-gliding tissue and the deeper the level that shear takes effect.

Various techniques have the motion across fibers, parallel to fibers, or in random fashion. Friction massage is perhaps the primary method of preventing or treating adhesion of scar tissue to bone and deeper structures. For this reason, physiatrists involved in amputee rehabilitation may encounter this application more often than other techniques.

All of these techniques have variants, and each has its proponents. A variant of tapotement, called vibration, differs in that the hands do not lift off the skin at the end of each cycle. Variants of pétrissage depend on the forces applied to the tissue trapped between the fingers or hands. The technique of picking up requires the tissue to be first compressed against deep structures, pinched, then released for advancement of the hands to a new area for treatment. Wringing requires one hand or finger of the practitioner to push the trapped tissue while the other hand pulls, perhaps shearing even deep tissue planes between them. Rolling, which is often done just to the skin in diagnosis of manipulable lesions, allows the tissue to roll between the fingers and thumb, often combined with gliding of the digits over the skin. In shaking, entrapped tissue is shaken by the operator’s hand as it glides down the length of tissue.

Each of these variants arises by changing the dominant force used, then adding a new sequence of shear, compression, or glide that basic techniques provide. For a given patient, the correct combination of these forces may be more important than the sequence of their application.

Decongestive Lymphatic Therapy

Decongestive lymphatic therapy, also known as complex decompessive therapy, includes manual lymphatic drainage (MLD) in addition to use of specifically prescribed compressive bandages and garments, as needed, along with exercises and skin care (126). MLD consists of gentle massage in the treatment of lymphedema. Proximal areas are treated first to prevent a damming effect. The treatment is designed to stretch the lymphatic collectors and to stimulate the drainage system. The massage is followed by the application of compression bandages and is incorporated into a comprehensive and individualized self-care program (127,128).

Eastern Massage

Eastern massage systems have evolved over many centuries and are integrated within the culture of the country where they are practiced. The systems for evaluation, diagnosis, and treatment for the most part are not based on known Western neurophysiology. It is evident that Western massage is incorporating some of the more valuable Eastern concepts (129). Many of these techniques resemble variants of classical Western massage, although they are often used for visceral complaints. There are many more variations of technique, as many as 24 in one Chinese system (130).

Acupressure

Acupressure applies massage forces, largely digital pressure, to the same points treated with the better known acupuncture
needles, and for the same reasons (112,113). Energy, or Chi, is said to circulate in the body along 12 meridians or pathways unrelated to the neuromuscular patterns known to Western medicine. Imbalances of energy found along the meridians are thought to cause disease and can be rectified by localized finger pressure accompanied by circular movements similar to deep friction. The pressure is generally increased until it is relatively heavy and then held constant. The similar Japanese form is called Shiatsu and the Indian form is Marma therapy (129,131).

**Reflexology and Auriculotherapy**
The meridian concept is also a part of the reflexology and auriculotherapy systems (112,113). The meridians are thought to have whole-body representation like a homunculus on the extremities, particularly the feet, hands (reflexology), and the ear (auriculotherapy). All parts of the body and its organs are mapped to points on the foot or ear, and massage of a point is thought to produce change in the organ or structure mapped to that point (131).

**Massage Variables**
There are variables of massage that need to be controlled by the operator (112,113). The physiatrist may wish to specify these if ordering massage, although often they are left to the provider. The environment should provide comfort and relaxation to the patient, with proper adjustment to the temperature, lighting, sounds, draping of the patient, and so forth. Media such as oil or powders may be used to reduce friction. Rhythm rate and pressure are applied according to the effect desired, that is, a rapid rate for stimulation versus a slower rate for relaxation. The amount of pressure is also varied. In lymphedema, it is thought that low pressure assists in the flow, whereas higher pressures may cause damage to the lymph vessels. The direction of stroking, as well as the sequencing of the regions, is important in aiding the flow of fluids. Peripheral to central stroking moves fluids toward the core, but central to peripheral sequencing prevents a proximal damming effect. A cross-fiber stroke helps reduce adhesions in the deep friction techniques.

The areas to be treated, as well as the duration and frequency of treatment, should be specified by the referring practitioner. Treatment goals are also helpful. The patient may wish to continue the treatment after goals have been attained. This could be helpful when the patient understands the transition from a treatment program to a wellness maintenance program.

**Physiologic Effects**

**Mechanical Effects**
Vascular changes are recognized as a clear mechanical effect of massage. Mechanical pressure on soft tissue displaces any fluids that are not chemically bound by the tissue or physically bound by compartmentalization. The fluid can move in a low-resistance direction under the static hand force, but a moving locus of pressure will create a pressure gradient. Assuming no significant resistance, pressure is lower ahead of (usually proximal to) the advancing hand. Once the mobilized fluid leaves the cells or interstitium, it can enter the low-pressure venous or lymphatic systems, with varicose mechanisms that help prevent fluid return. Areas are typically worked distal to proximal, centripetally, to keep the fluid movement toward the heart, which, along with the kidneys, is required to handle an additional load. The amount of fluid mobilized in any one treatment is probably quite small, and major effects to the heart have not been noted. Nonetheless, the physiatrist needs to be aware of this physiologic effect in the patient with cardiovascular or renal compromise. Proponents of the “decongestive lymphatic therapy” approach to lymphedema, espousing gently applied massage, tend to massage the proximal lymph area first, and then move distally. The belief is that proximal blockage areas in the lymph channels must be opened first, perhaps reflexively, to allow subsequent distal mobilization of the fluid and protein (122,123,132–134). Studies have shown a 59% to 69% reduction in limb volume with sustainable long-term responses (135,136) and others have not (137,138).

Kneading and stroking massage both decrease edema, and compression will convert nonpitting to pitting edema (139). Friction massage is probably not effective in this regard. In addition to that strictly mechanical effect, it is thought that histamine is released by the massage, causing superficial vasodilation to assist in the washing out of metabolic waste products (121,122). Venous return is thereby increased, and in healthy patients, stroke volume will increase (140). There is contradictory evidence that massage to one limb will increase the blood flow contralaterally (141,142). If this is true, the mechanism is uncertain. Effleurage may be more effective than either diathermy or ultrasound in increasing blood flow to congested or damaged tissue (143). Clearly, these effects on the mobilization of fluids are more important in the flaccid or mechanically inactivated limb because the normal compression supplied by skeletal muscle contraction is absent in those cases.

Recent findings suggest that massage may decrease blood viscosity and hematocrit and increase circulating fibrinolytic compounds (144,145). Although these are preliminary data, it would help explain the apparent success of massage in decreasing instances of deep vein thrombosis (146). Massage nonetheless is still contraindicated in the presence of existing thrombosis. Other blood compounds that may be transiently increased by massage include myoglobin, glutamic oxaloacetic transaminase, creatine kinase, and lactate dehydrogenase (147,148). One may suspect that these transient increases represent local muscle cell leakage from applied pressure. Lactate does decrease in massaged muscle cells (149). An increase in natural killer cell numbers has been documented in many massage studies (150). Massage is said to decrease muscle spasm and increase force of contraction and endurance of skeletal muscle (151,152). One can rationalize the decrease of spasm and the increase of endurance as occurring secondary to washout of metabolic waste products by the mobilization of fluids and increased blood flow. The enhanced contractural force, however, is based on remote publications and needs new investigation, validation, and elucidation, particularly because the mechanism is
not obvious. A decrease in muscle soreness is probably from the washout of metabolites, and the cycle may be broken permanently if the spasm was responsible for the original buildup of the metabolites and immobility of the muscle. A number of studies support the benefit of sports-specific massage, with other studies showing contradictory results (153).

The shearing forces of deep friction massage have not been studied adequately, but it is accepted that they can break down or prevent adhesions and increase fascial mobility (116,117,125,154,155). Limited local tissue damage, with hyperemia and mild inflammation, may occur when adhesions are reduced.

A special case of the mechanical effects of massage on the lymphatic system is the “lymphatic pump” often used as part of manipulation techniques for patients with respiratory compromise (156). This is a type of pêtrissage done to the chest and rib cage that is hypothesized to draw lymph into the patient’s thoracic duct and hence to the venous circulation because of the alternating increase and release of pressure on the chest cavity.

Similar to manipulation, it has been postulated that endogenous endorphins are released by massage (121,157). Here, too, there is little supportive evidence, and further studies are needed to establish this effect either way (158).

**Reflexive Changes**

Massage can stimulate cutaneous receptors and possibly the spindle receptors in superficial skeletal muscles (122). These produce impulses that reach the spinal cord, which, once there, conceivably can produce myriad effects (121,122,159). One such effect would be modulation of the facilitated segment, an area of impairment that develops a lower threshold for irritability when other areas are stimulated (160). Somatovisceral reflex effects on the viscera are possible in this model, although evidence for this is lacking. Our knowledge of somatovisceral and viscerosomatic reflexes is rudimentary enough to consider almost any effect, but allowing for such a system, massage could have distant visceral effects. Others might include input to the “gate” of the dorsal column of the spinal cord, thereby reducing pain like a mechanical TENS unit (159). The increased afferent input may influence the activity of the cord itself. Some of the vascular effects noted previously could be secondary to reflexive vasodilation, although one might justifiably question why vasodilation does not also occur. It does seem safe to acknowledge that reflexive effects occur in massage, with further details awaiting scientific validation. One method of massage that depends heavily on reflex mechanisms, and that may be encountered by the physiatrist, is the German *Bindegewebsmassage*, or connective tissue massage (112,161,162). This is based on the concept that autonomic vascular, rather than somatic, nerves are the channels for reflex systems that give rise to skin symptoms or palpatory changes, not necessarily along accepted dermatomes. The skin massage here is done very lightly, presumably in areas related to the viscera, to influence the connective tissue, trigger reflex mechanisms, and ultimately modify the target viscera. Studies of its effects are contradictory (163–166). Some of the Eastern approaches to massage noted earlier (reflexology, auriculotherapy, and Marma therapy) are also dependent on reflexive mechanisms, most of which have little or no basis in conventional Western neurophysiology and neuroanatomy.

These types of hypothetical postulates, which are difficult to prove and even more difficult to disprove, form the basis of massage and manipulation and of practitioners’ efforts to effect visceral or reflex changes with these methods. There are few, if any, complications of massage that cannot be explained on purely mechanical or well-established physiologic grounds. A recent discussion of the ability of somatic problems to mimic visceral pathology may be the first step in understanding some of these approaches (167).

**Psychological Effects**

Most people experiencing any form of massage will attest to the feeling of relaxation and well-being that it produces. Whether this is a placebo effect or the result of some hitherto undiscovered reflex is unknown, but this has stimulated interest in the use of massage for acute and chronic pain conditions as well as in anxiety states (111,150,163,167–169). Some practitioners incorporate a variety of other psychological techniques, such as guided imagery, into the massage treatment. Another massage approach which incorporates psychosomatic integration is Rolffing (after the originator) or structural integration (170). Rolffing incorporates deep friction massage aimed at realigning and balancing the patient’s body relative to the gravitational field, with predetermined sequences of treatments which may be unrelated to individual symptomatology. The rebalancing is thought to optimize both the psychological and physical states of the patient.

**Therapeutic Goals of and Indications for Massage Therapy**

Massage is used either as a therapeutic intervention alone or more commonly as an adjunct to other modalities. It is indicated by a need for one of its physiologic, reflexive, or psychological effects, as noted previously. Thus, it might be ordered to mobilize interstitial fluids, to reduce or modify edema (171–178), to increase local blood flow, to decrease muscle soreness or stiffness, to moderate pain, to prevent or eliminate adhesions, or to facilitate relaxation.

In a physiatric practice, these indications may be considered in patients with sprains, strains, and fractures; mechanical back pain; contractures; myofascial pain syndromes with accompanying tightness, myoedema, and spasms; flaccid or immobilized limbs; amputations; and sympathetic dystrophy (179). In these patients, massage may be used to alter the pathophysiology of the primary condition (e.g., contracture) or to prevent or modify the negative ramifications of the primary treatment modality (i.e., nonobstructive limb edema from immobilization). The physiatrist will most frequently be involved with the chronic sequelae of these impairments, and massage is, perhaps, more useful in chronic rather than acute conditions. In almost all cases, however, massage is secondary to and less important...
than the primary treatment therapies (i.e., immobilization, ranging, strengthening) and must be used in such a way that it acts synergistically with and not antagonistically to them. The sedative effect may be of complementary use in any of a large number of painful or anxiety-stimulating conditions.

**Contraindications**
Massage is contraindicated when it would cause worsening of a condition, unwanted destruction of tissue, or spread of the condition (112). Malignancy, thrombi, atherosclerotic plaques, and infected tissue could be spread by massage, and their presence is thought to be an absolute contraindication to massage. Relative contraindications include the treatment of scar tissue that is not fully healed; patients who are anticoagulated, either therapeutically or by disease; calcified soft tissues; skin grafts; inflamed tissue; atrophic skin; and tissue that is susceptible to further edema if circulation is increased by the massage technique.

Although not rigidly a contraindication, chronic pain situations must be approached cautiously by the physiatrist if massage is going to be used. Because it has a passive, hands-on nature and potentially strong psychological effects, the risk for dependency of the patient is always present. Patient satisfaction, of course, may increase with this type of approach, but objective improvement in the impairment may not occur. To use massage for chronic pain, endpoints must be established before institution of the treatment and treatment must be terminated when they are reached.

**Practical Problems in Physiatric Application of Massage**
Four practical issues may confront the physiatrist. First, patients who have already undergone massage require some action on the part of the physician. Many patients will have received massage directed at relaxation, sometimes ordered by physicians as a therapy for stress relief, but more often entered into by the patient by means of self-referral. Although this type of therapy is seldom directed primarily at musculoskeletal problems, the practitioners often have varying degrees of experience in athletic training or other such vocations and thus may address musculoskeletal issues in the course of the massage. An accurate history of the patient’s experience is relevant, and the physiatrist may wish to consider including the massage therapist in the overall care plan. At the very least, intervention when contraindications are present may become the responsibility of the physician.

The second practical issue facing the physiatrist committed to the use of the massage is the choice of a therapist. Competent practitioners in a therapeutic milieu may be physical therapists, occupational therapists, athletic trainers, nurses, or respiratory therapists. Massage therapists trained in therapeutic techniques may also be found in the clinical setting. When the massage is to be used as an adjunct, the choice is usually dictated by the concomitant therapy. Thus, the type of provider chosen often may depend on skills other than massage. Most physical therapists and many occupational therapists receive massage training during their professional education. Massage involves as much art as science, though, and different providers might obtain markedly different results with what appear to be identical techniques. Thus, trial and error and local reputation and credentials are probably the most useful criteria involved in the choice of a specific massage practitioner.

A third issue is the cost and insurance coverage of massage treatment. Massage alone is sometimes reimbursed but generally needs to be bundled into the overall therapy plan to be viable for most patients. This usually acts to the patient’s benefit for several reasons. Accompanying therapies tend to prevent dependency on massage, particularly in chronic pain conditions. Second, the other therapies involved usually have well-established functional endpoints that can be used to justify the overall program and validate its composite efficacy, an issue of some importance with massage alone. In those few cases in which improved function is not the main desired outcome (e.g., the cosmetic result of lymphedema reduction), the physician may have more difficulty establishing the usefulness of massage. Finally, a massage prescription may need to be generated, particularly if a long-term program is anticipated; such a prescription should include the relevant diagnoses, precautions, areas to be treated, verifiable goals, and specific massage variables discussed earlier. The physician cannot determine some of these variables a priori because tissue feedback often guides the provider, but the physician can delineate treatment parameters.

**Efficacy**
Long-term efficacy of massage has not been validated in most of the previously noted conditions. Most applications of interest to the physiatrist involve a limited course of therapy directed at permanent improvement of a condition. The situation is analogous to the use of an antibiotic to cure an infection as opposed to an antihypertensive medication designed to be used indefinitely. Long-term effects are therefore of paramount interest, but have not been investigated. One issue is that the underlying physiology of the tissues is not well understood so that it becomes difficult to understand how massage would affect that physiology. At this time, the therapeutic use of massage must be undertaken with the understanding that its value lies either in demonstrable short-term improvements (such as decreased edema, soreness, and sensitivity) or in longer-term verifiable outcomes of the accompanying therapies that it facilitates.

Given the interest in and use of massage in the face of increasing scrutiny of health care expenditures, it is important that the long-term efficacy of this modality be established scientifically, including accurate outcome measures. In addition, entry criteria to identify the subset of patients most likely to benefit from massage would be useful. Massage continues to provide a valuable modality for physical medicine practitioners since it is user friendly (albeit potentially time intensive and expensive), of low risk, and does not rely on technology (180,181). As a result, it is available worldwide and has the potential to continue to provide valuable future health benefits.
**TRACTION**

Traction is the act of drawing or pulling, or a pulling force. In medicine, forces are applied to the body generally to stretch a given part or separate two or more parts. Traction continues to be effectively used in the treatment of fractures of the extremities and the spine. In physiatry, traction is usually limited to the cervical or lumbar spine, with the hope of relieving pain in, or originating from, those areas, and this section addresses only spine traction.

Since the days of Hippocrates, attempts at correction of scoliosis have frequently involved traction in one form or another. Treatment of long bone fractures led to the development of many traction techniques in the 1800s (180). More recent attempts at spinal traction have included applying the pulling force manually (181), with free weights and a pulley (182), with the patient supplying the pull with his hands or feet (183), with motorized equipment (184), by inversion and gravity supplying the pull (185), and using partial body weight support in conjunction with treadmill walking (186).

**Physiologic Effects**

**Cervical Spine**

The most reproducible result of traction to the cervical spine is elongation. Cyriax reported that applying a force of 136 kg (300 lb) manually results in a 1 cm increase in cumulative interspace distance (181). A study to determine optimal weight for cervical traction concluded that the minimum weight to accomplish any vertebral separation is 11.6 kg (25 lb) (187). Greater forces (188), and in some cases much greater forces (189), have been used, but most studies have concluded that elongation of 2 to 20 mm of the cervical spine is achievable with 11.3 kg (25 lb) or more of tractive force. In another investigation, the anterior and posterior intervertebral spaces (IVS) and the facet joint space at neck positions of 30 degrees of flexion, neutral, and 15 degrees of extension were quantified under 13.6 kg (30 lb) of intermittent traction (8 seconds of traction followed by 6 seconds of rest) after a total traction time of 20 minutes. The anterior IVS was increased most in 30 degrees of flexion, ranging from 18% to 21%, with the largest IVS at C6-7. The effect of traction on the facet joint was, in all three positions, not statistically significant. Many subjects in the study noted that traction in neck extension was painful, and the posterior IVS actually decreased in some instances. The investigators, citing these reasons, in addition to the increased risk for complications from vertebral basilar insufficiency and the possibility of spinal instability, did not recommend the use of traction in extension (190). The vertebral separation accomplished after traction has been reported as temporary, with return to normal height the morning after traction (188). The effect of traction on paraspinal muscles is not well known. Cyriax postulated that traction leads to fatigue and then relaxation (181). However, another study found greater muscle relaxation in the cervical traction group with 0.9 kg (2-lb) placebo traction than in a group using 2.7 to 5.4 kg (6 to 12 lb) of traction (191,192).

**Lumbar Spine**

After friction is overcome with either adequate force of pull (193) or a split table (194), the major physiologic effect of traction of the lumbar spine is elongation. Investigators who report widening of the lumbar interspaces used between 31.7 and 136 kg (70 and 300 lb) of pull to obtain that result. The widening, when apparent, averages up to slightly more than 3 mm at one intervertebral foramen (195–198). The length of time that this separation persists remains indeterminate, with one study unable to document residual distraction 30 minutes after treatment (197) and another demonstrating residual separation at least 10 minutes after treatment, which was the endpoint of the observations (199).

Data have been generated on the dimensional and pressure changes caused by traction in lumbar discs, both in the normal spine and after herniation. A decrease in the intradiscal pressure of patients treated with 22.7 to 45.4 kg (50 to 100 lb) of traction on a split table has been documented (194), but there is evidence that some such applications actually cause an increase in intradiscal pressure that, in principle, would exacerbate nerve compression from a herniated component of the disc (200). Discograms performed during both compression and distractive traction showed large variations (>10%) of the radiologic disc area in patients with “disc syndrome,” but failed to demonstrate area changes in normal subjects (201). Reduction of prolapsed discs during lumbar traction has been noted by epidural contrast (202) and an unblended computed tomographic scanning technique at short-term follow-up (203). The latter study was marred by lack of both blinding and long-term follow-up. Thus, the evidence is inconclusive, with considerable information favoring at least temporary reduction of the herniated component of an abnormal lumbar disc under the influence of traction.

The issue of whether the disc changes size does not directly address pressure relief on the nerve root or reversal of abnormal physiology in radiculopathy. This much more clinically relevant question has received surprisingly little attention in the literature to date. However, a study by Knutsson et al. demonstrated an improvement in lower extremity voluntary muscle strength and normalization of bilateral somatosensory evoked potentials and skin temperature (by thermography) after autotraction (see below) in patients with lumbar and sacral radiculopathies (204). Unfortunately, the investigators were not blinded, and controls and long-term follow-up evaluations were not included. Nonetheless, this research breaks ground in the area of inquiry that is of most interest to physiatry. It should be noted that there is a tendency for some physicians to use lumbar traction as a means of enforcing bed rest (205,206). The benefits of strict bed rest for low back pain are questionable, and its physiologic effects are discussed in Chapters 55 and 57.

**Evidence-Based Outcome**

There are few scientifically rigorous studies in the literature that allow the effect of traction to be distinguished from the natural history of the pathology (usually radiculopathy) being treated.
Deyo has suggested criteria that would allow the true effects of traction to be delineated in the face of confounding variables. Studies meeting these criteria should use (a) randomized, controlled trials; (b) blindered outcomes assessment; (c) equivalent cointerventions; (d) monitored compliance; (e) minimal attrition; (f) minimal contamination; (g) adequate statistical power; (h) adequate description of study design, patients, and interventions; and (i) relevant, functional outcomes (207). There is probably only one group of studies evaluating traction in nonspecific low back pain that comes close to this rigorous level of evidence (208–210). No other traction outcome study to date has satisfied all these criteria. Despite this, there are randomized, controlled trials that meet many of these and give clinicians some insight into the efficacy of traction.

**Lumbar Traction**

In 1994, the AHCPR reviewed the literature on traction and made recommendations as part of the clinical practice guideline on acute low back problems in adults. Thirty-one articles were examined, of which six randomized, controlled trials met criteria for their review. Additionally, the panel did a meta-analysis of traction. The AHCPR conclusion was that “spinal traction is not recommended in the treatment of acute low back problems” (211). This is a strong, negative recommendation that does not appear to be refuted by other randomized, controlled trials using good methodology. Recent evidence reviews support the 1994 conclusions (212–215).

The highest quality trial is published in a series of three papers which report the methodology and results of a high-quality, randomized, controlled trial of 151 patients with nonspecific low back pain between 6 weeks’ and 6 months’ duration. A sham treatment of a brace that tightened in the back was used. Wide ranges of outcome measures, including work attendance at 4 weeks, 12 weeks, and 6 months, were used. The investigators could not show any statistically significant benefit from traction (216–218).

The studies that claim improvement after traction report modest, short-term improvements with limited or no improvement in functional outcome and have study design flaws. Some examples of these illustrate the lack of convincing evidence for traction efficacy. Eighty-two patients with acute sciatica and positive straight-leg-raising signs were assigned randomly to either corset and bed rest (not timed controlled placebo) or autotraction (219). Evaluation was performed 1 and 3 weeks after treatment by a blinded observer and after 3 months by questionnaire. Traction subjects were treated either as outpatients, who were usually driven home by ambulance, or as inpatients. Placebo controls did not have the time-intensive ambulance rides or overnight hospital stays. Although no outcome differences between those treated with traction and those treated as controls were apparent after 3 months of follow-up, the traction-treated group reported modest pain relief without documented functional improvement 1 and 3 weeks after treatment.

Pain relief has been reported in patients who were treated with traction for 6 days for a chief complaint of back pain and sciatica (220). This improvement was limited to a subgroup of women younger than 45 years of age, follow-up was only 2 weeks, and functional outcomes were not measured. Another investigation showed that patients with sciatica treated with traction and exercise had more favorable pain improvement than a cohort treated with hot packs, massage, and spinal mobilizing and strengthening exercises. The actual control group in this study received a placebo therapy of hot packs and rest and had similar improvement to the traction group. Nonetheless, these similar results were interpreted as a “significant priority” for those treated with pelvic traction (221).

Other controlled trials evaluating primarily patients with sciatica do not report clear benefit for traction (222–226). Of course the “absence of proof is not the proof of absence.” It is entirely possible that a subgroup of patients (e.g., lateral herniated discs with radiculopathy) can benefit from a particular type of traction for either short-term or long-term improvement in functional outcome. At this stage, the literature does not clearly delineate these subgroups and, overall, suggests no benefit from traction.

**Cervical Traction**

Few randomized, controlled trials address patient outcome after cervical traction. Methodology problems are similar to those noted previously regarding lumbar traction. A review of the randomized, controlled trials available will give clinicians an idea of the potential efficacy and limitations of cervical traction.

Perhaps the most ambitious trial was conducted by the British Association of Physical Medicine and published in 1966. A total of 466 neck and arm pain patients were randomly assigned into one of five groups: (a) traction; (b) placebo traction (positioning); (c) instruction in posture and cervical collar; (d) placebo as untuned short-wave diathermy; and (e) placebo tablets. A blinded assessment was performed at 2 and 4 weeks, with letter follow-up at 6 months. Outcome measures included range of motion, pain, and function. There was no statistical difference between the groups and no noticeable trends in subgroups. The authors suggest that none of these treatments alters the natural history of the problem. The major potential criticism of this study is that the traction treatment was not completely standardized. The investigators reported that the traction was “given every chance to succeed,” and thus the therapist had latitude to customize the traction (227).

Three randomized trials have reported improvements in patients treated with cervical traction. Goldie and Lundquist reported the results of a randomized, controlled trial in 73 neck and arm pain patients who were placed into one of three groups: (a) traction; (b) isometric exercises; and (c) instruction only (no time-controlled placebo) (228). There was slight improvement in both of the treatment groups over no treatment, with patient report, blinded physician report, and ranges of motion used as outcome measures. Of note, there was no difference in sick leave taken, although there was a trend for less time off work in the instruction only group. Zylbergold and Piper reported on a controlled trial of 100 patients randomized into
Parameters

As with many other areas of physiatric therapeutics, little has been proved about the most efficacious parameters for traction. Preferences of specific parameter settings (e.g., sitting vs. supine, etc.)
continuous vs. intermittent) depend more on the practitioner’s empiric observation than on objective data. Individual physicians should, within limits dictated by common sense and the experiences of others, develop personal guidelines regarding the amount of weight, duration of traction, and other parameters. Above all, the patient should be comfortable during traction therapy. Monitoring by the physical therapist is essential to ensure that the traction applied is not ineffective or aggravating a painful condition.

Positioning

Cervical

In cervical traction, the choice of sitting versus supine positioning should be based on patient comfort and ability to relax. Crue was the first to suggest the importance of neck position in cervical traction (240,241). This may be related to widening of the intervertebral foramina or IVS. The maximal effect of distraction seems to occur between 20 and 30 degrees of flexion with no accompanying rotation or side bending (190,242). Nearly all studies report difficulties with the cervical spine in extension; thus, neck extension during traction should be avoided (190).

Lumbar

Lumbar traction can be accomplished with upright body suspension, but chest discomfort from the harness is often a limiting factor (197); thus, the supine position is most commonly chosen. Hip flexion (15 to 70 degrees) is routinely incorporated to cause relative lumbar spine flexion, theoretically leading to optimal vertebral separation. Reliable studies comparing different positions of lumbar traction are lacking.

Continuous Versus Intermittent Traction

Cervical

A larger improvement in range of motion with less accompanying pain was noted in patients subjected to intermittent traction of 11.4 kg (25 lb) peak (10 seconds on, 10 seconds off, total duration of 15 minutes) than in cohorts subjected either to 15 minutes of manual traction or 15 minutes of static traction of 11.4 kg (25 lb). All three groups were treated with the cervical spine flexed at 25 degrees (229). Data show that a constant cervical distraction force of 13.6 kg (30 lb) effects maximum vertebral separation in 7 seconds or less and that no more is gained by applications of 30 or 60 seconds (199). If so, then strictly skeletal effects require no longer application times in order to be optimized, but, of course, accompanying musculoligamentous relaxations (181) may require much longer times of application. These studies, in conjunction with the fact that patients may prefer the improved comfort of intermittent pull, suggest some advantage to an intermittent over continuous protocol (187,229,243).

Lumbar

There is limited information on this topic for lumbar traction. Cyriax has reported that continuous traction is necessary to fatigue the muscles and allow strain to fall on the joints (181). Despite this claim, no statistical difference in x-rays was noted for normal subjects treated with either continuous traction of 45.4 kg (100 lb) for 5 minutes or intermittent traction of 45.4 kg (100 lb) peak for 15 minutes (10 seconds traction, 5 seconds rest) (182). Similarly, the data seem conclusive for the skeletal effects, but it is conceivable that some of the benefits of traction occur in the soft tissues, which may respond much differently in terms of sensitivity to time. Again, improved patient tolerance seems to favor the use of intermittent traction (184).

Weight

Cervical

If cervical traction is performed with the patient in the sitting position, about 4.5 kg (10 lb) is required to counterbalance the patient’s head (233). Amounts less than this may be used initially to condition the patient to the feel of the halter and pull. In the literature, there is a great deal of variation, with the amount of force reported varying from 2.7 to 200 kg (6 to 440 lb) of traction (244). Colachis and Strohm have shown that 13.6 kg (30 lb) of traction to a neck flexed to 24 degrees can cause vertebral separation, particularly posteriorly, but an increase of force to 22.7 kg (50 lb) produced no clear-cut additional separation (242). This correlates with the earlier finding that a minimum of 11.4 kg (25 lb) was needed before the vertebrae separated (187). An initial “test dose” of 2.3 or 4.6 kg (5 or 10 lb) of traction followed by a gradual increase in weight to 20.4 or 22.7 kg (45 or 50 lb) has been advocated (187).

Lumbar

Friction becomes an important consideration in the lumbar spine. In a widely quoted article, Judovich reported that a pull equal to about one half the weight of a body part is needed to overcome friction: for the lower body, this becomes 26% of total body weight (193). Either this amount of force needs to be applied before true traction on the spine is accomplished or a split table must be used (193,234,245). Regardless of whether the effect of friction is overpowered or bypassed, another 25% or more of body weight is then needed to cause vertebral separation. For example, Colachis and Strohm used 22.7 and 45.4 kg (50 and 100 lb) of lumbar traction with a split table and measured statistically significant vertebral separation with both weights (182). Posterior vertebral separation predominated at 22.7 kg (50 lb); anterior and posterior widening occurred at 45.4 kg (100 lb). The usual range of traction is between 31.8 and 68 kg (70 and 150 lb), averaging 45.4 kg (100 lb) (206,246). With tractive forces above 45.4 kg (100 lb), the counterforce, in the form of a chest or axillary harness, causes pain and is the limiting factor for many patients.

Duration

Cervical

The exact and optimal duration of traction has not been clearly shown. Recommendations have varied from 2 minutes to 24 hours for each session (229). Intermittent traction with the
For clinicians who are willing to recommend empirically valid outcome studies, it is likely that spinal traction was used for an elderly person, a careful screening process should be used. In general, the following are considered contraindications to cervical or lumbar traction: ligamentous instability; osteomyelitis; discitis; primary or metastatic bone tumor (251); spinal cord tumor; severe osteoporosis; untreated hypertension; severe anxiety; clinical signs of myelopathy; or inadequate expertise of the practitioner providing the treatment.

Cervical traction requires additional considerations. Patients with vertebral basilar artery insufficiency could theoretically experience a stroke due to traction, particularly with the neck in extension. In elderly patients, it is difficult to eliminate the diagnosis of vertebral basilar insufficiency. Thus, patients with a history at all suggestive of vertebral basilar insufficiency should not have cervical traction. Additionally, patients with rheumatoid arthritis and other connective tissue disorders are at high risk for atlantoaxial instability. This also may be difficult to diagnose, and thus patients with rheumatoid arthritis should undergo cervical traction with extreme caution, if at all. Patients with midline herniated nucleus pulposus or acute torticollis should not have cervical traction (176).

Cervical traction has been used for a wide spectrum of painful conditions based on its physiologic effects of vertebral separation, widening of the intervertebral nerve root foramina, and possible reduction of herniated disc material and muscle relaxation. Neck and arm pain has been historically used as the indication (227).

Lumbar traction has been used as an enforcement of bed rest in the treatment of acute low back pain. There are no data to suggest that painful or nonpainful radiculopathy with clear neurologic deficit will benefit from traction.

Contraindications

There are few anecdotes and no scientific reports that clearly delineate the contraindications of traction. Thus, clinicians must rely on empirical information and opinion to guide them. Old age is a relative contraindication, given the likelihood that at least one of the described conditions is more likely to be present in elderly people. Should a practitioner wish to prescribe traction for an elderly person, a careful screening process should be used.

In general, the following are considered contraindications to cervical or lumbar traction: ligamentous instability; osteomyelitis; discitis; primary or metastatic bone tumor (251); spinal cord tumor; severe osteoporosis; untreated hypertension; severe anxiety; clinical signs of myelopathy; or inadequate expertise of the practitioner providing the treatment.

Cervical traction requires additional considerations. Patients with vertebral basilar artery insufficiency could theoretically experience a stroke due to traction, particularly with the neck in extension. In elderly patients, it is difficult to eliminate the diagnosis of vertebral basilar insufficiency. Thus, patients with a history at all suggestive of vertebral basilar insufficiency should not have cervical traction. Additionally, patients with rheumatoid arthritis and other connective tissue disorders are at high risk for atlantoaxial instability. This also may be difficult to diagnose, and thus patients with rheumatoid arthritis should undergo cervical traction with extreme caution, if at all. Patients with midline herniated nucleus pulposus or acute torticollis should not have cervical traction (176).

The literature does not give clear indications for what type of neck or low back pain can benefit from traction. In fact, most studies strongly suggest that traction does not significantly influence the long-term outcomes of neck or low back pain (248). Thus, defining indications in the face of poor scientific evidence of efficacy is very difficult. Physicians who prefer sound scientific evidence for the efficacy of treatment probably rarely use traction for spinal pain. Balancing the scientific evidence, there are numerous anecdotes of cases in which traction seems to have helped individual patients, making many clinicians and patients willing to give traction a trial. Without additional scientifically valid outcome studies, it is likely that spinal traction as a treatment for neck and back pain will gradually diminish. For clinicians who are willing to recommend empiric treatments, the indications for traction are outlined subsequently.

Theoretically, if traction can separate vertebrae and decrease the size of herniated discs, then patients with this problem and radiculopathy would be the most likely to benefit. Indeed, most studies attempt to target this population, with little demonstrated efficacy. There is considerable disagreement among investigators. Weinberger has suggested that traction is “irrational, counterproductive, nonphysiologic, and traumatic” (249). Yates has stated that there is no reason to suggest that traction will be any more useful to treat degenerative spondylosis of the cervical spine than traction to a hip or knee with osteoarthritis (250).

Lumbar traction requires additional considerations. In general, the following are considered contraindications to cervical or lumbar traction: ligamentous instability; osteomyelitis; discitis; primary or metastatic bone tumor (251); spinal cord tumor; severe osteoporosis; untreated hypertension; severe anxiety; clinical signs of myelopathy; or inadequate expertise of the practitioner providing the treatment.

Cervical traction has been used for a wide spectrum of painful conditions based on its physiologic effects of vertebral separation, widening of the intervertebral nerve root foramina, and possible reduction of herniated disc material and muscle relaxation. Neck and arm pain has been historically used as the indication (227).
Lumbar traction has fewer factors that require special consideration. However, patients with restrictive lung disease or other breathing disorders should not be subjected to the pull of a chest harness. Additionally, the following have been listed as contraindications without supporting evidence: pregnancy, active peptic ulcers, hiatal hernia, other hernias, aortic aneurysm, gross hemorrhoids, and evidence for cauda equina compromise (251). If there is any remaining question of spine instability after evaluation of the patient, imaging studies, especially flexion-extension x-rays of the spine, should be considered.

Writing the Referral
The practitioner can no more write “cervical traction, please” as a treatment prescription than to ask for “physical therapy for back pain” or “medications for infection.” One is obliged to write a detailed parameter-specific prescription for cervical or lumbar traction. As with other physical therapy referrals, the following patient information is included: age, gender, diagnosis, underlying medical conditions, precautions to be taken, patient symptoms that would signal a need to discontinue the traction, and recommended medical follow-up.

Cervical or lumbar traction should not be used as a single modality but rather as one aspect of an overall rehabilitation program. Traction as a modality to improve activity, mobility, and function (e.g., return to work) is probably most effective. If patients are required to miss 2 to 3 hours of work per day to undergo traction and physical therapy, one could argue that this is an ineffective treatment and may actually be counter-productive to the patient’s overall recovery and well-being.

The patient’s interests are not served by writing a prescription for a home cervical traction unit and expecting a pharmacy clerk to explain its use, nor is lumbar traction reliably self-applied by the patient at home. The therapist’s input therefore becomes critical to the successful and safe use of traction. If traction for any reason exacerbates symptoms, it should be discontinued and the efficacy and rationale of continued traction should be closely examined. There is no evidence to date to suggest that the patient should “endure the side effects” of traction in order to enjoy its therapeutic effect.

The following specific parameters should be outlined in all traction referrals:

Position
Body position: sitting, supine, standing, or walking on treadmill
Neck position (cervical traction): between neutral and 30 degrees of neck flexion
Hip and knee position (pelvic traction): full extension to 90 degrees of flexion

Mode of Application
Continuous or intermittent: Intermittent is preferred for patient comfort. Various times have been used (e.g., 7 to 15 seconds of traction with 5 to 10 seconds rest or 30 to 60 seconds of force with 10 to 20 seconds of rest). Some equipment allows adjustment of the time required to reach maximum pull (rise time).

Weight
A range is necessary to allow patient acclimatization. Usually, a small amount of traction is used at the beginning, with additional weight added later.
Cervical: 2.3 to 20.5 kg (5 to 45 lb) (starting low and gradually increasing or decreasing each session or over a number of sessions).
Lumbar: 22.7 to 68.1 kg (50 to 150 lb).

Other Modalities
Superficial heat for relaxation may be used before or during traction.

Time
In most instances, 10 to 30 minutes.

Frequency and Duration
Three to five times per week for 3 to 6 weeks.

Physician Reevaluation
Two to 14 days.

Guidelines for Discontinuation
If symptoms become worse or new symptoms develop (e.g., pain, dizziness, weakness, autonomic symptoms), traction should be discontinued.

Goals and Endpoint
Delineating endpoints that the physiatrist can identify is helpful. Some of these are discussed above in the section on “Duration.” Working on protocols with therapists can be useful.

Home Traction
If a supervised trial of cervical traction is successful, a home unit can be prescribed with the contingency that the therapist checks the patient for correct head and neck positioning and for the use of adequate weight, initially and at intervals. The monitoring interval may vary depending on the patient but will likely increase over time.

The Future of Traction
Traction has a long history of clinical acceptance based on surprisingly little understanding of its mechanism of action or proof of its efficacy. Although it seems natural to associate separation of the vertebrae with relief of pressure on the spinal nerves in situations in which herniated discs are present in the lateral recesses, little or no evidence exists for this mechanism. Even if it does occur in this manner, the question arises as to whether the pressure relief is permanent and, if not, whether it is reasonable to believe that a neuropathic process will reverse itself in the short time the separation is present. If not, then might not some of the perceived benefits result from
changes in the musculoligamentous system through relaxation, vasodilation, neurologic reflexes, or other numerous mechanisms that might occur? Twenty years ago, these questions were unanswerable, but current imaging and physiologic instruments are able to evaluate neurologic, muscular, and soft-tissue structures. Unlike manipulation, which deals with normal but dysfunctional tissue (i.e., no clear neurologic damage), traction generally has been studied in patients with apparent abnormal structure. Perhaps traction really only addresses the soft tissues, and the wrong subset of patients (neck or back pain secondary to radiculopathy) has been evaluated.

It seems difficult if not impossible to document the actual prevalence of the use of traction in clinical practice; based on anecdotal discussions and the paucity of recent literature on the subject, it appears that its use has diminished significantly over the past few years. It would be unfortunate if traction were to fall completely out of favor because no one has used modern technology to study it. This is particularly true at a time when spine pain is a growing societal burden, and no other highly effective treatments appear on the horizon.

REFERENCES

PART VI | MANAGEMENT METHODS


82. Winkel D. Diagnosis and Treatment of the Spine. Gaithersburg, MD: Aspen; 1996.


CHAPTER 64 | MANIPULATION, MASSAGE, AND Traction


ANALGESICS

The treatment of patients with pain is a major focus of many outpatient physiatric practices and is becoming even more important since the Joint Commission on Accreditation of Healthcare Organizations (JCAHO) mandated pain as the fifth vital sign and Congress declared, as of January 1, 2001, that this is the Decade of Pain Control and Research. Physiatrists, therefore, should have familiarity with analgesics. This section is organized by grouping analgesic medications into classes and presenting them in simple alphabetical order (Table 65-1).

Overview of Pharmacologic Pain Management by the Physiatrist

Analgesics, particularly opioids, are being prescribed with greater frequency. Although pain can be classified as nociceptive or neuropathic, patients often present with mixed pain syndromes (1). The current optimal medication choice for neuropathic pain is unclear, even though a better understanding of the pathophysiology of neuropathic pain suggests that nonopioid agents, such as antidepressants and anticonvulsants, should be more efficacious than opioids or nonsteroidal anti-inflammatory drugs (NSAIDs). Although clinical studies have not consistently demonstrated this, the many confounding factors involved in neuropathic pain states make study interpretation difficult (2).

ANALGESIC MEDICATIONS

Acetaminophen (Tylenol)
Relevance to Physiatry

Although acetaminophen is unsatisfactory as a single agent for patients requiring a powerful analgesic, it can be an effective primary or adjuvant medication for pain of mild to moderate intensity. In particular, it is considered to be the initial medication of choice for patients with osteoarthritis (OA) of the knee or hip who present without obvious signs of inflammation (3). Acetaminophen is also an alternative for some patients who experience gastrointestinal (GI) side effects with NSAIDs or celecoxib or who are at particular risk for renal toxicity associated with these agents. In addition, it is often used in combination with both opioid and nonopioid analgesics so as to decrease potential side effects (and thereby lessen interference with a rehabilitation program) by lowering the dose requirement of these other medications. In contrast to aspirin, it is often used in the pediatric rehabilitation setting due to the lack of an association with Reye's syndrome.

Other potential uses for acetaminophen include headaches for which it can be used as a single agent or as a combination product with various narcotic analgesics as well as with butalbital and caffeine (i.e., Fioricet or Esgic). Acetaminophen is also the primary agent used to suppress fever in the inpatient rehabilitation setting.

Acetaminophen has negligible anti-inflammatory effects; it, therefore, cannot be substituted for anti-inflammatory agents when treating conditions such as rheumatoid arthritis.

Due to the potential for chronic acetaminophen overdosage leading to hepatotoxicity, patients must be counseled not to exceed dosage limitations when taking scheduled doses of acetaminophen. Patients also might not realize that over-the-counter (OTC) headache or cold/flu remedies often contain acetaminophen, and this can lead to inadvertent overdosage if they are also prescribed acetaminophen either in the form of Tylenol or in combination with analgesic medications.

Mechanism of Action and Pharmacokinetics

Despite clinical use of acetaminophen for over 120 years, its exact mechanism of action is still under investigation. Studies have confirmed a pathway whereby acetaminophen decreases production of prostaglandins by reducing prostaglandin H\textsubscript{2} synthase, a COX enzyme (4). This mechanism is notably distinct from NSAIDs, which physically block arachidonic acid from accessing COX enzyme’s active site, and is hypothesized to explain its diminished action at anti-inflammatory sites—where the activated immune system produces oxidizing agents that can reverse the reduction step.

Acetaminophen is rapidly and almost completely absorbed from the upper GI tract. It is then uniformly distributed throughout the body, partly bound by plasma proteins while the unbound portion exerts the therapeutic effects. Acetaminophen can penetrate both the placenta and blood-brain barrier (5). It is swiftly metabolized in the liver and excreted by the kidneys at its recommended dosage.

Preparations and Dosing

The brand name Tylenol is frequently used interchangeably with the generic term acetaminophen. There are three major oral Tylenol preparations and they are shown in Table 65-2, which lists the dosing regimens that would provide a maximum dose of
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>μg</td>
<td>microgram</td>
</tr>
<tr>
<td>ACEn</td>
<td>angiotensin converting enzyme</td>
</tr>
<tr>
<td>Ach</td>
<td>acetylcholine</td>
</tr>
<tr>
<td>ADP</td>
<td>adenosine 5′ diphosphate</td>
</tr>
<tr>
<td>AED</td>
<td>anti-epileptic drug</td>
</tr>
<tr>
<td>AF</td>
<td>atrial fibrillation</td>
</tr>
<tr>
<td>AF/flut</td>
<td>atrial fibrillation/flutter</td>
</tr>
<tr>
<td>APAP</td>
<td>acetaminophen</td>
</tr>
<tr>
<td>aPTT</td>
<td>activated partial thromboplastin time</td>
</tr>
<tr>
<td>ASA</td>
<td>aspirin</td>
</tr>
<tr>
<td>AV node</td>
<td>atrioventricular node</td>
</tr>
<tr>
<td>bid</td>
<td>twice per day</td>
</tr>
<tr>
<td>BZD(s)</td>
<td>benzodiazepine(s)</td>
</tr>
<tr>
<td>BP</td>
<td>blood pressure</td>
</tr>
<tr>
<td>BUN</td>
<td>blood urea nitrogen</td>
</tr>
<tr>
<td>Cmax</td>
<td>maximum concentration</td>
</tr>
<tr>
<td>Ca2+</td>
<td>calcium</td>
</tr>
<tr>
<td>CaCB(s)</td>
<td>calcium channel blocker(s)</td>
</tr>
<tr>
<td>cAMP</td>
<td>3′,5′-cyclic adenosine monophosphate</td>
</tr>
<tr>
<td>CBC</td>
<td>complete blood count</td>
</tr>
<tr>
<td>CHF</td>
<td>congestive heart failure</td>
</tr>
<tr>
<td>CNS</td>
<td>central nervous system</td>
</tr>
<tr>
<td>COPD</td>
<td>chronic obstructive pulmonary disease</td>
</tr>
<tr>
<td>COX-II</td>
<td>cyclooxygenase-II</td>
</tr>
<tr>
<td>CVA(s)</td>
<td>cerebrovascular accident(s)</td>
</tr>
<tr>
<td>DA</td>
<td>dopamine</td>
</tr>
<tr>
<td>DP</td>
<td>dipyridamole</td>
</tr>
<tr>
<td>DVT</td>
<td>deep venous thrombosis</td>
</tr>
<tr>
<td>ECG</td>
<td>electrocardiogram</td>
</tr>
<tr>
<td>ER</td>
<td>extended release</td>
</tr>
<tr>
<td>FDA</td>
<td>Food and Drug Administration</td>
</tr>
<tr>
<td>GERD</td>
<td>gastroesophageal reflux disease</td>
</tr>
<tr>
<td>GI</td>
<td>gastrointestinal</td>
</tr>
<tr>
<td>H/S</td>
<td>at bedtime</td>
</tr>
<tr>
<td>H2</td>
<td>histamine 2 receptor</td>
</tr>
<tr>
<td>HDL</td>
<td>high-density lipoprotein</td>
</tr>
<tr>
<td>HIV</td>
<td>human immunodeficiency virus</td>
</tr>
<tr>
<td>HTN</td>
<td>hypertension</td>
</tr>
<tr>
<td>IM</td>
<td>intramuscular</td>
</tr>
<tr>
<td>INR</td>
<td>international normalized ratio</td>
</tr>
<tr>
<td>IR</td>
<td>immediate release</td>
</tr>
<tr>
<td>IV</td>
<td>intravenous</td>
</tr>
<tr>
<td>JCAHO</td>
<td>Joint Commission on Accreditation of Healthcare Organizations</td>
</tr>
<tr>
<td>LDL</td>
<td>low-density lipoprotein</td>
</tr>
<tr>
<td>LFTs</td>
<td>liver function tests</td>
</tr>
<tr>
<td>LMWH</td>
<td>low–molecular-weight heparin</td>
</tr>
<tr>
<td>LSD</td>
<td>lysergic acid diethylamide</td>
</tr>
<tr>
<td>MAO(s)</td>
<td>monoamine oxidase inhibitor(s)</td>
</tr>
<tr>
<td>mg(s)</td>
<td>milligram(s)</td>
</tr>
<tr>
<td>MI(s)</td>
<td>myocardial infarction(s)</td>
</tr>
<tr>
<td>mL(s)</td>
<td>milliliter(s)</td>
</tr>
<tr>
<td>MT</td>
<td>multiple trauma</td>
</tr>
<tr>
<td>Na+</td>
<td>sodium</td>
</tr>
<tr>
<td>NE</td>
<td>norepinephrine</td>
</tr>
<tr>
<td>NMDA</td>
<td>N-methyl-D-aspartate</td>
</tr>
<tr>
<td>NNT</td>
<td>number needed to treat</td>
</tr>
<tr>
<td>NO-NSAIDs</td>
<td>Nitric oxide NSAIDs</td>
</tr>
<tr>
<td>NSAID(s)</td>
<td>nonsteroidal antiinflammatory drug(s)</td>
</tr>
<tr>
<td>OA</td>
<td>osteoarthritis</td>
</tr>
<tr>
<td>OTC</td>
<td>over the counter</td>
</tr>
<tr>
<td>PO</td>
<td>by mouth</td>
</tr>
<tr>
<td>PR</td>
<td>by rectum</td>
</tr>
<tr>
<td>PCP</td>
<td>phencyclidine</td>
</tr>
<tr>
<td>PE</td>
<td>pulmonary embolus</td>
</tr>
<tr>
<td>PRN</td>
<td>as needed</td>
</tr>
<tr>
<td>PSVT</td>
<td>paroxysmal supraventricular tachycardia</td>
</tr>
<tr>
<td>PT</td>
<td>prothrombin time</td>
</tr>
<tr>
<td>PVC</td>
<td>premature ventricular contraction</td>
</tr>
<tr>
<td>q</td>
<td>every</td>
</tr>
<tr>
<td>qd</td>
<td>once per day</td>
</tr>
<tr>
<td>qid</td>
<td>four times per day</td>
</tr>
<tr>
<td>Q wk</td>
<td>weekly</td>
</tr>
<tr>
<td>qod</td>
<td>every other day</td>
</tr>
<tr>
<td>RVR</td>
<td>rapid ventricular rate</td>
</tr>
<tr>
<td>SC</td>
<td>subcutaneous</td>
</tr>
<tr>
<td>SCI</td>
<td>spinal cord injury</td>
</tr>
<tr>
<td>SL</td>
<td>sublingual</td>
</tr>
<tr>
<td>SP</td>
<td>substance P</td>
</tr>
<tr>
<td>SR</td>
<td>sustained release</td>
</tr>
<tr>
<td>SSRI(s)</td>
<td>selective serotonin reuptake inhibitor(s)</td>
</tr>
<tr>
<td>SVT</td>
<td>supraventricular tachycardia</td>
</tr>
<tr>
<td>t1/2</td>
<td>half-life</td>
</tr>
<tr>
<td>TBI</td>
<td>traumatic brain injury</td>
</tr>
<tr>
<td>TCA</td>
<td>tricyclic antidepressant</td>
</tr>
<tr>
<td>TD</td>
<td>transdermal</td>
</tr>
<tr>
<td>THA</td>
<td>total hip arthroplasty</td>
</tr>
<tr>
<td>tid</td>
<td>three times per day</td>
</tr>
<tr>
<td>TKA</td>
<td>total knee arthroplasty</td>
</tr>
<tr>
<td>UFH</td>
<td>unfractionated heparin</td>
</tr>
<tr>
<td>V tach</td>
<td>ventricular tachycardia</td>
</tr>
<tr>
<td>VLDL</td>
<td>very low-density lipoprotein</td>
</tr>
<tr>
<td>WBCs</td>
<td>white blood cells</td>
</tr>
<tr>
<td>VF</td>
<td>ventricular fibrillation</td>
</tr>
</tbody>
</table>
Acetaminophen Preparations and Dosing for Adults

<table>
<thead>
<tr>
<th>Tylenol Formulation</th>
<th>Acetaminophen Content</th>
<th>Dose (Maximum No. of Units)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tylenol</td>
<td>325 mg</td>
<td>2 tabs q4h</td>
</tr>
<tr>
<td>Extra-strength Tylenol (ES-Tylenol)</td>
<td>500 mg</td>
<td>1 tab q3hr or 2 tabs q6h</td>
</tr>
<tr>
<td>Tylenol arthritis</td>
<td>650 mg</td>
<td>2 tabs PO q8h</td>
</tr>
</tbody>
</table>

4 g per 24 hours for those patients with normal hepatic function. For those with abnormal liver function, specific dosing information should be consulted prior to prescribing. Acetaminophen is also available in liquid and suppository preparations.

**Relevant Side Effects and Drug Interactions**

Acetaminophen has an extremely favorable side-effect profile when used within recommended dosage limitations (6). In those with normal hepatic function, the maximum daily recommended dose is 4 g/day. Precaution must be taken when it is used

1. Chronically in excess of the recommended dose
2. At more than approximately 2 g/day in patients who consume excessive amounts of alcohol (defined as more than three alcoholic drinks/day)
3. When taken as a single dose in excess of approximately 15 g. Fatalities in children due to accidental overdose were reported and OTC cold medicines in the United States were subsequently withdrawn for children under 2 years of age

Acetaminophen is metabolized into an intermediate, N-acetyl-benzoquinonemine, that accumulates and may cause fatal hepatic necrosis (7). While the liver is the primary target of toxicity, chronic use of acetaminophen is also associated with renal failure (8). However, aside from cases of overdosage, acute nephrotoxic effects of acetaminophen are not common in the nonalcoholic population (9). The National Kidney Foundation, for this reason, has not amended its 1996 recommendation, making acetaminophen the drug of choice for analgesia in those with renal dysfunction (10).

Acetaminophen has a very favorable medication interaction profile. One potentially important exception, however, is warfarin as large acetaminophen doses can potentiate its effect by prolonging its half-life (11). Although this is believed to only be clinically significant in patients with a relatively high international normalized ratio (INR), the INR should be monitored in patients who are chronically taking acetaminophen while also on warfarin (12).

**Antidepressants**

**Relevance to Physiatry**

Because depression is a common consequence of illnesses and injuries, particularly if they are associated with functional loss, physiatrists should familiarize themselves with this medication class. Antidepressants are also used off-label to treat chronic nonmalignant pain syndromes and neuropathic pain (13–15). Not only do these agents treat the psychiatric component of chronic pain, but there is also evidence that they have independent analgesic effects (16,17). This chapter specifically addresses their use as analgesics.

Analgesic action of tricyclic antidepressant (TCA) medications for neuropathic pain has been extensively studied. Secondary-amine TCAs, nortriptyline and desipramine, are preferred to tertiary-amine TCAs (e., amitriptyline) because they have been shown to be equally effective and have fewer side effects (e.g., drowsiness) (18,19). Another TCA, doxepin, has shown some efficacy as a topical analgesic in managing chronic neuropathic pain (20).

A relatively newer class of antidepressants, selective serotonin reuptake inhibitors (SSRIs), is hypothesized to affect brain stem pain-modulating systems. Initial interest in using this class of medication for analgesia developed from a few promising case reports. However, a major systematic review concluded that there is very limited evidence that SSRIs are effective for management of neuropathic pain and more studies are warranted to draw definitive conclusions (21).

The serotonin-norepinephrine reuptake inhibitors (SNRIs) appear to be replacing SSRIs for neuropathic pain conditions. Venlafaxine (Effexor) and duloxetine (Cymbalta) are two SNRIs that have been studied in the treatment of pain due to postherpetic neuralgia and diabetic neuropathy. Duloxetine is currently approved for the treatment of diabetic neuropathy and venlafaxine was found to be as effective as imipramine in treatment of polyneuropathy (22,23). The popularity of SNRIs is also vested in their cost effectiveness and favorable side-effect profiles (24).

Other antidepressants that have been studied as analgesics include several that do not fall into any one particular chemical class, including trazodone and bupropion. Both of these have received somewhat less attention than the previous three classes of antidepressants but deserve further comment. Additional antidepressants in this category include mirtazapine and nefazodone. Although 5 years have passed since the last publication of this book, there is still only one case report in the peer-reviewed literature on mirtazapine and one basic science study on nefazodone in the setting of pain (25,26).

Trazodone is chemically unrelated to other antidepressants. It is rarely used for depression but is more commonly used as a hypnotic. Although there is some literature on it as an analgesic, a review of 59 randomized placebo-controlled trials of antidepressants as analgesics concluded that trazodone is not effective (27,28).

A placebo-controlled crossover trial confirmed that bupropion SR (sustained release, 150 to 300 mg daily) is effective and well tolerated in treating neuropathic pain (29). In contrast, it was found to be ineffective in treating non-neuropathic, chronic pain (30).

**Mechanism of Action and Pharmacokinetics**

TCAs increase aminergic transmission by inhibiting serotonin and norepinephrine (NE) reuptake to different...
degrees at presynaptic nerve-ending terminals. For example, amitriptyline primarily inhibits serotonin reuptake whereas nortriptyline primarily inhibits NE reuptake. As a result, TCAs elevate pain thresholds in depressed and nondepressed patients. Analgesic doses are usually lower than those for primary depression. These agents are rapidly absorbed and bind avidly to plasma albumin. Metabolism first involves demethylation of tertiary amine to secondary amine, followed by hydroxylation, glucuronidation, and eventual renal excretion as inactive metabolites.

SSRIs selectively inhibit serotonin reuptake, with less of an effect on NE reuptake. This selectivity offers the advantage of a superior side-effect profile. End results include prolonged decreased production of serotonin and down-regulation of presynaptic and postsynaptic receptors. Paroxetine and sertraline are the most frequently used agents in this class and both have a chemical structure that is unique among the SSRIs as well as other antidepressants. As a whole, this class is well absorbed orally and then undergoes hepatic metabolism followed by renal excretion.

Bupropion’s mechanism of action is still uncertain but evidence suggests that it inhibits dopamine (DA) and NE reuptake; its effect is greater on the former than on the latter (31). Bupropion is heptatically metabolized into an active metabolite, 4-hydroxybupropion, which is excreted in the urine.

Trazodone possibly acts via serotonin reuptake inhibition and mixed serotonin agonist-antagonist effects. Although extensively metabolized in the liver, it has variable clearance that may lead to accumulation in some patients.

SNRIs act as NE reuptake inhibitors via α2-adrenergic receptor blockade, serotonin reuptake inhibitors, and it binds to opioid receptors. This combined mechanism of action is somewhat similar to that of tramadol.

Preparations and Dosing
Dosage, side effects, and miscellaneous information about the most commonly used antidepressants for neuropathic and chronic pain are shown in Table 65-3.

Relevant Side Effects and Drug Interactions
Antidepressants in general are associated with a high incidence of sexual dysfunction that is often underreported in the product literature (32). Antidepressants that inhibit serotonin reuptake (e.g., SSRIs, SNRIs, trazodone) can cause “serotonin syndrome,” a hyperexcitable state of nervousness and insomnia.

TCA side effects are mainly anticholinergic and include dry mouth, blurred vision, tachycardia, constipation, aggravation of glaucoma, and urinary retention. They also cause anti-histamine side effects such as sedation (therefore they are often prescribed as a single bedtime dose) and weight gain, related to an increased appetite for carbohydrates. TCAs also exert some quinidine-like cardiac effects, including atrioventricular (AV) conduction-time prolongation. Cognitive/behavioral alterations (e.g., agitation and memory impairment) appear at plasma TCA concentrations greater than 0.450 μg/mL (33);

many TCAs, notably dothiepin, are fatal at concentrations greater than 1 μg/mL (34). Nortriptyline is generally considered superior to all TCAs because it is more potent and has a comparatively wide therapeutic range (35). Elderly and otherwise medically fragile patients should probably be started on nortriptyline rather than amitriptyline for the above reason and because side effects such as orthostatic hypotension and significant morning sedation, which can potentially interfere with rehabilitation efforts in this patient population, are relatively less common.

Since SSRIs have a relatively specific effect on serotonin reuptake without a significant effect on NE reuptake, their side-effect profile is generally superior to TCAs—especially with respect to cardiovascular issues—and they are much safer in cases of overdose. However, abrupt cessation of SSRI has been reported to cause SSRI discontinuation syndrome in some individuals. This syndrome includes dizziness, light-headedness, insomnia, fatigue, anxiety/agitation, nausea, headache, and sensory disturbance.

Bupropion has caused seizures and interference with cardiac conduction (ventricular arrhythmias and third-degree heart block). Idiosyncratic reactions including Stevens-Johnson syndrome, rhabdomyolysis, and severe hepatotoxicity have also been reported (36–38). There is evidence that bupropion is a cytochrome P450–2D6 (CYP2D6) inhibitor, but the clinical significance of this has yet to be established (39). The sustained-release (SR) bupropion is generally better tolerated than the immediate-release (IR) form.

Trazodone can be quite sedating and possesses other mild anticholinergic effects, but these are less generally than those from TCAs. It also exhibits α-adrenergic blocking properties, which can cause penile or clitoral priapism (40).

TCAs, SSRIs, and bupropion should not be used in patients taking monoamine oxidase inhibitors (MAOIs) and should be instituted cautiously in patients who have been off of MAOIs for at least 2 weeks. The only exception to this rule is nortriptyline, which can be safely combined with MAOIs or sertraline (an SSRI) (35). Concomitant use of other TCAs or any of SSRIs and MAOIs can cause hyperpyretic crises, seizures, and death. TCAs should also be used cautiously in patients taking other anticholinergic medications, neuroleptics, or central nervous system (CNS) depressants.

It is not known whether interactions occur between trazodone and MAOIs. Trazodone may increase serum digoxin and phenytoin levels and can cause either an increase or a decrease in prothrombin times (PTs) in patients on warfarin (41).

Venlafaxine’s most common side effects are from increased serotonin levels (irritability, insomnia, and sexual dysfunction) and also include constipation and nausea. There are several case reports of false-positive phencyclidine (PCP) results from ingesting high dosage of venlafaxine (42,43).

Corticosteroids
Relevance to Physiatry
Anti-inflammatory effects of corticosteroids are usually more important to the physiatrist than their mineralocorticoid and
### TABLE 65.3 Antidepressants Used in the Treatment of Neuropathic Pain

<table>
<thead>
<tr>
<th>Generic Name</th>
<th>Dose (mg)</th>
<th>Neuropathic Pain and (Depression)</th>
<th>Miscellaneous</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>[TCAs]</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Amitriptyline (Elavil/Saroten/Endep/Vanatrip) | 10–100 @ H/S (150–300/d); begin @ 12.5–25 qH/S, and titrated as tolerated | Dry mouth and sedation very common  
Demethylated to nortriptyline  
First metabolite of amitriptyline; less side effects but not as potent  
Minor side effects |
| Nortriptyline (Aventyl, Pamelor) | 10–30 @ H/S (50–150/d) |                                                                               |
| Doxepin (Sinequan) topical | Topical application of 3.3% doxepin, 0.025% capsaicin, and 3.3% doxepin/0.025% capsaicin produces analgesia of similar magnitude. The combination produces more rapid analgesia. Cream: 50 mg/g |                                                                               |
| **[SSRIs]**           |           |                                   |                                                                               |
| Citalopram (Celexa)   | 20–40 qd (20–60 qd) | Relatively short half-life  
Very popular when first released; blamed in the press as a contributing factor to several high-profile murders | Least studied of the SSRIs for pain  
Most selective of the SSRIs |
| Fluoxetine (Prozac)   | 20 qd (20–80 qd) | 90 qw of Prozac Qwk | Tablets and oral concentrate; serotonin syndrome (hyperserotonergic state) with tramadol coadministration; also used for obsessive-compulsive disorder (OCD) and post-traumatic stress disorder (PTSD) |
| (Prozac Weekly)       |            | (Sarafem)                        |                                                                               |
| Fluvoxamine (Luvox)   | 100 qd (50–150 bid) | 20–40 qd (20–60 qd) throughout menstrual cycle or 14 prior to menses |                                                                              |
| Paroxetine (Paxil)    | 20–50 qd (20–50 qd) |                                                                               |
| (Paxil CR)            | 25 mg PO qam max 62.5 mg/d |                                                                               |
| Sertraline (Zoloft)   | 50–150 qd (50–200 qd) |                                                                               |
| **[Other antidepressants]** |           |                                   |                                                                               |
| Bupropion SR (Wellbutrin SR) | 150–300 qd (100–450 qd) | SR formulation have a better side-effect profile vs. IR preparation, esp. for sexual dysfunction and seizures; also used for smoking cessation (Zyban) |                                                                              |
| Wellbutrin, Zyban, Buproban | Start 100 mg PO bid IR tab increase to tid |                                                                               |
| Wellbutrin XL         | 150–300 qd (100–450 qd) |                                                                               |
| Trazodone (Desyrel)   | Start 50–150 mg/d PO in divided doses; @ H/S (200–300 bid), usual effective dose is 400–600 mg/d | Priapism that can be severe; less anticholinergic side effects vs. TCAs |                                |

(continued)
androgenic/estrogenic effects. Physiatrists today are using corticosteroids for a range of injection procedures, including fluoroscopic-guided spinal injection procedures and peripheral joint injection procedures. They are also prescribing them in either short, tapering oral courses (e.g., Medrol Dosepak) for patients with radiculopathy and other localized musculoskeletal conditions; or chronically for systemic inflammatory diseases. Besides their oral and injectable forms, corticosteroids can additionally be delivered transdermally by iontophoresis or phonophoresis.

**Mechanism of Action and Pharmacokinetics**

Corticosteroids bind to receptors within a target cell’s nucleus and cause an alteration in protein synthesis. These altered proteins then exert various mineralocorticoid, androgenic/estrogenic, and glucocorticoid effects. Corticosteroids are classified into one of these three categories depending upon their predominant effect. The focus of this section is the glucocorticosteroid class. At physiologic, but not pharmacologic, doses, glucocorticosteroid exerts anti-inflammatory and immunosuppressive effects via the following mechanisms:

1. Inhibition of prostaglandin and leukotriene synthesis (research suggests that this occurs by preventing arachidonic acid release from phospholipids; this contrasts with NSAIDs and COX-II inhibitors, both of which act at a later step in prostaglandin synthesis via inhibition of COX isoenzymes)
2. Inhibition of chemotactic factor release, leading to a diminished attraction of white blood cells (WBCs) to sites of inflammation
3. Decrease in circulating lymphocytes and monocytes
4. Reduction of vascular permeability by acting as vasoconstrictors or by inhibiting vasodilator release (e.g., histamines and kinins)
5. Stabilization of lysosomal membranes (occurs only at higher steroid doses)

Oral glucocorticoids are hepatically metabolized and renally excreted at a rate proportional to the particular agent’s water solubility. Hence, longer-acting glucocorticoids are less water soluble.

---

**TABLE 65.3 Antidepressants Used in the Treatment of Neuropathic Pain* (Continued)**

<table>
<thead>
<tr>
<th>Generic Name</th>
<th>Dose (mg) Neurpathic Pain and (Depression)</th>
<th>Miscellaneous</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mirtazapine (Remeron)</td>
<td>Initial dose 15 mg taken at bedtime. The dose may be increased in 15-mg increments every 1 or 2 wk as needed. Typical doses range between 15 and 45 mg. Dosages above 45 mg/d are not recommended</td>
<td>Side effects are sleepiness and nausea. Other common side effects are dizziness, increased appetite, and weight gain. Less common adverse effects include weakness and muscle aches, flu-like symptoms, low blood-cell counts, high cholesterol, back pain, chest pain, rapid heartbeats, dry mouth, constipation, water retention, difficulty sleeping, nightmares, abnormal thoughts, vision disturbances, ringing in the ears, abnormal taste in the mouth, tremor, confusion, upset stomach, and increased urination</td>
</tr>
<tr>
<td>Nefazodone (Serzone)</td>
<td>Initial dose of nefazodone is 100 mg taken by mouth twice daily. The dose may be increased in 100 or 200 mg increments once a week. Most commonly, final dosages range between 300 and 600 mg taken by mouth each day</td>
<td>Side effects: dizziness, difficulty sleeping, weakness, or agitation. Other common adverse effects are sleepiness, dry mouth, nausea, constipation, blurred vision, and confusion</td>
</tr>
<tr>
<td>Venlafaxine (Effexor)</td>
<td>18.75–75 qd, divided bid or tid (37.5–75 divided bid or tid) 75 mg/d divided bid-tid MAX 375 mg/d 25, 37.5, 50, 75,100</td>
<td>An extended-release form (Effexor XR) is used for depression but not studied yet for pain</td>
</tr>
<tr>
<td>Duloxetine (Cymbalta)</td>
<td>Total dose of 40 mg/d (given as 20 mg bid) to 60 mg/d (given either once a day or as 30 mg bid), no evidence that doses &gt;60 mg/d confer any additional benefits</td>
<td>Side effects: impaired thinking or reactions. Be careful if you drive or do anything that requires you to be awake and alert</td>
</tr>
</tbody>
</table>

*Note: Only those generally considered used for neuropathic pain are shown in the table.*
CHAPTER 65 | PHARMACOTHERAPY OF DISABILITY

PREPARATIONS AND DOSING

The two most commonly used oral steroid preparations in many physiatric practices are prednisone and methylprednisolone. The latter is often prescribed as Medrol Dosepak of 4-mg tabs, which provides an initial 24 mg of Medrol (equivalent to 30 mg of prednisone) and tapers to 0 mg over 7 days. The popularity of Dosepak comes from the fact that patients’ instructions are conveniently printed on the package and it eliminates the need for patients to count out a different set of pills each day. Potential drawbacks to Medrol Dosepak are its higher expense compared to generic prednisone and a limited peak dose of merely 30-mg prednisone. Some physicians overcome the low peak dose by prescribing two Medrol Dosepak to be taken simultaneously.

Corticosteroid selection can be made on the basis of equivalent cortisone dose, relative anti-inflammatory potency, relative mineralocorticoid potency, and onset and duration of action (44,45) (Table 65-4). For comparison, physiologic steroid doses are equivalent to 30 mg/day of hydrocortisone (7.5 mg/day of prednisone), whereas stress doses are equivalent to 300 mg/day of hydrocortisone (75 mg/day of prednisone). General dosing guidelines have also been developed (see Tables 65-4 and 65-5), but recent popularity of injectable preparations.

**TABLE 65.4 Corticosteroid Preparations**

<table>
<thead>
<tr>
<th>Corticosteroid Generic (Trade) Name</th>
<th>Route</th>
<th>Equiv. Oral Dose (mg)</th>
<th>Relative Potencies: Antiinflammatory (Mineralocorticoid)</th>
<th>Onset</th>
<th>Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Betamethasone (Celestone)</td>
<td>PO/IM</td>
<td>0.6–0.75</td>
<td>20–30 (0)</td>
<td>Very fast</td>
<td>Long</td>
</tr>
<tr>
<td>Cortisone (Cortone)</td>
<td>PO</td>
<td>25</td>
<td>0.8 (2)</td>
<td>Fast, slow (LA)</td>
<td>Short</td>
</tr>
<tr>
<td>Dexamethasone (Decadron, Decadron-LA)</td>
<td>PO/IM/IV</td>
<td>0.75</td>
<td>20–30 (0)</td>
<td>Long</td>
<td></td>
</tr>
<tr>
<td>Hydrocortisone (Cortef, Solu-Cortef, Hydrcortone)</td>
<td>PO/IM/IV</td>
<td>20</td>
<td>1 (2)</td>
<td>Moderate (Cortef), fast (Solu-Cortef), slow (Hydrcortone)</td>
<td>Short</td>
</tr>
<tr>
<td>Methylprednisolone (Medrol, Medrol Dosepack, SoluMedrol)</td>
<td>PO/IM/IV</td>
<td>4</td>
<td>5 (0)</td>
<td>Slow</td>
<td>Intermediate</td>
</tr>
<tr>
<td>Prednisolone (Hydeltra)</td>
<td>PO/IM/IV</td>
<td>5</td>
<td>4 (1)</td>
<td>Moderate</td>
<td>Intermediate</td>
</tr>
<tr>
<td>Prednisone (Deltasone, Orasone)</td>
<td>PO</td>
<td>5</td>
<td>4 (1)</td>
<td>Fast</td>
<td>Intermediate</td>
</tr>
<tr>
<td>Triamcinolone (Aristocort, Kenacort, Kenalog)</td>
<td>PO/IM</td>
<td>4</td>
<td>5 (0)</td>
<td>Slow</td>
<td>Intermediate</td>
</tr>
</tbody>
</table>

**TABLE 65.5 Corticosteroid Dosing Guidelines**

- Use these only after less toxic therapy has been ineffective or is not an alternative
- Use the smallest corticosteroid amount that can control symptoms
- Administer the corticosteroid locally rather than systemically whenever possible
- Short-term use: Dosing qd (preferably in the a.m.) is more convenient and causes less adrenal suppression than qid dosing at quarter the total dose
- Chronic use: Dosing qd is less likely to suppress adrenal function
- Do not use the term *steroids* because of this word’s negative connotations. Although the terms cortisone and prednisone may also have negative connotations, explain that osteoporosis and truncal obesity only occur with chronic use
- Forewarn patients that oral steroids typically cause a metallic taste
- Adrenal suppression is likely for dose, potency, and duration as follows:
  - Doses ≥100 mg hydrocortisone (25 mg prednisone) daily × 3 d
  - Doses ≥30 mg hydrocortisone (7.5 mg prednisone) daily × 30 d
- Wean patients off over weeks or months if taking steroids for more than several weeks
- If unsure that patient has become adrenally suppressed, refer to endocrinologist for metyrapone or insulin-tolerance testing. Recovery of adrenal function is variable
- For corticosteroid injections:
  - Can decrease chance of corticosteroid arthropathy with limit of: 3/y; 20/lifetime
  - Never inject directly into a tendon and avoid weight-bearing peritendinous injections (e.g., Achilles, patellar, posterior tibial) or risk tendon rupture
corticosteroids has raised concern that there is a lack of uniform guidelines for treating intra-articular joint conditions (46,47).

Relevant Side Effects and Drug Interactions
Most side effects occur after prolonged administration and many are simply manifestations of Cushing’s syndrome as shown in Table 65-6. Among these conditions, steroid myopathy and avascular necrosis are particularly pertinent to physiatry. The first, necrosis of the femoral or humeral head, is a rare idiosyncratic event that can occur after a short course of prednisone. Physiatrists who routinely perform electrodiagnostic studies are likely to be familiar with the need to “rule out steroid myopathy.”

Skin depigmentation and subcutaneous atrophy are dermatological complications that can occur with corticosteroid injections but can be minimized by adding local anesthetic or normal saline vehicle into the injectate and flushing the needle of residual corticosteroid with saline or local anesthetic injection before removal from the skin. Skin changes from chronic oral corticosteroids can lead to pressure ulcers and easy bruising.

Acceleration of corticosteroid metabolism occurs with medications that induce hepatic microsomal enzymes, especially phenobarbital, phenytoin, carbamazepine, and rifampin. In contrast, corticosteroid potency is increased by NSAIDs and exogenous estrogens (47). Clinicians should consider discontinuing NSAIDs or switching to a COX-II inhibitor if and exogenous estrogens (47). Clinicians should consider discontinuing NSAIDs or switching to a COX-II inhibitor if concomitant corticosteroid use is needed as corticosteroids are risk factors for NSAID-induced GI bleeding.

Lastly, although not a true side effect, a potential problem with corticosteroids is that they mask forewarning inflammation of various disorders. Thus, there is a tendency for patients to overvalue temporary relief and ignore the underlying disorder. An example is a patient who has received a subacromial steroid injection and soon resumes repetitive overhead activities that initially led to impingement.

<table>
<thead>
<tr>
<th>TABLE 65.6 Corticosteroid Side Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Organ System</strong></td>
</tr>
<tr>
<td>Cardiovascular</td>
</tr>
<tr>
<td>Endocrine/metabolism</td>
</tr>
<tr>
<td>Gastrointestinal</td>
</tr>
<tr>
<td>Musculoskeletal</td>
</tr>
<tr>
<td>Skin</td>
</tr>
<tr>
<td><strong>Side Effect</strong></td>
</tr>
<tr>
<td>Behavior and mood alteration</td>
</tr>
<tr>
<td>Fluid retention; HTN</td>
</tr>
<tr>
<td>Adrenal atrophy; amenorrhea; appetite increase; glucose tolerance impairment; hyponatremia and hypokalemia; weight gain leading to “moon facies”</td>
</tr>
<tr>
<td>Aggravation of peptic ulcer disease</td>
</tr>
<tr>
<td>Avascular necrosis; bone demineralization; steroid myopathy</td>
</tr>
<tr>
<td>Acne, depigmentation, and subcutaneous atrophy with injection, fatty deposition leading to “buffalo hump”; hirsutism, skin thinning</td>
</tr>
</tbody>
</table>

Membrane-Stabilizing Medications: Antiarrhythmics

Relevance to Physiatry
There are three circumstances under which a physiatrist might prescribe antiarrhythmics: first, in the inpatient rehabilitation setting, for patient who needs ongoing treatment for existing cardiac conditions; second, for a patient with neuropathic pain who responds to off-label use of type I antiarrhythmics (i.e., mexiletine, tocainide, lidocaine, and phenytoin); third, for a patient with myotonia-associated pain from certain neuromuscular disorders. Intravenous (IV) lidocaine, as an analgesic agent, is not discussed in detail in this chapter because it is infrequently used other than as a predictive test for mexiletine treatment in highly specialized pain management clinics (48). In contrast, the low sensitivity of IV lidocaine infusion renders the test inappropriate for definitive diagnosis of neuropathic pain (49).

Literature pertaining to oral antiarrhythmics for neuropathic pain is still limited to mexiletine because other oral analogues (e.g., flecainide and tocainide) have some potentially lethal adverse effects (50). It is noteworthy that, prior to tocainide’s withdrawal from U.S. market over safety concerns, there were successful reports on using the agent as treatment for myotonic pain in paramyotonia congenita and Thomsen-Becker myotonia (51).

Earlier case reports and prospective studies suggested that mexiletine is efficacious and safe in various neuropathic pain states including pain from peripheral nerve damage, diabetic neuropathy, alcoholic neuropathy, phantom limb pain, multiple sclerosis complicated by painful dysesthesias, and thalamic pain syndrome (50,52–57). A 2005 systematic review of local anesthetics found mexiletine (at a median dose of 600 mg/day) to be “superior to placebo in relieving neuropathic pain… as effective as other analgesics used for this condition” (58). Critics, nonetheless, persist to this day because most studies were of relatively short duration and fewer than 400 patients have been studied altogether (59).

Mechanism of Action and Pharmacokinetics
Type I antiarrhythmics (e.g., lidocaine, mexiletine, and tocainide) block Na+ channels in nerve and muscle cell membranes with a subsequent reduction in the number of abnormal ectopic impulse generated by dysfunctional peripheral nerves. Lidocaine and oral analogues differ in that the latter have low first-pass metabolism, which enhances their oral bioavailability.

Preparations and Dosing
Mexiletine is available in 150-, 200-, and 250-mg caplets. Neuropathic pain doses are lower (150 to 300 mg tid) than those used for arrhythmias (200 to 400 mg tid). It can be initiated as a 150-mg-per-day regimen, titrated weekly.

Relevant Side Effects and Drug Interactions
Mexiletine’s potential side effects are acute in onset and can involve the GI, neurologic, and cardiovascular systems. One mexiletine study of experimentally induced pain found that
analgesic dosing caused side effects at an average of 993 mg/day, whereas another study found negligible side effects at doses up to 900 mg/day (60). GI side effects include nausea, anorexia, and gastric irradiation in up to 40% of patients. Another 10% of patients experience neurologic side effects similar to those of other class I antiarrhythmics, including dizziness, visual disturbances, anxiety in patients with history of anxiety disorder, tremor, and altered coordination. Individuals with abnormal cardiac conduction are also at risk of mexiletine-induced excitation; mexiletine is absolutely contraindicated in second- and third-degree heart block uncontrolled by a pacemaker. Patients in fact should obtain an electrocardiogram (ECG) prior to starting this medication.

Mexiletine's pharmacologic disposition is susceptible to multiple alterations. For example, opioid analgesics, atropine, and antacids slow its absorption while metoclopramide enhances it. Phenytoin, rifampin, and smoking increase its metabolism. Mexiletine, in turn, may significantly reduce the clearance of theophylline and caffeine (61).

**Membrane-Stabilizing Medications: Anticonvulsants**

**Relevance to Physiatry**

When anticonvulsants are used to treat neuropathic pain, they are generally referred to as *membrane-stabilizing medications*. Their initial use as antineuralgic drugs in the 1960s was derived merely from positive clinical observations. Current research efforts are dedicated to understanding this phenomenon. A popular theory now attributes similarities between neuropathic pain and epilepsy pathophysiology models to the efficacy of anticonvulsants in treating both conditions (62). As supportive evidence continues to accumulate, anticonvulsants have marked a new era in the treatment of pain.

Various anticonvulsants have yielded good results thus far and pregabalin is the latest addition to the list. Specific drugs will be discussed individually as there are some significant differences among them with respect to side effects and mechanisms of action (Tables 65-7 and 65-8).

**Gabapentin (Neurontin)**

**Relevance to Physiatry**

Gabapentin is still among the first-line treatments for neuropathic pain. It has been shown to be as effective as TCAs and carbamazepine and has a highly favorable side-effect profile and reputation for minimal drug interactions (63). Although evidence is lacking for its efficacy in acute pain states (64), there is now growing evidence supporting its efficacy in treating painful diabetic neuropathy (65–68). It has also been investigated for spasticity reduction in spinal cord injury (SCI) patients (69,70).

**Mechanism of Action and Pharmacokinetics**

See Tables 65-7 and 65-8.

Gabapentin's mechanism of action is still poorly understood. It was originally thought to inhibit gamma-aminobutyric acid (GABA) receptors because of its structural similarity to GABA, a major CNS excitatory neurotransmitter. While there is evidence suggesting that it does indeed interact with presynaptic GABA-B receptors to reduce glutamate release, other studies point to more possibilities. Specifically, gabapentin has been implicated as a calcium channel blocker (CaCB) and hippocampal CA1 neural enhancer. There is also evidence that it may raise the interneuron pool excitability threshold of polysynaptic reflexes (71).

Gabapentin is transported from the GI tract into the bloodstream by the amino acid transport system. It does not bind to plasma protein, is not metabolized, and is ultimately excreted by the kidneys at the rate of creatinine clearance.

**Preparations and Dosing**

The reported dosage that provides adequate relief from neuropathic pain ranges from 900 to 2,400 mg per day, divided tid. When initiating therapy, a 300-mg dose at bedtime on the first day, then bid dosing on the second day, and tid dosing thereafter can be used as a way of helping patients accommodate to its CNS depressive effects. A maximum of 3,600-mg daily dose has been well tolerated in a small number of patients for short duration.

**Relevant Side Effects and Drug Interactions**

CNS depression (e.g., somnolence, dizziness, ataxia, and fatigue) is the main side effect. Nystagmus has also been reported. Side effects are generally transient with resolution in 2 weeks. There have been only rare reports of adverse events (e.g., rash, leukopenia, increased blood urea nitrogen (BUN), thrombocytopenia, and nonlethal ECG abnormalities) that required its discontinuation. Given the rarity of serious adverse events, routine laboratory monitoring and monitoring of serum gabapentin levels are not indicated.

There are two documented drug interactions, but both are deemed to be clinically insignificant. Cimetidine minimally decreases renal gabapentin excretion, and Maalox reduces gabapentin bioavailability by 20%. This lack of significant drug interactions is expected given its pharmacokinetics profile.

**Carbamazepine (Tegretol)**

**Relevance to Physiatry**

Carbamazepine was the first anticonvulsant to be studied in neuropathic pain clinical trials. Results from these trials confirmed its indication in treating trigeminal neuralgia (Food and Drug Administration [FDA] approved), glossopharyngeal neuralgia, painful diabetic neuropathy, and postherpetic neuralgia. It has not been as extensively studied in other neuropathic conditions. Traditional belief is that carbamazepine is especially effective for neuropathic pain that is acute and lancinating, as is often found in postamputation neuroma. Its relative lack of CNS side effects compared to other anticonvulsants offers an obvious advantage with respect to functional activities. The unfortunate combination of potential hematologic toxicity, the need for periodic blood work, baseline and periodic eye examinations, and numerous medication interactions renders carbamazepine a less attractive choice.
**TABLE 65.7**  
Anticonvulsant Efficacy and Pharmacokinetics in Neuropathic Pain

<table>
<thead>
<tr>
<th>Efficacy</th>
<th>Medication</th>
<th>Specific Neuropathic Pain Uses</th>
<th>Pharmacokinetics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Efficacious</td>
<td>Gabapentin (Neurontin)</td>
<td>Especially diabetic neuropathy and postherpetic neuralgia (FDA approved)</td>
<td>Not protein bound or metabolized; renal excretion</td>
</tr>
<tr>
<td></td>
<td>Carbamazepine (Tegretol)</td>
<td>Trigeminal neuralgia (FDA-approved), glossopharyngeal neuralgia, painful diabetic neuropathy, and postherpetic neuralgia</td>
<td>Binds and prolongs inactivation of voltage-dependent sodium channels. The number of action potentials is consequently decreased. Highly plasma protein-bound, variable $t_1/2$ as it induces its own metabolism.</td>
</tr>
<tr>
<td></td>
<td>Pregabalin (Lyrica)</td>
<td>FDA approved in treating diabetic neuropathy pain, postherpetic neuralgia pain. Used off-label for other neuropathic pain conditions. Mounting evidence it can manage pain due to fibromyalgia.</td>
<td>Does not bind to plasma proteins and nearly the entire dose is excreted unchanged in the urine, with elimination following first-order kinetics.</td>
</tr>
<tr>
<td>Unclear</td>
<td>Clonazepam (Klonopin)</td>
<td>Some efficacy in trigeminal neuralgia</td>
<td>Good absorption; highly plasma protein-bound, lipid soluble, and hepatically metabolized.</td>
</tr>
<tr>
<td></td>
<td>Lamotrigine (Lamictal)</td>
<td>Primarily indicated in treatment for epilepsy and bipolar disorder. Provides minimal, if any, therapeutic effects in acute and chronic pain. Some efficacy in trigeminal neuralgia; peripheral neuropathy poststroke syndromes</td>
<td>Good oral absorption; hepatic conjugation; renal excretion.</td>
</tr>
<tr>
<td></td>
<td>Oxcarbazepine (Trileptal)</td>
<td>Efficacy in newly diagnosed and refractory trigeminal neuralgia and possible merit in areas of neuropathic pain and bipolar disorder other neuropathic pain conditions</td>
<td>Hepatically metabolized to its active metabolite; renal excretion.</td>
</tr>
<tr>
<td></td>
<td>Phenobarbital (Solfoton)</td>
<td>No clinical studies in humans in peer-reviewed literature</td>
<td>Moderate protein binding; hepatic metabolism; pH-dependent renal excretion.</td>
</tr>
<tr>
<td></td>
<td>Phenytoin (Dilantin)</td>
<td>Conflicting results in trigeminal neuralgia and diabetic neuropathy</td>
<td>Metabolism saturable at high plasma levels; thus large concentration increases from additional small doses.</td>
</tr>
<tr>
<td></td>
<td>Tiagabine (Gabitril)</td>
<td>Two small trials showed beneficial outcome in painful sensory neuropathy</td>
<td>Highly protein bound; at least 2 metabolic pathways.</td>
</tr>
<tr>
<td></td>
<td>Topiramate (Topamax)</td>
<td>Conflicting results in studies of neuropathic pain and diabetic neuropathy. Some evidence in refractory intercostal neuralgia, trigeminal neuralgia, and trigeminal autonomic cephalgias</td>
<td>Rapidly absorbed orally, a third of the drug is metabolized by the hepatic CYP450 system into inactive metabolites, remainder excreted renally unchanged.</td>
</tr>
<tr>
<td></td>
<td>Valproate (valproic acid, Depakene)</td>
<td>Efficacy in neuropathic cancer pain but not paraplegia central pain. Mixed results in neuropathic pain, postherpetic neuralgia, and polyneuropathy</td>
<td>A lipid-soluble compound, rapidly absorbed and becomes tightly protein bound. Metabolized in the liver through oxidation and glucuronidation pathways. Active metabolites and a small, unchanged portion are renally eliminated.</td>
</tr>
<tr>
<td></td>
<td>Zonisamide (Zonegran)</td>
<td>Studies in neuropathic pain are lacking</td>
<td>Renally excreted intact and as a glucuronide metabolite.</td>
</tr>
</tbody>
</table>

**Mechanism of Action and Pharmacokinetics**

See Tables 65-7 and 65-8.

Carbamazepine binds and prolongs inactivation of voltage-dependent sodium channels. The number of action potentials is consequently decreased. Because of its high lipid solubility, it is slowly absorbed into the body following oral administration and becomes highly protein bound. It induces the hepatic CYP450 enzymes and this increases the metabolism of multiple medications—including its own.
agranulocytosis (rare), hepatotoxicity, skin reactions (e.g., Stevens-Johnson syndrome and toxic epidermal necrolysis), and (72), to a lesser degree, renal dysfunction. Prior to initiating carbamazepine, a complete blood count (CBC), liver function tests (LFTs), BUN and urinalysis, reticulocyte count, and serum iron levels are recommended. CBC and LFTs should be reviewed periodically and, if toxicity is suspected, consideration should be given for medication discontinuation.

Due to the pharmacokinetics described previously, carbamazepine interacts with many medications (Table 65-9). In addition, it should not be given to individuals with TCA hypersensitivity due to potential crossreactivity, nor should it be used within 2 weeks of an MAOI.

Clonazepam (Klonopin)

Relevance to Physiatry

This benzodiazepine (BNZ) has been used to provide relief in neuropathic pain states (especially in patients with trigeminal neuralgia who are either intolerant to or have failed carbamazepine, baclofen, or phenytoin) and movement disorders (e.g., sleep-related nocturnal myoclonus, restless legs syndrome, tardive dyskinesia, phantom limb pain, and opioid-related myoclonic jerks).

Mechanism of Action and Pharmacokinetics

See Tables 65-7 and 65-8.

Clonazepam is avidly bound to plasma protein and is highly lipid soluble. It is acetylated into nonactive metabolites in the liver and is gradually excreted by the kidneys. Low level of metabolites is also excreted in breast milk. In spite of this, there have not been reports of adverse outcome in newborns; lactation is therefore not a contraindication (73).

Preparations and Dosing

Carbamazepine is available in tablets, chewable tablets, extended-release capsules, syrup, suspensions, and rectal suppositories preparations. Although there are no dosing guidelines for trigeminal neuralgia, it is commonly used to treat this condition with 100 mg bid initially, and then the dose is gradually increased to a maximum of 400 mg tid. Given the potential for hematologic toxicity, the maintenance dose should be reduced to the minimum effective level. The extended-release form, Tegretol-XR, can be given bid to achieve the same total daily doses as described above.

Relevant Side Effects and Drug Interactions

Severe toxicity can occur with carbamazepine including leukopenia and thrombocytopenia, aplastic anemia and

<table>
<thead>
<tr>
<th>Proposed Mechanism</th>
<th>Medications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Na+ channel blocker</td>
<td>Carbamazepine; lamotrigine; oxcarbazepine; phenytoin; valproate; zonisamide</td>
</tr>
<tr>
<td>Ca++ channel blocker</td>
<td>Gabapentin; oxcarbazepine; zonisamide</td>
</tr>
<tr>
<td>GABA receptor activity</td>
<td>Barbiturates; BNZs</td>
</tr>
<tr>
<td>GABA metabolism</td>
<td>Gabapentin; tiagabine; valproate</td>
</tr>
<tr>
<td>Glutamate receptor activity</td>
<td>Carbamazepine; lamotrigine; topiramate</td>
</tr>
<tr>
<td>Glutamate metabolism</td>
<td>Gabapentin</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Interaction</th>
<th>Decreased Serum Level</th>
<th>Increased Serum Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medications whose serum levels are affected by carbamazepine</td>
<td>Acetaminophen, alprazolam, amitriptyline, bupropion, buspirone, citalopram, clobazam, clonazepam, clozapine, cyclosporine, delavirdine, desipramine, diazepam, dicumarol, doxycycline, ethosuximide, felbamate, felepidine, glucocorticosteroids, haloperidol, itraconazole, lamotrigine, levethyroxine, lorazepam, methadone, midazolam, mirtazapine, nortriptyline, olanzapine oral and other hormonal contraceptives, oxcarbazepine, phenytoin, primidone, protease inhibitors, quetiapine, risperidone, theophylline, tiagabine, topiramate, tramadol, triazolam, trazodone, valproate, warfarin, zonisamide</td>
<td>Clomipramine, phenytoin, primidone</td>
</tr>
<tr>
<td>Medications that affect serum carbamazepine levels</td>
<td>Cisplatin, doxorubicin HCL, felbamate, methsuximide phenobarbital, phenytoin, primidone, rifampin, theophylline</td>
<td>Acetazolamide, azole antifungals, CaCBs, cimetidine, clarithromycin, dalfopristin, danazol, delavirdine, diltiazem erythromycin, fluoxetine, fluvoxamine, grapefruit juice, isoniazid, itraconazole, ketoconazole, loratadine, nefazodone niacinamide, nicotinamide, protease inhibitors, propoxyphene, quinine, quinupristin, troleandomycin, valproate, verapamil, zileuton</td>
</tr>
</tbody>
</table>
Preparations and Dosing
Clonazepam is available as 0.5-mg, 1-mg, and 2-mg tablets. These tablets distinguish themselves with a unique K-shaped perforation in the middle of the pill. For movement disorders, clonazepam is begun at either 0.5 mg at bedtime or 0.5 mg tid, and it can be titrated up to 2 mg tid. A daily dose of 0.5 to 1.0 mg is recommended to treat trigeminal neuralgia (74).

Relevant Side Effects and Drug Interactions
Ataxia and personality changes can develop early in the treatment course but may subside with long-term use. At the other end of the spectrum, withdrawal often causes a flu-like syndrome, and abrupt discontinuation of a chronic, high-dose regimen can even lead to seizures. Moreover, common to all BNZs, chronic clonazepam use can result in psychological addiction and physical tolerance. Caution should be exercised when clonazepam is given along with another CNS depressant.

Lamotrigine (Lamictal, Lamictal CD)
Relevance to Physiatry
Lamotrigine is primarily indicated in treatment for epilepsy and bipolar disorder. Multiple, well-designed clinical trials have investigated its efficacy in neuropathic pain, but a recent systematic review concluded that lamotrigine provides minimal, if any, therapeutic effects in acute and chronic pain (75). In addition, fatal skin reactions including Stevens-Johnson syndrome and toxic epidermal necrolysis have been reported.

Oxcarbazepine (Trileptal)
Relevance to Physiatry
Oxcarbazepine is a structural analog of carbamazepine. It has been used in the treatment of epilepsy since 1990. There is convincing evidence of its efficacy in newly diagnosed and refractory trigeminal neuralgia (76). In addition, it has emerging merit in areas of neuropathic pain and bipolar disorder (77).

Mechanism of Action and Pharmacokinetics
See Tables 65-7 and 65-8.
Liver microsomes metabolize oxcarbazepine into an active metabolite, 10-monohydroxy metabolite (MHD), which exerts the desired pharmacologic effects. This process minimally induces hepatic CYP450 enzymes. MHD is excreted by the kidneys.

Preparations and Dosing
Oxcarbazepine is available as 150-, 300-, and 600-mg scored tablets, and as a 60 mg/mL suspension. Rapid titration (over 7 to 10 days) from the initial dose (ranging from 75 to 300 mg bid) to a maximum of 1,200 mg bid is recommended. Therapeutic range for trigeminal neuralgia extends from 600 to 1,800 mg per day.

Relevant Side Effects and Drug Interactions
Although a carbamazepine analogue, oxcarbazepine is not associated with serious hematologic toxicity. CNS and GI disturbances are instead reported; between 20% and 25% of patients discontinue oxcarbazepine due to its side effects.

The coadministration of felodipine with CaCBs and/or oral contraceptives should be avoided. The specific mechanism is undetermined, but verapamil (a CaCB) is known to decrease oxcarbazepine concentration by 20% and oxcarbazepine, in turn, decreases CaCB felodipine concentration by 30%. As for contraceptives, oxcarbazepine stimulates their metabolism and, in effect, diminishes their efficacy (78). Antiepileptic drug (AED) coadministration, a regular practice in the setting of epilepsy but not neuropathic pain, also warrants caution because CYP450 enzymes inducers (e.g., carbamazepine, phenobarbital, and phenytoin) decrease mean plasma oxcarbazepine concentrations by 40%.

Phenobarbital (Luminal, Solfoton)
Relevance to Physiatry
There is limited evidence for the effectiveness of phenobarbital in pain management (79). As is true of all barbiturates, human neuropathic pain studies are lacking and sedative properties limit their usage. Phenobarbital thus has a very limited role in neuropathic pain management.

Phenytoin (Dilantin)
Relevance to Physiatry
Besides its well-known anticonvulsant use, phenytoin is also used off-label as a neuropathic pain agent. It was the first anticonvulsant used as an antinociceptive agent and was confirmed (via controlled clinical trials) to be effective in managing trigeminal neuralgia and diabetic neuropathy more than 20 years ago. Recent clinical trials have ironically shown conflicting evidence (80). In addition to this ambiguity, there is a significant potential for medication interactions. In contrast, phenytoin offers the advantage of relative low cost and once a day dosing.

Mechanism of Action and Pharmacokinetics
See Tables 65-7 and 65-8.
Oral absorption of phenytoin in the stomach is slow because it is weakly acidic. The drug is mainly bound to plasma proteins and it readily traverses the blood-brain barrier. Liver hydroxylates phenytoin into nonactive metabolites, which are excreted by the kidneys; this metabolic pathway is progressively saturated as the concentration of phenytoin increases.

Preparations and Dosing
Phenytoin is available as tablets, extended-release capsules, chewable tablets, injectable solution, and oral suspension. Neuropathic pain doses are often less than those used for seizures, but a specific therapeutic range has not yet been defined. Given the complex pharmacokinetics, it is important to monitor serum phenytoin levels because small increases in the dose can produce unexpectedly large increase in plasma concentrations.
CHAPTER 65 | PHARMACOTHERAPY OF DISABILITY

(i.e., glutamate, noradrenaline, and substance P [SP]) (85) to produce its net inhibitory effect on neurons.

Pregabalin has an oral bioavailability of greater than 90%. Concomitant food intake reduces the absorption rate but the total amount absorbed remains constant. Pregabalin does not bind to plasma proteins and nearly the entire dose is excreted unchanged in the urine, with elimination following first-order kinetics.

**Preparations and Dosing**

Pregabalin is available as capsules in the following strengths: 25, 50, 75, 100, 150, 200, 225, and 300 mg. For the treatment of painful diabetic neuropathy, it is generally started at 50 mg tid for 1 week then increased to a maximum of 1,000 mg tid. Therapeutic level of pregabalin for postherpetic neuralgia can be achieved with an initial dose of 75 mg bid or 50 mg tid, and increased to 150 mg bid or 100 mg tid after 1 week; the maximum dose for this purpose is 600 mg/day.

**Relevant Side Effects and Drug Interactions**

Side effects can be classified into three different categories: dose-related toxic effects, true side effects, and idiosyncratic reactions:

- **Toxic effects** generally occur between plasma levels of 20 and 40 \( \mu g/mL \), but there is marked individual variation. The effects include sedation, ataxia, and nystagmus. Ingesting high doses over a prolonged period can result in painful peripheral neuropathy.

- **True side effects** from long-term use include hirsutism, osteomalacia, and hypocalcemia (secondary to interference with vitamin D metabolism), megaloblastic anemia (secondary to interference with vitamin B\(_{12}\) metabolism), and gingival hyperplasia (secondary to interference with fibroblastic activity).

- **Idiosyncratic reactions** include blood dyscrasias and a rare clinical picture, which resembles malignant lymphoma.

Phenytoin should not be prescribed to pregnant women as there are conflicting reports on its teratogenic effects, including the fetal hydantoin syndrome. Cerebellar ataxia can occur at seizure management doses and may interfere with rehabilitation efforts.

Refer to Table 65-10 for drug interactions related to serum protein binding and hepatic metabolism.

**Pregabalin (Lyrica)**

**Relevance to Physiatry**

A second analogue of GABA, pregabalin, is FDA approved for use in treating diabetic neuropathy pain, postherpetic neuralgia pain, and epilepsy. It is also being used off label for other neuropathic pain conditions. Furthermore, there is mounting evidence that this potent “sibling” of gabapentin can help to manage pain due to fibromyalgia and symptoms of anxiety disorder (81–83).

**Mechanism of Action and Pharmacokinetics**

Based on its structure, pregabalin is expected to be a CaCB (84). However, it is also theorized to have other mechanisms of action, including modulating the release of various neurotransmitters (i.e., glutamate, noradrenaline, and substance P [SP]) (85) to produce its net inhibitory effect on neurons.

Pregabalin has an oral bioavailability of greater than 90%. Concomitant food intake reduces the absorption rate but the total amount absorbed remains constant. Pregabalin does not bind to plasma proteins and nearly the entire dose is excreted unchanged in the urine, with elimination following first-order kinetics.

**Preparations and Dosing**

Pregabalin is available as capsules in the following strengths: 25, 50, 75, 100, 150, 200, 225, and 300 mg. For the treatment of painful diabetic neuropathy, it is generally started at 50 mg tid for 1 week then increased to a maximum of 1,000 mg tid. Therapeutic level of pregabalin for postherpetic neuralgia can be achieved with an initial dose of 75 mg bid or 50 mg tid, and increased to 150 mg bid or 100 mg tid after 1 week; the maximum dose for this purpose is 600 mg/day.

**Relevant Side Effects and Drug Interactions**

Dizziness, somnolence, and dry mouth have been frequently reported (86). Complaints of headache, weight gain, edema, blurred vision, and difficulty with concentration are also sometimes reported. Pregabalin is classified as a schedule V controlled substance because it leads to euphoria in selected individuals (87). As is true for gabapentin, pregabalin has no known drug interactions.

**Tiagabine (Gabitril)**

**Relevance to Physiatry**

Tiagabine became available at the turn of the century. Although it demonstrated antihyperalgesic and antinociceptive activity for neuropathic pain in animal models, placebo-controlled trial data are lacking; there are only two small clinical trials to date documenting its beneficial outcome (88–90). New research also proposes that tiagabine is indicated for Stiffman syndrome, bruxism, and tonic spasms in multiple sclerosis; however, more studies are needed to draw definite conclusions (91–93).
**Mechanism of Action and Pharmacokinetics**

See Tables 65-7 and 65-8.

Tiagabine is well absorbed orally; concomitant consumption of food decreases the absorption rate but the fraction absorbed remains constant. It is highly bound to plasma proteins. Metabolism of tiagabine involves oxidation and glucuronidation. Both pathways produce inactive metabolites, which are excreted by the biliary system (major) and the kidneys (minor).

**Preparations and Dosing**

Tiagabine is available in multiple strength tablets (2, 4, 12, 16, and 20 mg). Its neuropathic pain dosing schedule has not been established. The two published clinical trials investigated its efficacy in the range from 4 to 24 mg daily. When used as an anticonvulsant, however, it should be initiated at 4 mg every day, then increased to 8 mg at the beginning of week 2, and increased by 4 to 8 mg at weekly intervals thereafter, until clinical response is achieved or up to 32 mg/day.

**Relevant Side Effects and Drug Interactions**

Side effects are mild. CNS and GI disturbances (e.g., tiredness, somnolence, nausea, and abdominal pain) are the primary reasons that participants withdraw from tiagabine-related clinical trials. Tiagabine is also free of significant drug interactions.

**Topiramate (Topamax)**

**Relevance to Physiatry**

Structurally derived from D-fructose, topiramate is FDA approved for seizure management and migraine prophylaxis, but is also used off-label for neuropathic pain and myoclonic jerks (94). Though several small studies support its efficacy in neuropathic pain, including diabetic neuropathy, larger studies have been disappointing (95,96). Minor studies also found that topiramate provides pain relief in refractory intercostal neuralgia, trigeminal neuralgia, and trigeminal autonomic cephalgias; these findings will likely invite additional future research (97,98). Other possible indications of less physiatric relevance include obesity and bipolar disorder (99).

**Mechanism of Action and Pharmacokinetics**

See Tables 65-7 and 65-8.

Topiramate is rapidly absorbed into the body and does not bind well to plasma proteins (<20%). Only a third of the drug is metabolized by the hepatic CYP450 system into inactive metabolites. The remainder is excreted unchanged in the urine.

**Preparations and Dosing**

Topiramate is available as tablets (25, 50, 100, and 200 mg) and sprinkle capsules (15 and 25 mg). Its recommended daily dose as an adjunctive therapy for seizure prophylaxis is 200 mg bid and gradual titration from an initial daily dose of 25 mg over 8 weeks can reduce adverse cognitive effects. There are no dosing guidelines in neuropathic pain, but topiramate has been used between 200 and 400 mg per day.

**Relevant Side Effects and Drug Interactions**

Topiramate causes two general CNS-related side effects: delayed psychomotor actions (e.g., difficulty concentrating and sluggish speech) and somnolence. Weight loss (secondary to appetite suppression) and paresthesia can develop with chronic intake. The reversible effects discussed above are generally more common at seizure prophylaxis doses (100). Occasional cases of acute myopia and angle closure glaucoma have recently associated with topiramate use as well (101). Last but not least, animal models suggest that it may be teratogenic.

Topiramate is a carbonic anhydrase inhibitor; concomitant use with another carbonic anhydrase inhibitor should be avoided to prevent increased risk of renal stone formation. It also has a mild inductive effect on hepatic CYP450 enzymes and can hence increase metabolism of digoxin and oral contraceptives.

**Valproate (Valproic Acid, Depakene)**

**Relevance to Physiatry**

Valproate has been used as a third-line agent for epilepsy and migraine prophylaxis. It is more recently being studied for use in neuropathic pain, postherpetic neuralgia, and polyneuropathy (102–106). Similar to several other anticonvulsants though, evidence of its efficacy is weak. Previously, it was the subject of a small randomized clinical trial in paraplegic central pain and was not effective (107). It has yield variable results in experimentally induced central pain (108).

**Mechanism of Action and Pharmacokinetics**

See Tables 65-7 and 65-8.

Valproate, a lipid-soluble compound, is rapidly absorbed and becomes tightly protein bound. It is then metabolized in the liver through oxidation and glucuronidation pathways. Active metabolites and a small, unchanged portion are then renally eliminated.

**Preparations and Dosing**

Valproate is widely available in the following forms: caplet, sprinkle capsule, delayed-release tablet, extended-release tablet, syrup, and parenteral preparation. It is often initiated at 250 mg/day and is then titrated to a maximum dose of 1,000 mg bid.

**Side Effects and Drug Interactions**

Nausea, tremors, drowsiness, and weight gain are commonly encountered in the clinical setting. More serious complications include hepatotoxicity among the children population, pancreatitis, and prolonged bleeding time. Frequent monitoring of the previous parameters is recommended. Rare instances of valproate-induced encephalopathy are presumed to be caused by inhibition of ammonia metabolism (109). However, there are recent reports of encephalopathy in the absence of hyperammonia (110).

Valproate is an inhibitor of the oxidation and glucuronidation pathways; it inhibits the metabolism of phenytoin,
carbamazepine, and lamotrigine. It also decreases the clearance of amitriptyline and nortriptyline.

Zonisamide (Zonegran)

Relevance to Physiatry

Zonisamide is FDA approved as an adjunctive therapy in partial seizures. A small, randomized clinical trial in 2005 concluded that it produces statistically insignificant effects when used in diabetic neuropathy (111). There has been renewed interest lately and at least one ongoing research dedicated zonisamide’s analgesic mechanism in animal models (112).

Mechanism of Action and Pharmacokinetics

See Tables 65-7 and 65-8.

Preparations and Dosing

Zonisamide (100 mg) capsules are given daily for the first 2 weeks, after which the dose may be increased to 200 mg/day for at least 2 weeks. It can be increased to 300 and 400 mg per day, with the dose stable for at least 2 weeks to achieve steady state at each level. Evidence from controlled trials as an anticonvulsant suggests that 100 to 600 mg per day doses are effective, but there is no suggestion of increasing efficacy above 400 mg/day.

Relevant Side Effects and Drug Interactions

Zonisamide is contraindicated in sulfonamide allergy. Rare cases of aplastic anemia and agranulocytosis have been reported. It can also cause adverse psychiatric CNS events (depression and psychosis, psychomotor slowing (concentration difficulties and speech/language problems, especially word-finding difficulties), and somnolence and fatigue. Concomitant phenytoin or carbamazepine use increases zonisamide clearance.

LOCAL ANESTHETICS—INJECTABLE ANESTHETICS

Relevance to Physiatry

In the outpatient physiatric setting, injectable and sometimes topical anesthetics are frequently used to provide local anesthesia for a variety of procedures and as a diagnostic tool during intra-articular, soft tissue, and nerve block procedures. They can also be combined with corticosteroids in intra-articular and soft tissue injections to attain immediate pain relief. Injectable anesthetic can also be found as a component of proliferant solutions used in prolotherapy.

Local anesthetics are classified as esters (e.g., procaine) or amides (e.g., bupivacaine and lidocaine) based upon their chemical structure. Amide anesthetics are preferred over ester anesthetics because the latter are associated with a higher incidence of allergic reactions. Cross-sensitivity between the two classes does not occur (113). The amide class is further subdivided according to each drug’s duration of action. Lidocaine, a short-acting injectable anesthetic, is commonly used in percutaneous infiltration anesthesia. Long-acting injectable anesthetics (e.g., bupivacaine) are often reserved for procedures in which a longer degree of postprocedure pain relief is desirable. For example, after a positive lidocaine diagnostic shoulder impingement test, bupivacaine can be added to the corticosteroid to provide prolonged pain relief while corticosteroid gradually takes effect. This section will examine the amide anesthetics in greater detail.

Bupivacaine (Marcaine, Sensorcaine)

Bupivacaine has been widely used for half a century now. Two comparable variants, levobupivacaine and ropivacaine, were subsequently introduced to circumvent the drawbacks of bupivacaine-related side effects, but bupivacaine remains a viable, inexpensive choice (114). Though spinal anesthesia for surgical procedures is considered to be the best indication for bupivacaine, a blend of bupivacaine and corticosteroid can also be injected as part of intra-articular, soft tissue, and some spinal injection procedures to provide long-term relief; candidates for this procedure are typically individuals who responded to prior lidocaine injection (115). Bupivacaine is also used as part of comparative local anesthetic medial branch blocks as a precursor to possible radiofrequency ablation, as is discussed in Chapter 68.

Lidocaine (Xylocaine)

Lidocaine has been used as an anesthetic for over half a century. It has long replaced the first synthetic anesthetic, procaine, as the short-acting injectable anesthetic of choice because of its favorable side-effect profile. Lidocaine is regularly used in outpatient offices for procedures such as abscess drainage and laceration repair, but indications most relevant to physiatry include regional anesthesia for musculoskeletal procedures and nerve blocks.

Mechanism of Action and Pharmacokinetics

All amide anesthetics are speculated to be sodium channel blockers. They selectively inhibit tetrodotoxin-resistant sodium channels on dorsal root ganglia. These axonal structures are involved in generating nociceptive and temperature sensation. Local anesthetics cause a differential neural block—that is, sensory block with minimal loss of motor function (116,117).

Many factors influence an injectable anesthetic’s activity including lipid solubility, level of ionization, molecular size, and vasodilation capacity that are proportional to its potency, onset, and duration of action. Amides undergo extensive hepatic metabolism to become active metabolites—whereas esters are hydrolyzed by plasma enzymes to para-aminobenzoic acid (a potential allergen). Serum levels peak between 5 and 25 minutes postinjection, depending on route of administration and rate of renal excretion.

Preparations and Dosing

Table 65-11 summarizes the dosing guidelines for using lidocaine and bupivacaine in percutaneous infiltration anesthesia.
7.5% Sodium Bicarbonate

Lidocaine leaves a memorable burning sensation upon intradermal and subcutaneous injection because lidocaine solutions are acidic in part due to the preservative contained in multidose bottles. The addition of epinephrine worsens the sting by further decreasing the pH. This unpleasant feeling can be minimized by buffering a lidocaine solution with 7.5% sodium bicarbonate at a 9:1 ratio (e.g., 2 mL sodium bicarbonate added to 20 mL of 1% lidocaine) \(^{(119,120)}\). The mixture should be within 24 hours to avoid risk of contamination due to the unclear effect upon the preservative when the solution has been buffered.

### Relevant Side Effects and Drug Interactions

Amide anesthetics have a wide therapeutic index. Although toxicity can occur and it is dose related, when used alone, lidocaine and bupivacaine have maximal recommended doses of 5 and 2 mg/kg, respectively \(^{(121,122)}\). Spinal injections carry higher likelihood of significant toxicity from possible inadvertent intrathecal administration. Toxicity manifestation can be divided into two categories: local and systemic. Local toxicity encompasses a variety of irreversible neurovascular changes (e.g., paresthesia). Systemic toxicity, albeit uncommon, can affect the CNS followed by the cardiovascular system \(^{(114)}\). The patient may initially complain of drowsiness, tremors, and altered special senses; but, as serum level of the offending agent increases, arrhythmia, seizure, and respiratory/cardiac arrest can develop \(^{(123)}\). Lipid infusion can reportedly treat late-phase systemic toxicity \(^{(124)}\).

Because local anesthetics are CNS depressants, they should not be combined with another CNS depressant. Anesthetics can enhance the action of neuromuscular-blocking agents. Preparations containing epinephrine should not be given concomitantly with MAOIs or TCAs due to risk for severe HTN.

### Other Medications Sometimes Used in Conjunction with Local Anesthetics

**Epinephrine**

Epinephrine counteracts the anesthetic’s vasodilation property and therefore slows systemic absorption rate at the site of injection. Coadministration of this agent with anesthetics can potentiate and prolong analgesic effects, as the effective dose is maintained longer. Epinephrine is also used to dilute high-dose preparations to prevent anesthetics-associated systemic side effects. Concentrations between 2 and 10 \(\mu\)g/mL (i.e., ratio from 1:500,000 and 1:100,000) are generally used.

Despite its benefits, epinephrine may expose patients to additional side effects such as wound infection, tachycardia, and hypertension (HTN). It is generally believed that tissue ischemia can occur when the mixture is injected into body regions, the digits in particular, that have compromised or limited blood supply. The other concern is that epinephrine solutions contain sodium metabisulfite, which can cause an allergic reaction in certain individuals \(^{(118)}\).

### TABLE 65.11 Commonly Used Local Anesthetics

<table>
<thead>
<tr>
<th>Generic (Trade) Name</th>
<th>Applicable Preparations and Concentrations</th>
<th>Onset of Action (Duration)</th>
<th>Usual Dosage (mL): Bursal Injection (A); (IP); (Ish); (SA); (T)(^a)</th>
<th>Usual Dosage (mL): Joint Injection Small (Large)</th>
<th>Dosage: Percutaneous Infiltration (Maximum Amount)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bupivacaine (Marcaine, Sensorcaine)</td>
<td>0.25%, 0.5%, and 0.75%</td>
<td>5 min (2–4 h)</td>
<td>(A) 2½–4½; (IP) 4–4½; (Ish) 2½–4; (SA) 4–6; (T) 4½–9</td>
<td>1–2 mL (2–4 mL)</td>
<td>Up to 70 mL</td>
</tr>
<tr>
<td>Lidocaine (Xylocaine)</td>
<td>0.5%, 1%, 1.5%, and 2%</td>
<td>½–1 min (½ h)</td>
<td>(A) 2½–4½; (IP) 4–4½; (Ish) 2½–4; (SA) 4–6; (T) 4½–9</td>
<td>1–2 (2–4 mL)</td>
<td>Up to 60 mL</td>
</tr>
<tr>
<td>Ropivacaine (Naropin)</td>
<td>0.5%, 0.75, and 1%</td>
<td>5 min (2–4 h)</td>
<td>Not yet described</td>
<td>Not yet described</td>
<td>Up to 100 mL</td>
</tr>
</tbody>
</table>

\(^{a}\)A), anserine bursa; (IP), iliopsoas bursa; (Ish), ischial bursa; (SA), subacromial bursa; (T), trochanteric bursa.
LOCAL ANESTHETICS—TOPICAL ANESTHETICS

In addition to the injectable preparations, local anesthetics are available as a cream and as a transdermal (TD) patch that are sometimes employed for preprocedural soft tissue anesthesia. Topical lidocaine patch has also become a first-line treatment for postherpetic neuralgia and is becoming increasingly used off-label for various neuropathic and musculoskeletal pain conditions (125,126).

Eutectic Mixture of Local Anesthetics

Eutectic mixture of local anesthetics (EMLA) cream is an emulsion composed of 2.5% prilocaine and 2.5% lidocaine droplets. This topical compound has a lasting anesthetic effect for up to 4 hours (127). It can be applied to intact skin an hour prior to painful outpatient procedures. Numerous studies confirmed its efficacy over placebo cream, ethyl chloride, subcutaneous lidocaine, and iontophoresis (128). Although EMLA was deemed less effective than intradermal lidocaine, it is nonetheless preferred by patients (129).

Lidocaine Cream and Patch

Lidocaine cream (LMX) is regarded as an equivalent of EMLA but with faster onset of action (i.e., half an hour vs. 1 hour) (130). The patch form, with a dual mechanism of nociceptive sensation blockade and a mechanical barrier from friction against injured skin, is FDA approved for treatment of postherpetic neuralgia pain. There is also new evidence that lidocaine patches are effective for other neuropathic pain, low back pain, and OA of the knee (125,131,132). Two small, uncontrolled studies reported significant pain relief in reflex sympathetic dystrophy (RSD)/complex regional pain syndrome, stump neuroma pain, intercostal neuralgia, post-thoracotomy pain, and meralgia paresthetica as well (133,134).

The branded lidocaine patch, Lidoderm, contains 700 mg of 5% lidocaine. The recommended guideline to treat postherpetic neuralgia is to apply a maximum of three simultaneous patches on intact skin for 12 hours a day. Clothing may be worn over the application area. Weaker preparations of both the cream and patch can be found OTC.

LET

Both LET solution and gel are formulated with 4% lidocaine, 0.1% epinephrine, and 0.5% tetracaine. The gel has obvious advantages over the liquid form and is widely used half an hour before suturing uncomplicated facial and scalp lacerations (135). LET can also ameliorate pain from infiltrative anesthesia. Doses between 1 and 3 mL are effective in the aforementioned procedures. LET is not effective when used to anesthetize large wounds on the trunk or extremities at recommended doses.

LT Peel

A 7% lidocaine and 7% tetracaine cream dries within half an hour of application and is then peeled off. There is evidence that it is superior to EMLA in adults undergoing cutaneous procedures (136).

Relevant Side Effects and Drug Interactions

Topical agents are associated with potential side effects that are similar to their injectable counterparts except that the TD mode of delivery has a predilection to cause local skin irritation. This generally consists of mild redness and irritation at the application site. Importantly, these local reactions are usually not due to anesthetic allergy. Overall, topical anesthetics are safe. No serious adverse events have been observed in more than 120,000 patch hours in patients who used the patch for up to 8.7 years (137,138).

However, topical anesthetics should be used cautiously in patients who are taking class I antiarrhythmics (e.g., tocainide and mexiletine) as the toxic effects of both drugs are potentially synergistic.

MUSCLE RELAXANTS

Relevance to Physiatry

Unlike antispasticity agents, muscle relaxants are not indicated for true skeletal muscle spasticity. They are instead intended for short-term use in musculoskeletal conditions where muscle “tightness” is one of the primary pain generators. A recent meta-analysis supports this indication and showed that muscle relaxants provided modest short-term pain relief in the treatment for patients with back pain (139). Although there are no published studies comparing the relative efficacy of acetaminophen, NSAIDs and muscle relaxants appear to offer some benefits in patients with non-specific back pain.

Muscle relaxants act by decreasing muscle excitability and thus diminishing tension-induced pain. Unlike antispasticity agents, muscle relaxants offer the advantage of not compromising muscle strength. Unfortunately, sedation limits their application and sometimes obliges physicians to prescribe them for bedtime use only. Cyclobenzaprine, methocarbamol, and carisoprodol are some of the most commonly prescribed muscle relaxants.

Mechanism of Action and Pharmacokinetics

Muscle relaxants are a unique group of medications as they have different mechanisms and pharmacokinetics (Table 65-12). The only common denominator among them is that they all act in some fashion on the CNS, rather than at the muscle fibers level, to interrupt nociceptive signals.

Preparations, Dosing, Relevant Side Effects, and Drug Interactions

See Table 65-12.
This has led to the development of moderate affinity channel blockers (e.g., glycine B) and NR2B selective antagonists, which selectively block peripheral NMDA receptors (147). This new generation of drugs currently displays a better side-effect profile in animal models.

**Ketamine (Ketalar)**

Ketamine hydrochloride is primarily used in veterinary practices as a tranquilizer. It is approved for use in children and individuals with poor health as a general anesthetic (IV or intramuscular [IM]) and as a preoperative sedative (PO or parenteral). It causes a state of consciousness known as dissociative anesthesia. It does not have an official pain indication. There is no predetermined dosing guideline for its analgesic use, but a dose ranging from 6.5 to 13 mg/kg IM is administered for 12 to 25 minutes of surgical anesthesia. One study used 0.4 mg/kg IM doses to treat trigeminal neuralgia (148). There is also a case report on its successful use in complex regional pain syndrome (aka RDS) (149).

Ketamine gel has been reported in a case series involving several different neuropathic pain conditions (150). Ketamine is classified as a schedule III controlled substance because its psychomimetic effects are comparable to snorting PCP.

**N-METHYL-D-ASPARTATE-RECEPTOR ANTAGONISTS**

**Relevance to Physiatry**

Tissue and nerve injury enhance the release of glutamate. This excitatory amino acid then binds to N-methyl-D-aspartate (NMDA) receptors in the spinal cord and, in turn, modulates pain sensation. NMDA receptor antagonists (also known as NMDA glutamatergic antagonists) act to inhibit this pathway. There is some preliminary evidence of success in using the oral form of these agents to treat neuropathic and cancer pain (140). Other potential applications include oral/epidural preemptive analgesia prior to surgery and coadministration with opioids to improve postoperative pain relief (141–144). In addition to their independent analgesic effect, NMDA receptor antagonists are also synergists with opioids and can prevent tolerance to opioids (145).

Ketamine, dextromethorphan, memantine, and amantadine are examples from this category of medication. Two opioids (i.e., methadone and dextropropoxyphene) also possess NMDA antagonistic properties. These agents are discussed below. Other NMDA-receptor antagonists exist, but their use is significantly limited by side effects.

Several NMDA-receptor antagonist clinical trials have been discontinued due to psychomimetic adverse effects as well as ataxia and coordination impairment (146). This has led to the development of moderate affinity channel blockers (e.g., glycine B) and NR2B selective antagonists, which selectively block peripheral NMDA receptors (147). This new generation of drugs currently displays a better side-effect profile in animal models.

**Ketamine (Ketalar)**

Ketamine hydrochloride is primarily used in veterinary practices as a tranquilizer. It is approved for use in children and individuals with poor health as a general anesthetic (IV or intramuscular [IM]) and as a preoperative sedative (PO or parenteral). It causes a state of consciousness known as dissociative anesthesia. It does not have an official pain indication. There is no predetermined dosing guideline for its analgesic use, but a dose ranging from 6.5 to 13 mg/kg IM is administered for 12 to 25 minutes of surgical anesthesia. One study used 0.4 mg/kg IM doses to treat trigeminal neuralgia (148). There is also a case report on its successful use in complex regional pain syndrome (aka RDS) (149). Ketamine gel has been reported in a case series involving several different neuropathic pain conditions (150). Ketamine is classified as a schedule III controlled substance because its psychomimetic effects are comparable to snorting PCP.
Dextromethorphan
Dextromethorphan attenuates acute pain at doses of 30 to 90 mg, divided every 4 to 6 hours (5–10 mg/mL), and reduces analgesic requirements in postoperative patients without major side effects, but it has a suboptimal analgesic effect in treatment of chronic pain (151). There is some preclinical evidence of neuroprotective properties in the setting of perioperative brain injury, amyotrophic lateral sclerosis, and methotrexate neurotoxicity (152).

Amantadine
Amantadine is an antiviral agent with NMDA receptor antagonist properties. It is also used in Parkinson’s disease and traumatic brain injury (TBI). Amantadine was unsuccessful as an agent used to prevent postmastectomy pain neuropathic syndrome (153).

Memantine
Memantine is a moderate affinity NMDA-receptor antagonist. It is indicated for moderate Alzheimer’s disease. There is some evidence from case reports and small, controlled trials on its application in neuropathic pain (154,155).

NONSTERoidal ANTI-INFLAMMATORY DRUGS

Relevance to Physiatry
Oral NSAIDs are frequently used in the outpatient musculoskeletal medicine setting. At high doses, NSAIDs display both anti-inflammatory and analgesic properties, generally without causing sedation. All NSAIDs share the same mechanism of action and overall side-effect profile, but they also have individual characteristics that distinguish them from each other. No one NSAID has been demonstrated as being superior in terms of efficacy to others. Ideally, physiatrists should be familiar with at least one agent from each NSAID class. This will allow the physician to comfortably switch a patient off an agent from one NSAID class to an agent from a different class. This strategy can be emplaced if they do not respond and/or have side effects with use of an NSAID from one class.

Mechanism of Action and Pharmacokinetics
NSAIDs exert their primary effects by inhibiting the synthesis of prostaglandins and other related inflammatory compounds (e.g., thromboxanes and leukotrienes). The four primary properties of NSAIDs are analgesia for mild to moderate pain, anti-inflammatory effects, antipyresis, and reversible platelet inhibition. Anti-inflammatory effects also contribute to analgesia by preventing inflammatory-mediated sensitization of nociceptors. Oral NSAIDs are absorbed in the upper GI tract. A large percent of the drug becomes bound to plasma protein while the unbound portion exerts its pharmacological effects. NSAIDs undergo hepatic metabolism and renal excretion. NSAIDs are available as short-acting and long-acting preparations with a half-life ranging from 30 to 50 hours at steady state. The consequence of accumulating long-acting agents (e.g., oxaprozin [Daypro] and piroxicam [Feldene]) in the human body is unclear, but to date, there are no reports of any additional significant adverse events beyond those typically seen with NSAIDs (156).

Preparations and Dosing
Table 65-13 shows a classification scheme of NSAIDs based on their chemical structure. Additional information on the individual classes is as follows:

• Salicylates: These include aspirin and three nonacetylated salicylates. Compared to other NSAIDs, nonacetylated salicylates are less potent but cause less GI and platelet inhibition. It is unclear if any one particular nonacetylated salicylate in this category is more advantageous than the other two.

• Propionic acids: This is the most popular NSAID class due to the OTC availability of ibuprofen and naproxen, and the direct marketing of the agent to the general public.

• Acetic acids: This class is the most potent and most potentially toxic of all NSAIDs. It includes two drugs that can be administered via IM and parenteral routes (i.e., ketorolac and indomethacin) and two prodrugs (i.e., sulindac and nabumetone) are converted into their active counterparts.

• Fenemates: Meclofenamate and mefenamic acid offers no advantage over other NSAIDs but can cause significant GI toxicity and dysmenorrhea pain, respectively.

• Oxicams: Only piroxicam and meloxicam are currently available in the United States. Piroxicam has convenient once-daily dosing but is associated with severe dermatologic reactions such as exfoliative dermatitis and pemphigus vulgaris. The risk of adverse effects is lower for meloxicam. It was FDA approved for treatment of pain due to OA in 2004.

Relevant Side Effects and Drug Interactions
Individuals regularly using NSAIDs have up to five times the risk of developing GI complications (157). NSAIDs act directly to increase gastric acid secretion and indirectly to inhibit prostaglandin, which protects the GI tract lining. The direct effect varies among NSAIDs and only occurs with oral administration. The direct effect, on the contrary, remains constant regardless of the rate of administration. The elderly and patients with a history of peptic ulcer disease are particularly vulnerable to epigastric discomfort and ulceration. Recent studies indicate that piroxicam and ketorolac confer the highest GI risks at low dose and ibuprofen confers the least (158). The theory behind relative GI risks is that certain NSAIDs undergo extensive biliary excretion of their active metabolites and this, in turn, prolongs mucosal contact.

Simple strategies to decrease risk of GI complications include taking NSAIDs with meals and selecting enteric-coated
### TABLE 65.13 Nonsteroidal Antiinflammatory Drugs (NSAIDs)

<table>
<thead>
<tr>
<th>Drug Name</th>
<th>Dose (Oral in mg)</th>
<th>Other Properties and Side Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>[Salicylates: acetylated]</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aspirin (Ecotrin, Anacin, Bayer)</td>
<td>325–650 q4–6h 81–325 daily cardioprotective</td>
<td>Used: especially for antipyretic and cardioprotective effects Other formulations available: 800 mg controlled release (prescription only) 975 mg enteric-coated (prescription only) suppositories: 120, 200, 300, 600 600 mg combined with narcotics and muscle relaxants Side effects: allergy esp. if triad of nasal polyps, hay fever, asthma; GI toxicity but enteric-coated and buffered forms exist; tinnitus; Reye syndrome in children</td>
</tr>
<tr>
<td><strong>[Salicylates: nonacetylated]</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diflunisal (Dolobid)</td>
<td>500–1,000 load then 250–500 q8–12h</td>
<td>Relatively weak anti-inflammatory effect; lacks antipyretic activity</td>
</tr>
<tr>
<td>Salsalate (Disalcid, Salflex)</td>
<td>3,000 divided q6–12h 1,500 bid</td>
<td>Relatively weak anti-inflammatory effect; no platelet inhibition</td>
</tr>
<tr>
<td>Salicylate combination (Trilisate)</td>
<td>200–300 divided bid–qtid 200–800 tid–qld</td>
<td>Relatively weak anti-inflammatory effect, no ASA-allergic reactions; liquid preparation available (500 mg/5 ml)</td>
</tr>
<tr>
<td><strong>[Propionic acids]</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flurbiprofen (Ansaid)</td>
<td>200–300 divided bid–qtid 200–800 tid–qld</td>
<td>Available in ophthalmic solution (Ocufen; TO form available Inexpensive and widely used; frequent dosing; (OTC): Advil, Motrin IB; Nuprin, Rufen; TO form available</td>
</tr>
<tr>
<td>Ibuprofen (Motrin)</td>
<td>25–75 tid–qgd 250–500 bid (EC-Naprosyn) 375–500 bid</td>
<td>Accrulates if poor renal function</td>
</tr>
<tr>
<td>Ketoprofen (Ordis, Actron)</td>
<td>25–75 tid–qgd 250–500 bid (EC-Naprosyn) 375–500 bid</td>
<td>High-incidence GI side effects; advantage of enteric-coated form?, although expensive; (OTC): Aleve;</td>
</tr>
<tr>
<td>Naproxen (Naprosyn, Aleve)</td>
<td>750–1,000 qd 275–550 bid</td>
<td>Naprelan has Intestinal Protective Drug Absorption System (IPDAS); IR- and SR components</td>
</tr>
<tr>
<td>Naproxen-Na (Naprelan)</td>
<td>1,500 bid</td>
<td>qd or bid dosing</td>
</tr>
<tr>
<td>Ibuprofen (Motrin)</td>
<td>50 bid- tid or 75 bid</td>
<td>LFT monitoring if prolonged use; side effects in up to 20%;</td>
</tr>
<tr>
<td><strong>[Acetic acids]</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diclofenac (Cataflam, Voltaren) (Voltaren-XR)</td>
<td>100–200 qd 200–400 bid-tid 400–1,200 qd</td>
<td>Most potent and toxic NSAID; PR preparation (Indotec); drug of choice in ankylosing spondylitis; indicated in other highly inflammatory conditions (e.g., acute gouty arthritis); Prevents heterotopic ossification s/p total hip replacement (THR) and used for myositis ossificans; dose-related CNS/hematologic side effects in up to 25–50%; GI toxicity</td>
</tr>
<tr>
<td>Indomethacin (Indocin) (Indocin-SR)</td>
<td>25–50 tid 75 qd</td>
<td>Arthrotec = diclofenac (50 or 75 mg) + misoprostol (200 μg) Gastric-sparing properties?</td>
</tr>
<tr>
<td>Ketorolac (Toradol)</td>
<td>15–30 IV/IM q6h or 10 q4–6h pm Lower doses if age &gt;65 or renal dysfunction</td>
<td>FDA-approved only for 5 consecutive days; GI bleeding at higher doses; rapid analgesia with IM form—decrease dose for age ≥ 65, renal dysfunction, weight &lt;110; IV preparation also available</td>
</tr>
<tr>
<td>Nabumetone (Relafen)</td>
<td>1,000 initially then 1,500–2,000 qd or divided bid</td>
<td>qd or bid dosing; nonacidic prodrug that undergoes hepatic biotransformation into active metabolite; preliminary studies suggest that unlike other NSAIDs, no evidence of enterohepatic recirculation of active metabolite—this may be an advantage</td>
</tr>
<tr>
<td>Sulindac (Clinoril)</td>
<td>150–200 bid</td>
<td>Prodrug; possibly renal-sparing because urinary excretion, primarily as biologically inactive forms, may be more GI toxic Frequent dosing; frequent GI toxicity</td>
</tr>
<tr>
<td>Tolmetin (Tolectin)</td>
<td>200–600 tid</td>
<td>Frequent dosing; diarrhea common</td>
</tr>
<tr>
<td><strong>[Fenemates]</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Meclofenamate</td>
<td>50–100 tid–qid or 500 initially, then 250 mg q6h pm for ≤1 wk</td>
<td>Frequent dosing; used for dysmenorrheic pain</td>
</tr>
<tr>
<td>Mefenamic acid (Ponstel)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>[Oxicams]</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Piroxicam (Feldene)</td>
<td>20 qd</td>
<td>qd dosing; accumulation in older adults possibly due to enterohepatic recirculation; dermatologic side effects and cases of serum sickness; PR form (Fexicam)</td>
</tr>
<tr>
<td>Meloxicam (Mobic)</td>
<td>7.5 qd</td>
<td></td>
</tr>
</tbody>
</table>
preparations. Prophylactic medications can be given if the patient has concomitant use of corticosteroids, warfarin, or a history of GI bleeding or peptic ulcer disease (Table 65-14). The classes of available medications include antacids, H₂ blockers, misoprostol, proton pump inhibitors (PPIs), and sucralfate (Table 65-14). Only misoprostol and PPIs are FDA approved for gastric ulcer prevention and H₂ blocker has not been shown to be effective in chronic NSAID users (159). Misoprostol is recommended over PPIs in patients without active Helicobacter pylori infection (160). Ideally, the selected prophylactic medication should be taken for the duration of NSAID therapy in order to provide maximum protection.

Less common GI side effects involve the esophagus, the nonduodenal portion of small bowel, colon, and liver. Esophageal side effects include esophagitis and benign esophageal strictures. Irritable bowel disease can be unmasked while the small bowel and colon can develop ulcers, erosions, and web-like strictures. NSAID enteropathy is not believed to occur via an acid mechanism and is therefore not prevented by antacids, H₂ blockers, or PPIs.

A large study in 2005 found that NSAIDs are the second main cause of drug-induced liver injury. Despite this, hepatotoxicity is rare except in individuals with a history of liver disease (161). Clinically, significant hepatic enzymes elevation does occur with certain NSAIDs, particularly diclofenac. When employing these NSAIDs, LFTs are recommended throughout the course of treatment. There is however currently no established specifically recommended schedule for liver function testing.

In large part due to studies of coxibs (see below), all NSAIDs carry a black box warning that warns against potential cardiovascular and GI side effects. This is discussed in more detail in the coxib section below.

Individuals with preexisting kidney disease or comorbid medical conditions that impair renal blood flow (e.g., CHF and hypovolemia) are prone to acquire NSAID-induced renal toxicity. Acute renal failure, nephrotic syndrome, and interstitial nephritis are examples. It has been suggested but not proven that sulindac is somewhat renal sparing compared to other NSAIDs (162).

In light of all adverse effects discussed above, a significant amount of research has been dedicated to find an alternative to traditional NSAIDs. The effort yielded only one product, the COX-II inhibitor celecoxib, that is still available on the market and it will be further explored in the following section. More options are currently under investigation. They include dual COX and 5-lipoxygenase (5-LOX) inhibitors, synthetic lipoxins, nitric oxide-releasing NSAIDs, and hydrogen sulfide-releasing NSAIDs (163). The mechanism of action for these drugs involves either combining current NSAIDs with a moiety (i.e., nitric oxide and hydrogen sulfide) that releases gastroprotective mediators or targeting new processes of the inflammatory process.

True NSAID allergic reactions occur in 1% of the population and range from simple skin rashes and rhinitis to anaphylaxis. NSAIDs should not be used in patients allergic to aspirin.

**Cyclooxygenase-II Inhibitors**

**Relevance to Physiatry**

COX-II inhibitors (also known as coxibs) represent an alternative to NSAIDs. Coxibs in general were generally thought to confer the advantage of less GI toxicity compared to traditional NSAIDs. This appeal led to them becoming the most frequently prescribed new medication within the first year of being introduced in 1999. However, much has changed since the turn of the century. For example, the two other previously available coxibs, rofecoxib (Vioxx) and valdecoxib (Bextra), were withdrawn from the market in 2004 due to an increased risk of cardiovascular toxicity as is discussed in more detail below. Celecoxib, the only currently available coxib at the time of this writing, is FDA approved for the treatment of pain associated with rheumatoid and OA, and for acute pain. Celecoxib is especially used in those individuals who require oral anti-inflammatories but are on concomitant anticoagulation therapy, are at high risk for GI side effects, or are to undergo certain injection procedures such as fluoroscopically-guided spinal injection procedures or injection procedures into deep joint structures, where the risk of inadvertent bleeding makes the use of traditional NSAIDs potentially dangerous (164,165). There also continues to be interest in coxibs because COX-II expression is implicated in colon cancer and Alzheimer’s disease (165).
Mechanism of Action and Pharmacokinetics

Coxibs reduce prostaglandin synthesis by selectively inhibiting one isoform of cyclooxygenase enzyme over another—namely, COX-II over COX-I. This mechanism contrasts with that of NSAIDs, which inhibit COX-II and COX-I equally. COX-I is constitutively expressed in all human tissues including the GI tract. Only a low level of COX-II is constitutively expressed in brain, kidney, bone, and female reproductive tissues; however, the expression of COX-II can be induced at sites of inflammation. By sparing COX-I, coxibs achieve comparable anti-inflammatory and analgesic effects, with less adverse GI toxicity. The lack of effect on thromboxane synthesis explains the absence of antiplatelet effect.

Food has no significant effect on either peak plasma concentration or absorption at therapeutic doses. Higher doses (≥400 mg bid) should, however, be taken with food to improve absorption.

Preparations and Dosing

Celecoxib is available as 100-mg, 200-mg, and 400-mg capsules. It is currently only approved for the indications as shown in Table 65-15. The lowest effective dose should be used, especially in patients with HTN or CHF because coxibs may cause renal prostaglandin-mediated fluid retention. In patients with moderate hepatic impairment, doses should be decreased by approximately 50%.

Relevant Side Effects and Drug Interactions

Clinical trials found that coxibs caused less gastropathy than traditional NSAIDs; the most notable trials included the VIGOR (Vioxx GI Outcomes Research), TARGET (Therapeutic Arthritis Research and GI Event Trial), and CLASS studies (Celebrex Long-term Arthritis Safety Study) (166–168). There is however growing evidence that coxibs impair healing of damaged gastric mucosa and may negatively affect both the small and the large intestines (169–171). These findings have led to further questioning as to whether there really is an advantage of less GI toxicity from coxibs over traditional NSAIDs (163).

Coxibs can cause a dose-dependent risk of adverse cardiovascular events, including myocardial infarction and stroke (172,173). Rofecoxib was withdrawn upon completion of the 3-year APPROVe study, which showed that daily intake of 25 mg of this coxib doubled the risk of thrombotic events (174). It was also shown that doubling the dose of celecoxib increased adverse cardiovascular events by (risk ratio, 3.4; 95% confidence interval, 1.5 to 7.9) above that of the 200 mg qd Celebrex dose (2.6; 95% CI, 1.1 to 6.1) (175,177). The pathophysiology behind this effect is that coxibs inhibit vascular endothelium production of prostaglandin I₂, a lipid that counteracts thromboxane A₂ to prevent platelet aggregation leading to atherosclerosis, and cause a similar blood pressure (BP) elevation as NSAIDs (176). It is not known if cardioprotective aspirin or low-salt diet can mitigate coxibs’ cardiovascular risks (177).

This issue of increased cardiovascular risk led to a black box warning that is now part of the package inserts of celecoxib and all NSAIDs as follows.

Cardiovascular Risk

- NSAIDs and celecoxib may cause an increased risk of serious cardiovascular thrombotic events, myocardial infarction, and stroke, which can be fatal. This risk may increase with duration of use. Patients with cardiovascular disease or risk factors for cardiovascular disease may be at greater risk.
- NSAIDs and celecoxib are contraindicated for the treatment of perioperative pain in the setting of coronary artery bypass graft (CABG) surgery.
- GI risk.
- NSAIDs and celecoxib cause an increased risk of serious GI adverse events including bleeding, ulceration, and perforation of the stomach or intestines, which can be fatal. These events can occur at any time during use and without warning symptoms.
- Elderly patients are at greater risk for serious GI events.

Traditional NSAIDs and coxibs have equal potential to cause renal toxicity (e.g., fluid retention leading to edema, renal HTN, interstitial nephritis, and papillary necrosis) (178,179). Elderly patients with renal insufficiency or hepatic failure and individuals at risk for renal failure should avoid coxibs (180).

Mean INR can increase up to 10% for a subset of patients receiving both warfarin and coxibs, and there have been infrequent reports of increased INR associated with bleeding events (159). This notion however was subsequently challenged by the results of a randomized, controlled, crossover trial to assess the effect on INR of celecoxib in 15 patients who were receiving warfarin therapy (181). Nonetheless, patients on anticoagulation therapy should have their INRs monitored when a coxib is initiated or its dose changed.

---

**Table 65.15** Celecoxib: Recommended Doses for Approved Indications

<table>
<thead>
<tr>
<th>Indication</th>
<th>Celecoxib (Celebrex): Capsules (mg) 50; 100; 200</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acute pain or primary dysmenorrhea</td>
<td>400 mg first dose, then another 200 mg on day 1 pm, then 200 mg bid pm thereafter</td>
</tr>
<tr>
<td>Ankylosing spondylitis</td>
<td>200 mg qd or 100 mg bid, then 400 mg qd if no effect after 6 wk</td>
</tr>
<tr>
<td>Osteoarthritis</td>
<td>200 mg qd or 100 mg bid</td>
</tr>
<tr>
<td>Familial adenomatous polyposis</td>
<td>400 mg bid</td>
</tr>
<tr>
<td>Rheumatoid arthritis—adults</td>
<td>100–200 mg bid</td>
</tr>
<tr>
<td>Rheumatoid arthritis—juvenile (≥2 years old)</td>
<td>≥10 kg to ≤25 kg: 50 mg bid; &gt;25 kg: 100 mg bid</td>
</tr>
</tbody>
</table>
Tizanidine (Zanaflex)  
Relevance to Physiatry
Tizanidine was developed as an antispasticity agent and is approved for use in spasticity due to multiple sclerosis or SCI. It has also been studied for acute low back pain accompanied by muscle "spasm" and neuropathic pain conditions, including trigeminal neuralgia and phantom pain syndrome (186–188). An open-label study in humans concluded that tizanidine may be an effective treatment for pain associated with idiopathic peripheral neuropathy, and another animal study showed that it relieved thermal hyperalgesia in rats with induced neuropathic pain (189,190). However, there are no large, randomized clinical trials to date on its use in neuropathic pain. Other indications that are not as applicable to physiatry include its use along with long-acting NSAIDs for detoxification of rebound headache, chronic tension headache, and narcotic withdrawal (191–193).

Mechanism of Action and Pharmacokinetics
Tizanidine is a centrally acting α₂-adrenergic agonist which presumably reduces spasticity by increasing presynaptic motor neuron inhibition. Following oral administration, 95% of the dose undergoes first-pass hepatic metabolism to inactive metabolites and peak plasma level is reached in 1.5 hours. Concomitant food intake decreases peak plasma concentration and increases the time it takes to reach peak level (194).

Preparations and Dosing
Tizanidine is available in 2-, 4-, and 6-mg scored tablets. Dosing for spasticity is accomplished by gradually increasing a starting dose of 4 mg by 2 to 4 mg until a daily maximum of 36 mg divided tid is reached; the desired therapeutic effect is achieved or dose-limiting side effects occur. Although the maximum total daily dose is generally denoted as 36 mg, clinical experience beyond 24 mg is limited. Dosing guidelines for myofascial and neuropathic pain have not yet been established as these uses are considered off-label.

Side Effects and Medication Interactions
Hypotension, dry mouth, CNS depression, and muscle weakness are commonly associated with α₂-adrenergic antagonism. Muscle weakness associated with tizanidine is comparatively less than with other antispasticity agents (195).

The most clinically significant side effect, though rare, is hepatotoxicity. In controlled clinical studies, 5% of participants had serum transaminase elevated three times that of normal (or twice if baseline levels were elevated); most cases resolved upon medication withdrawal but some individuals developed nausea, vomiting, anorexia, and jaundice. Liver function testing is recommended during the first 6 months (i.e., baseline, 1, 3, and 6 months) of initiating tizanidine therapy and then periodically monitored. Tizanidine should be used with caution in patients with impaired hepatic function.

Oral contraceptives lower tizanidine clearance (196). There are also limited case reports of tizanidine interaction with fluvoxamine, lisinopril, and ciprofloxacin (197–199).

OPioid Analgesics  
Relevance to Physiatry
Narcotic analgesics are often referred to as opioid or opiate analgesics, as some are opium derived. They are indicated for moderate to severe acute and ongoing malignant pain, especially that which is of a nociceptive quality (i.e., generally
dull or aching). In contrast, opiate analgesics are considered second-line treatments for various neuropathic pain syndromes such as phantom leg pain, diabetic neuropathy, and postherpetic neuralgia (200–202). These indications are partly adapted from the World Health Organization (WHO) three-step ladder that guides analgesic treatment (203). The first step consists of nonopioid analgesics and adjuvants. Weak opioids (e.g., codeine) comprise the second step of pain management when the first-step options have failed. Step three consists of strong opioids (e.g., morphine) plus or minus the first-step options.

The lack of end-organ toxicity (especially GI, hepatic, and renal) with pure narcotic analgesics, other than meperidine, makes them an appealing choice for chronic therapy. However, the potential for serious adverse events (e.g., respiratory depression) represents a drawback (204).

More common side effects associated with opioid analgesics include nausea, vomiting, constipation, sedation, euphoria, tolerance, and physical/psychological dependence. There may be minor variations to this side-effect profile for individual agents (i.e., some have more pronounced side effects than others at equivalent doses) (205,206). Since all pure opioid agonists have equianalgesic doses with each other, selection of a specific agent is based upon the desired route of administration, duration of action, and desired side effect.

Other physiatric applications for opioids include diarrhea treatment and cough suppression—the latter is particularly pertinent for patients with nonproductive cough that is interfering with sleep and, thus, their rehabilitation. The recent discovery of opioid receptors beyond the CNS has led to much interest in “peripheral opioid analgesia”; there is now evidence that IM, intra-articular, and IV injection of low-dosed opioid provide localized, peripheral analgesia, and anti-inflammation (207–211).

**Mechanism of Action and Pharmacokinetics**

This class of medication mimics the action of opioids naturally produced by the body. Endorphins, enkephalins, and dynorphins are examples of these endogenous opioids. All opioids exert their effects by binding, with varying degrees of affinity, to three primary receptor types: mu (µ), kappa (κ), and delta (δ). Each receptor type has a unique CNS distribution and results in different physiologic responses, via variable biochemical pathways. Receptors have also been identified in immune cells, peripheral sensory nerves, and joints; these sites account for the peripheral analgesic and anti-inflammatory effects of opioids (212,213). The unifying property among all opioid receptors is that they are coupled to a G-protein, which inhibits adenyl cyclase. An activated receptor acts via its specific biochemical pathway to inhibit neuronal excitability and thus, blocks pain impulse transmission.

Opioid analgesics can be categorized based on their affinity for a receptor and their intrinsic activity (i.e., amount of receptor stimulation they can produce). Morphine and methadone are designated as full agonists because they have high affinity for receptors and produce strong analgesia. Partial agonists, such as codeine, have lower affinity and are thus less potent than full agonists. Pentazocine is an example of a mixed agonist/antagonist that can activate unoccupied opioid receptors while blocking occupied ones. Antagonists (e.g., naloxone) are effective in reversing the effects of a full opioid agonist.

Slightly basic opioids are generally well absorbed in the small intestine. Short-acting agents demonstrate maximal effects between 30 and 60 minutes of administration and have durations of approximately 4 hours. Long-acting or SR agents achieve peak effects within 2 to 24 hours and last for 12 to 72 hours—the first value correlates with oral administration and the second value correlates with TD application. Many opioids undergo first-pass metabolism in the liver. Hepatic conjugation is the primary route of metabolism for most opioids but metabolism can also occur in the kidneys, lungs, and CNS. Active and inactive metabolites are excreted in urine and/or bile.

**Preparations and Dosing**

Opioid analgesics can be administered orally, intramuscularly, intravenously, subcutaneously, intrapinally, intranasally (as naloxone), rectally (as hydromorphone), transdermally (as fentanyl patch), and transmucosally (as fentanyl buccal tablet). The preferred route of administration and duration of action (i.e., long acting versus short acting) will vary for a given patient. Short-acting agents can be used to manage acute pain syndromes and episodes of breakthrough pain. Long-acting agents are generally more convenient for patients with chronic conditions and their sustained effects often help prevent pain-related nocturnal awakening.

Opioid preparations most relevant to physiatry, along with their usual dosage ranges and relative potencies, are shown in Table 65-17; although partial agonists and mixed agonist/antagonist are listed on the table, they have limited application. Required oral doses are greater than parenteral ones because of first-pass effect following oral administration. The management of neuropathic pain usually requires higher doses than for nociceptive pain.

Patients requiring chronic opioid therapy are often initiated on a short-acting agent that is titrated over several weeks to achieve adequate analgesia, and then converted to an equivalent dose of the long-acting agent (214). Doses can be titrated as high as necessary to relieve pain as long as they are tolerated well. Studies have shown that daily doses up to 120 mg of morphine equivalent can provide pain relief and improve sleep, without impairing cognitive function (202,214,215). Opioid rotation, which involves replacing one opioid with one that has incomplete crosstolerance for the original, is a technique that is used in an attempt to avoid tolerance buildup and to overcome dose-limiting side effects (216,217). An individual is considered to be tolerant to opioid therapy when he/she is consuming more than 60 mg oral morphine daily, more than 25 mcg TD fentanyl hourly, more than 30 mg oxycodone daily, more than 8 mg hydromorphone daily, or equivalent dose of another opioid for more than 1 week (218).

WHO recommends that all opioid therapies follow a fixed-interval dosing schedule, rather than being dispensed...
### TABLE 65.17 Narcotic Analgesics

<table>
<thead>
<tr>
<th>[Narcotic Class]</th>
<th>Subclass</th>
<th>Generic (Trade) Name</th>
<th>Usual Dosage Range (mg) (Time = h)</th>
<th>Relative Potency PO (IM) [Other]</th>
</tr>
</thead>
<tbody>
<tr>
<td>[AGONIST]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Codeine</td>
<td></td>
<td>15–60 PO/IM/IV/SC q4–6h</td>
<td>130 (75) [75 IV/SC]</td>
<td></td>
</tr>
<tr>
<td>Fentanyl-TD (Duragesic, Actiq, IONSYS)</td>
<td>1 patch q72h</td>
<td>Unknown (0.1) [0.1 IV/SC]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TD patch: see below</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Morphine</td>
<td></td>
<td>8–22 40–68 PO/IV q4–6h</td>
<td>7.5 (1.5) [1.5 IV/SC]</td>
<td></td>
</tr>
<tr>
<td>Meperidine</td>
<td></td>
<td>0.5–2 IM/SC or slow IV q4–6h</td>
<td>300 (75) [75 IV/SC]</td>
<td></td>
</tr>
<tr>
<td>Meperidine (Demerol, Pethidine)</td>
<td>1–1.8 mg/kg PO/IM/SC q3–4h; slow IV q3–4h</td>
<td>30 (10) [10 IV/SC]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hydromorphone (Dilaudid)</td>
<td>2–4 PO q4–6h; 3 PR q6–8h</td>
<td>5 mg/5 mL liquid PO q6h</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Dilaudid-5)</td>
<td></td>
<td>0.5–2 IM/SC or slow IV q4–6h</td>
<td>30 (10) [10 IV/SC]</td>
<td></td>
</tr>
<tr>
<td>Meperidine</td>
<td></td>
<td>1–1.8 mg/kg PO/IM/SC q3–4h; slow IV q3–4h</td>
<td>30 (10) [10 IV/SC]</td>
<td></td>
</tr>
<tr>
<td>Morphine sulfate</td>
<td></td>
<td>2.5–10 mg IM/SC/PO q3–4</td>
<td>5–15 (2.5–10) [2.5–10 IV/SC]</td>
<td></td>
</tr>
<tr>
<td>SR tabs (MS Contin; Oramorph SR)</td>
<td>30 PO q8–12h</td>
<td>30 (10) [10 IV/SC]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SR caps (Kadian)</td>
<td></td>
<td>20 PO q12–24h</td>
<td>30 (10) [10 IV/SC]</td>
<td></td>
</tr>
<tr>
<td>SR caps (Avinza)</td>
<td></td>
<td>30 PO</td>
<td>30 (10) [10 IV/SC]</td>
<td></td>
</tr>
<tr>
<td>Oral solution (Roxanol)</td>
<td>Various concentrations: 10–30 q4h</td>
<td>30 (10) [10 IV/SC]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Immediate release (MS IR)</td>
<td>10–30 q4h</td>
<td>30 (10) [10 IV/SC]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oxycodeone</td>
<td></td>
<td>5–30 PO q4h</td>
<td>20 (N/A) [N/A IV/SC]</td>
<td></td>
</tr>
<tr>
<td>IR tabs (OxyIR, OxyFAST)</td>
<td>5–30 PO q4h</td>
<td>10–40 PO q12h</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IR tabs (Percolone)</td>
<td>10–30 PO q4h</td>
<td>10–40 PO q12h</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Controlled release (Kadian)</td>
<td>20 PO q12–24h</td>
<td>10–40 PO q12h</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extended release (Avinza)</td>
<td>30 PO initially</td>
<td>10–40 PO q12h</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sustained release (OxyContin)</td>
<td>10–40 PO q12h</td>
<td>10–40 PO q12h</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Propoxyphene</td>
<td></td>
<td>65–100 PO q4h prn</td>
<td>130</td>
<td></td>
</tr>
<tr>
<td>(Darvon Pulvules)</td>
<td></td>
<td>100 PO q8h prn</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Darvon-N)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[PARTIAL AGONISTS]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Buprenorphine (Buprenex)</td>
<td>0.3–0.6 IV/IM q6h</td>
<td>N/A (0.3–0.4) [0.3–0.4 IV/SC]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[MIXED AGONIST-ANTAGONISTS]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Butorphanol tartrate nasal spray (Stadol NS)</td>
<td>1 intranasal spray (per nostril) q3–4h</td>
<td>N/A [2] [2 IV/SC]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pentazocine</td>
<td></td>
<td>1 tab PO q3–4h prn</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>(Talwin)</td>
<td></td>
<td>30 IV/IM q3–4h prn</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>(Talwin NX) (50 mg pentazocine/0.5 naloxone)</td>
<td>1 tab PO q3–4h prn</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[ANALGESIC COMBINATIONS]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Narcotic/acetaminophen</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Propoxyphene/acetaminophen (Darvocet)</td>
<td>1 tab PO q4h</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(IV – 100 = 100/6501, A500 = 100/500)</td>
<td>1 tab PO q4h</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(IV – 50 = 50/325)</td>
<td></td>
<td>1 tab PO q4h</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>Hydrocodone/acetaminophen (Lortab)</td>
<td>1 tab PO q4–6h</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2.5/500, 5/500, 7.5/500, 10/500)</td>
<td>1 tab PO q4–6h</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anexsia (hydrocodone/acetaminophen)</td>
<td>1 tab PO q4–6h</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(5/25, 5/50, 7.5/325, 7.5/500)</td>
<td>1 tab PO q4–6h</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hydrocodone/acetaminophen (Lorcet)</td>
<td>1 tab PO q4–6h</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(5/500, 7.5/650, 10/650)</td>
<td>1 tab PO q4–6h</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(continued)
as needed (prn) (203). This strategy helps maintain a stable serum medication level and, in turn, minimizes side effects experienced at peak levels and breakthrough pain at trough levels. Another recommendation is for health care providers to recognize individual variability in opioid metabolism and pain tolerance. Therapy must therefore be individualized to maximize analgesia while minimizing side effects. Other guidelines within this document are designed to help optimize opioid analgesics use and minimize the side effects.

**INDIVIDUAL AGENTS**

**Partial Agonists**

**Codeine**

Codeine is a natural opioid present in opium that can be used as an analgesic, antitussive, or antidiarrheal agent. It possesses less than one-seventh the analgesic potency of morphine and is often either combined with acetaminophen or used concomitantly with NSAIDs to manage mild to moderate pain (219). It has a relative lack of respiratory depression and a low abuse potential (due to its relative lack of euphoria)—thus posing a lower abuse potential. Codeine has a high oral bioavailability because its phenyl ring shields the molecule from first-pass metabolism. Following administration, codeine is demethylated to morphine, an active metabolite and an inactive metabolite, norcodeine; these metabolites are then hepatically conjugated and most are renally excreted. Individuals unable to convert codeine to morphine would experience negligible analgesic effects.

Potential adverse effects associated with codeine are typical of narcotic analgesics in general. Codeine is regulated under the Controlled Substances Act. Codeine is considered to be unsafe when consumed at high doses or for prolonged time during pregnancy. Some codeine-containing preparations may contain sodium metabisulfite, a potentially allergenic preservative. SR codeine, codeine Contin, is available outside the United States for chronic pain management (220).

**Propoxyphene (Darvon, Darvon Pulvules, Dolene)**

This narcotic analgesic is a synthetic structural analogue of methadone. It has two isomers: one has antitussive effects while the other has one half to two thirds of the analgesic potency of codeine (219). The latter isomer offers an alternative to individuals who cannot metabolize codeine and is often combined with acetaminophen or aspirin to manage mild to moderate pain. Propoxyphene, unlike codeine, undergoes extensive first-pass metabolism and is converted into norpropoxyphene, which has local anesthetic effects and may accumulate with repeated dosing. Toxic levels of the active metabolite from chronic administration or acute overdose can lead to arrhythmias, cardiogenic shock, respiratory depression, hallucination, seizures, and death. Contrary to the general rule that opioid agonists do not have an analgesic ceiling, propoxyphene doses should not exceed 390 mg/day. The concomitant intake of propoxyphene with another CNS depressant should be avoided, and caution should be exercised when prescribing it to patients with preexisting hepatic dysfunction because propoxyphene therapy has been associated with hepatotoxicity, including rare instances of reversible cholestatic jaundice. Several nursing home specialists advise against the use of propoxyphene in geriatric population (221). In practice, however, some physicians may select propoxyphene when it the only viable opioid option.
**Full Agonists**

**Fentanyl-TD (Duragesic), Fentanyl-Transmucosal (Actiq, Fentora)**

Fentanyl is a potent (i.e., >80 times of morphine), short-acting opioid with therapeutic uses including analgesia, preprocedural anxiolysis, sedation, and supplemental anesthesia. It is available as an injectable preparation, a nebulized preparation, a TD patch, and a transmucosal lozenge/tablet. Management of terminally ill patients with dypsnea and is an off-label use of nebulized fentanyl (222).

A TD delivery system for fentanyl is possible because it has a low molecular weight and is highly lipophilic, which enables the drug to be absorbed through the skin and subsequently distributed throughout the body. The fentanyl patch, with an average duration of 72 hours, is used to manage chronic pain that requires regular administration of narcotic analgesics. The advantages of TD fentanyl include lower incidences of constipation, nausea, and drowsiness (223). Cutaneous reactions are limited to localized dermatitis, but systemic adverse effects may occur. The patch cannot be used in acute pain syndromes because initial absorption is delayed for 17 to 48 hours, and it should not be used due to the risk of significant hypoventilation in narcotic-naive patients. There is, however, a recent trend of employing a patient-controlled iontophoresis TD system to manage acute postoperative pain within inpatient settings (224–226).

When initiating a fentanyl patch therapy for narcotic-naive patients, one should start with the lowest available dose (currently 12.5 mcg/hour). There is a wide variability in absorption among patients and titration can be accomplished by combining various available patch strengths (i.e., 25, 50, 75, and 100 mcg/h) (227). As for patients who are accustomed to opioid therapy, the initial dose is estimated from previous equianalgesic morphine doses. Fever, diaphoresis, cachexia, morbid obesity, and ascites can all affect the rate of TD absorption and dosage should hence be adjusted accordingly (228). Physiatrists involved in pain management should also become familiar with dosage titration, discontinuation strategies, and drug interactions for the fentanyl patch.

With an average onset of action in 5 minutes, transmucosal delivery provides a rapid route of absorption into the bloodstream. The lozenge and the buccal tablet are only FDA approved for treatment of breakthrough cancer pain in patients who are tolerant to their current opioid therapies, but various off-label uses have been reported (229–232). Dosing should be individualized to achieve adequate analgesia with tolerable side effects. If the pain is not relieved following transmucosal administration, a second dose of lozenge and buccal tablet can be administered 15 and 30 minutes after completion of the first, respectively (233). Patients should not exceed four units per day. Single doses between 100 and 800 mcg are generally well tolerated, causing mild-to-moderate opioid-associated side effects (234).

**Hydromorphone (Dilaudid)**

This semisynthetic opioid is derived from morphine. These two compounds have similar pharmacokinetic profiles, except that hydromorphone has a shorter duration of action but five times the analgesic efficacy. An open-labeled study found that hydromorphone causes less nausea, emesis, and constipation (235,236). Hydromorphone may be a valuable alternative for patients with “pseudoallergy” (e.g., flushing and itching due to histamine release) to morphine in managing moderate to severe chronic pain because it appears to induce minimal histamine release (237). It is also preferred for patients with renal failure because its metabolite is nontoxic (238). Finally, hydromorphone’s superior water solubility allows high doses to be dissolved into solution and delivered to opioid-tolerant patients.

Several hydromorphone-containing products are FDA approved as antitussives and a nebulized form is available for off-label use in terminally ill patients with dypsnea (222,239). Although hydromorph Contin remains available in Canada, a similar extended-release formulation was withdrawn from the U.S. market in 2005 due to a high risk of overdose when consumed with alcohol.

**Meperidine (Demerol)**

Once a popular fast-acting agent used for management of moderate-to-severe postoperative pain, meperidine is of limited use in physiatry for three major reasons. First, it has only one-tenth morphine’s potency yet the two have similar addiction and physical dependence risk (240). Second, its short duration of action necessitates frequent meperidine dosing and this practice cycles through peaks and troughs in serum concentrations. Third, meperidine is metabolized to normeperidine, a toxic metabolite that causes CNS hyperexcitability that can manifest as seizure, anxiety, tremors, and myoclonus (241). Patients with impaired renal function are obviously susceptible to meperidine toxicity, but healthy individuals are also vulnerable because normeperidine accumulates with use greater than several days. Finally, naloxone, an opioid antagonist, cannot prevent or reverse overdose and it may even worsen CNS toxicity (33).

**Methadone**

Methadone, along with dextropropoxyphene, is both an agonist to μ-receptors and a weak, noncompetitive antagonist to NMDA receptors. Discovery of this unique dual action has led to increased use of methadone for the treatment of neuropathic pain, including phantom limb pain, and burn pain in the recent years (242–244). Methadone is also becoming increasingly used as a second-line agent in treating cancer pain unresponsive to conventional opioids (44). Traditional uses include management of severe, chronic pain, and suppression of withdrawal symptoms from heroin and morphine. For the latter reason, there is social stigma associated with methadone.

This schedule II controlled opioid has several advantages including its lack of active metabolites, its high lipid solubility, its excellent absorption after oral and rectal delivery, and its availability at low cost (45). Despite these advantages, the conversion schedule (i.e., dose and time course) between methadone and other opioids remains indeterminate. Confusion
arises from the fact that parenteral methadone is proposed, but not proven to be equianalgesic to morphine and traditional opioid conversion charts have underestimated the potency of methadone, leading to safety concerns (245). Another drawback to methadone is its long half-life because unpredictable drug accumulation at the beginning of a regimen carries a risk of respiratory depression. Myoclonus and electrocardiographic changes (e.g., QT prolongation and torsade de pointes) have been reported as well (246). Methadone is also increasingly being abused by recreational drug users and has been associated nationally with increase in overdoses and deaths. Because it does not lead to quick onset or significant euphoria, its substance abuse potential had been considered as low. Unfortunately, it can be hours before a user feels any effect, and by the time they have actually overdosed, no one is with them.

**Oxycodone (OxyIR, OxyContin, Eth-Oxydose)**

Oxycodone is a morphine derivative that is available as a generic short-acting form, a short-acting IR form (OxyIR), a controlled-release form (OxyContin), as well as several combination-analgesic products. Pure oxycodone products offer the advantage of not being limited by the potential for toxicity, which is associated with increasing doses of aspirin and acetaminophen. Its efficacy in managing neuropathic and somatic pain has been established, but OxyContin is only approved for management of chronic, moderate-to-severe pain (247).

OxyContin has high oral availability due to minimal first-pass metabolism. Easy access has historically led to its widespread abuse in the United States (248). This in fact has led the FDA to place a black box warning on the product, categorize it as a schedule II controlled substance, and order the discontinuation of the 160-mg tablets. Despite these problems with diversion, OxyContin is still a valuable medication with a fast onset and long duration of action—the apparent dichotomy of prompt yet sustained analgesia can be explained by the AcroContin delivery system, whereby oxycodone is released relatively quickly after ingestion and continues to be steadily released over the subsequent 12 hours. When dosed around the clock, it provides sustained serum levels through the night and potentially minimizes narcotic side effects often associated with peak serum drug levels. A relatively short half-life also allows for OxyContin to reach steady state in a short time period, thereby achieving its full analgesic potential within a day or two of treatment initiation. An additional benefit is that, unlike MS Contin, its absorption is independent of pH; this allows patients to take the opioid with or without food.

The short-acting IR form, OxyIR, can be used as a premedication prior to physical therapy for patients whose pain level interferes with meaningful participation in therapeutic exercises. One-fourth to one-third the 12-hour dose of OxyContin can also be prescribed to treat breakthrough pain. If more than 2 rescue doses are needed during any 24-hour period, the OxyContin regimen should be titrated.

The liver extensively metabolizes oxycodone into oxymorphone, which has even greater analgesic potency, and noroxycodone, a weak analgesic. Neither metabolite causes end-organ toxicity. Efficacy and adverse events are believed to be similar for both OxyIR and OxyContin (249,250). Compared to morphine, they cause less nausea and vomiting but more constipation (251).

**Partial Agonists**

**Buprenorphine (Buprenex, Subutex, Suboxone)**

Buprenorphine, a mild-to-moderate analgesic, has been available via parenteral administration for many years. This partial agonist, with agonistic activity at κ- and δ-receptors, became available to the United States in a sublingual form (Subutex) in 2002, and it was FDA approved for maintenance of opioid addiction (252). A TD delivery system for this medication can also be found in European markets. A recent review of various clinical studies concluded that oral, IV, intrathecal, and TD buprenorphine are all efficacious in managing neuropathic pain. Further research is however, required to develop guidelines for different syndromes (253). IV and oral administrations, additionally, have sustained antihyperalgesia effect in inflammatory models (254).

Slow dissociation from the μ-receptor makes buprenorphine a long-acting agent. Fewer withdrawal signs occur because buprenorphine minimally affects GI motility and sphincter tone. It does not have an analgesic ceiling, but it can cause respiratory depression at high doses and this cannot be readily reversed with naloxone (255,256).

**MIXED AGONISTS-ANTAGONISTS**

These analgesics are used for moderate to severe pain, but do not offer any superiority over opioid agonists. They have a lower respiratory depression risk compared to traditional agonists, but this benefit is minimal given that all agonist users build tolerance to respiratory depression over time. There are several disadvantages inherent to class, including an analgesic ceiling and the potential to precipitate a withdrawal syndrome among previous opioid users. It is nonetheless important to be somewhat familiar with several key members.

The prototype agent, pentazocine (Talwin, Talacen), is an antagonist at μ-receptors and an agonist at κ- and δ-opioid receptors. It is approximately equianalgesic to codeine on a milligram-per-milligram basis. Several different pentazocine combinations are available as shown in Table 65-17. It is unclear how pentazocine-containing medications can be best used in pain management. A pentazocine-methylphenidate combination—known by various street terms including crack-ers, poor man’s heroin, and T’s and rits—is subject to illicit use because the compound produces an effect similar to that of heroin mixed with cocaine. A new formulation, Talwin NX, blends pentazocine with naloxone to maintain analgesic effects but minimizes abuse.

Butorphanol tartrate (Stadol NS) is a mixed agonist-antagonist (i.e., agonist at κ-opioid receptors and a mixed agonist-antagonist at μ-opioid receptors) that offers the flexibility of intranasal administration. Of note, widespread abuse of the
nasal spray led to its classification as a schedule IV controlled substance in the late 1990s. Butofuranol has not been widely used by physiatrists and there remains scant literature on its use in musculoskeletal pain (257,258). Instead, it has a niche in the general surgical setting, where it is used as a preoperative/postoperative sedative and analgesic, a supplement to balanced anesthesia, a conscious sedative, and a postanesthesia shaking suppressor. Additional applications include analgesia during labor, relief of moderate postpartum pain, and as a treatment for migraine headache.

**Relevant Side Effects and Drug Interactions**

Before discussing opioid-related side effects, it is important to clarify three terms: tolerance, addiction, and dependence. *Tolerance* can be a double-edged sword. It potentially negatively affects treatment when an increasing amount of a drug is needed to produce a given therapeutic effect. Tolerance to side effects (other than constipation), however, is of obvious benefit. *Dependence* is the onset of withdrawal symptoms when a drug is abruptly removed. *Addiction* is the habitual use of a substance to achieve a certain effect—usually euphoria—that the patient perceives as pleasurable. Confusion of these terms has led to bias against the use of narcotic analgesics, particularly for nonmalignant pain. Fear of patient addiction is also the main reason why physicians tend to underprescribe narcotics (259). There is controversy as to whether psychological addiction actually develops in patients with chronic pain without a past history of substance abuse (260–262).

Health care providers may need to explain the concept of tolerance to patients and their families to quell the anxiety associated with increased narcotic requirement that often occurs after approximately 1 month of treatment. Tolerance begins after the first dose but does not become clinically apparent until the second or third week, and it generally lasts for up to 2 weeks after the regimen concludes. Since there is incomplete cross-tolerance among the different narcotic agents, analgesia can often be sustained with a schedule of opioid rotation (263). Finally, there is a notion that tolerance can be avoided altogether if opioid doses are matched with the patient’s needs, such that there is no excess medication to cause euphoria and then tolerance.

Abrupt cessation of opioid therapy in a physically dependent patient can lead to withdrawal symptoms, which can sometimes be subtle and manifest only as complaints of mild, nonspecific muscle aches. Onset and duration of the withdrawal process correlates with the half-life of the specific drug, but autonomic symptoms can be blunted with the use of oral or TD clonidine, at a dose of 0.1 to 0.2 mg per day. Similar to tolerance, there is a notion that if doses are matched with the patient’s needs then physical dependence will never develop. However, traditional teaching is that most patients who take opioids for more than 1 month will have some degree of physical dependence. In order to avert withdrawal symptoms in physically dependent patients, detailed guidelines have been developed for weaning patients off opioid treatment (264,265).

The overall side-effect profile for narcotic analgesics is relatively favorable, especially in the elderly population. Constipation is the most common narcotic side effect and it is also the only one to which tolerance will not develop over time. Since it can profoundly negatively affect patients’ lives, the initiation of prophylactic bowel stimulants or osmotic agents should be strongly considered with opioid therapy. If constipation develops despite prophylaxis, it should be treated aggressively.

Other prevalent GI side effects include nausea with or without vomiting. Prophylaxis against nausea is not routinely employed because patients usually build tolerance to nausea over time. Treatment of nausea depends upon its etiology. If it is due to constipation, then the latter should be treated; in contrast, if nausea is due to a primary effect of the medication (i.e., stimulation of the chemotrigger zone), then prochlorperazine is the first-line agent for treatment of this. If the patient remains nauseous after addressing the above and other possible etiologies, agents such as haloperidol could be considered. If the patient additionally develops agitation, chlorpromazine would be an appropriate second-line agent of choice. Metoclopramide can be employed if the suspected etiology is gastric outlet obstruction, secondary to the antmotility effect of opioids. Nonoral adjuvant analgesics (e.g., IV indomethacin or IM ketorolac) may be more appropriate for patients suffering from narcotic-related nausea.

Patients on opioid therapy can also experience significant non–GI-related adverse events. Orthostatic hypotension, for example, can reach a degree where it limits rehabilitative transfer and ambulation training. Respiratory depression, manifested as reduced respiratory rate in the early stage, can also become severe enough to cause respiratory arrest. Habitual users are less likely than narcotic-naïve patient to suffer the extremes of either event because of a phenomenon known as *tolerance*, where patients adapt to a drug’s effects over time. Evolving theories attribute receptor desensitization and receptor down-regulation to tolerance (266).

Opioids can induce a variety of additional side effects. For example, opioids can influence various hormones of the hypothalamus-pituitary-gonad system (267,268). Both acute administration and chronic use of opioids, in the absence of pain, can be immunosuppressive (269). Repeated administration may produce hyperalgesia (266,270,271). Potential CNS side effects include sedation and euphoria. Sedation can be countered using stimulants such as caffeine, dextroamphetamine, and methylphenidate. Euphoria is comparatively more problematic as it is the basis behind psychological addiction.

**Tramadol Hcl (Ultram); Tramadol Hcl and Acetaminophen (Ultracet)**

**Relevance to Physiatry**

Tramadol is a centrally acting, synthetic analgesic with FDA approval for management of moderate-to-severe pain. It is a unique drug because it is not classified as a controlled substance (though its μ-receptor binding affinity is similar to
codeine) and it has an additional mechanism of action akin to many antidepressants. Given the dual mechanism, tramadol can be an effective analgesic for both nociceptive and neuropathic pain (272). Randomized-controlled trials have demonstrated its efficacy in treating postherpetic neuralgia, phantom limb pain, diabetic neuropathy, and polyneuropathy of various etiologies (273–276). As for chronic nociceptive pain, the American College of Rheumatology (ACR) recommends tramadol for OA patients who failed to achieve adequate benefit from nonnarcotic analgesic medications (277). Finally, this weak opioid agonist can also be used to treat migraine, moderately severe episodic breakthrough pain and it has been studied in acute dental and surgical pain (278,279).

Ulteracet contains 37.5 mg of tramadol and 325 mg of acetaminophen. Its exact role in pain management is evolving. Its advantages, like any other combination analgesics, are synergistic analgesia and reduced dose-dependent side effects (280). Only three randomized, placebo-controlled studies have been conducted to date using Ulteracet and they show that it is as effective as and better tolerated than acetaminophen with codeine (at 300 mg per 30 mg) for chronic, nonmalignant low back pain, OA pain, and fibromyalgia (281–283). An ongoing randomized, multicenter is investigating Ulteracet’s efficacy in treating pain associated with rheumatoid arthritis (284).

The extended-release form, Ultras ER, has a convenient once-daily dosing schedule for managing chronic pain conditions. It is shown to have similar tolerability and effectiveness at relieving moderate to severe OA pain as the original tramadol (285).

All things considered, tramadol is potentially advantageous compared to NSAIDs and COX-II inhibitors because it does not cause GI bleeding or exacerbation of HTN or CHF (286). There may also be decreased tolerance to tramadol’s therapeutic effect with chronic use in treatment of pain due to knee and hip OA (287,288).

**Mechanism of Action and Pharmacokinetics**

Tramadol has two complementary analgesic mechanisms including weak activation of μ-receptors and pain impulse transmission modification, via weak inhibition of NE and serotonin reuptake. This compound can therefore be thought of as both a synthetic opioid and a TCA.

Tramadol has high bioavailability. Following oral administration, one-fifth of the drug is protein bound and it undergoes extensive first-pass hepatic metabolism. Onset of analgesia is apparent within 1 hour and a mean peak plasma concentration is reached within 1.5 to 2 hours. The liver demethylates and glucuronidates tramadol into several metabolites, only one of which have analgesic properties (289). All metabolites and the unchanged portion are renally excreted. Tramadol has a half-life of 6 hours and steady state is achieved within 2 days when it is taken four times daily.

**Preparations and Dosing**

Tramadol is available in the United States as 50-mg IR tablets. Extended-release tablets (Ultras ER in denominations of 100, 200, and 300 mg) recently became available in U.S. market. IV, IM, rectal, and subcutaneous preparations are available internationally.

The usual dosing range for IR tramadol is between 50 and 100 mg, every 4 to 6 hours. The maximum recommended dose is 400 mg daily. Dosing adjustments are recommended for patients older than 75 years old (<300 mg/day), with a creatinine clearance less than 30 mL/minute (administer every 12 hours, with a maximum daily dose of 200 mg), and a history of hepatic dysfunction. Patients with nonacute pain initiated on a starting dose of 50 mg every day, increased by one 50 mg dose every 3 days until the maximum daily dose is reached; this strategy minimizes side effects. Patients with acute pain can be treated with an initial 50-mg dose, followed by another 25 to 50 mg dose if adequate analgesia is not achieved within the first hour.

The starting dose for extended-release therapy is 100 mg daily, which can be increased at a rate of 100 mg/day for every 5 days. The maximum recommended dose is 600 mg/day but a daily dose of 400 mg has been safely prescribed for patients under 75 years old in several studies (290,291).

**Relevant Side Effects and Drug Interactions**

Nausea, drowsiness, and constipation are the most common side effects (292). The frequency of both is notably less than with traditional opioids. Vomiting, somnolence, abdominal pain, headache, dry mouth, dyspnea, and vertigo have also been reported (280). Respiratory depression and pruritus are potential side effects associated with all opioids, but tramadol has a much lower risk of either incident (293). Serious, and rarely fatal, anaphylactoid reactions have occurred following the first tramadol dose; individuals with a history of anaphylactoid reactions to opioids may be at higher risk.

Seizure is a rare but significant side effect of tramadol. Caution should be exercised if tramadol is prescribed to patients with epilepsy, a seizure history or seizure risk factors (e.g., TBI, metabolic disorders, alcohol/drug withdrawal, and CNS infections). SSRIs or TCAs should not be prescribed with tramadol because they may further lower the seizure threshold. CNS depressants and MAOIs should be avoided as well because they pose risks of and hypertensive crisis, respectively. One should also be aware that carbamazepine markedly induces tramadol’s metabolism, such that twice the usual tramadol dose might be needed.

Tramadol is a nonscheduled opioid and it has limited abuse potential because it rarely produces euphoria. However, abrupt cessation of therapy can lead to opioid withdrawal symptoms—though they may not as severe as with other opioids—and atypical withdrawal symptoms (e.g., hallucinations and paranoia) (294).

**TD AND TOPICAL ANALGESIC MEDICATIONS**

The skin can be used to deliver medications locally, to an underlying target tissue, or systemically. Depending upon the
medication, delivery can be via an exogenous, disposable TD delivery system, or topical application of cream or ointment. While the terms TD and topical are often used interchangeably, TD delivery aims to achieve systemic therapeutic levels similar to those of oral administration and can be administered distal to the target site over an extended period of time. The fentanyl patch is an example of a true TD delivery system while topical NSAIDs, the topical NSAID patch Flector (topical diclofenac epolamine), and the lidocaine patch are examples of topical application.

TD delivery and topical application have become increasingly popular as it circumvents the unpredictability of GI tract absorption and hepatic first-pass metabolism. Topical medications can also achieve higher therapeutic concentration at local sites than do systemically administered medications.

All TD delivery systems comprise three elements: a backing, the active drug, and the adhesive. First-generation reservoir systems release active drug via a rate-limiting membrane. Second-generation matrix systems have the active drug embedded in polymer layers and are directly applied to the skin. Both systems deliver the active drug at a constant rate and thereby maintain constant drug-plasma level. This feature, in turn, extends the therapeutic activity for drugs with short half-lives and improves patient compliance by increasing the dosing interval.

An increasing number of medications will likely be delivered via the skin in the future as new mechanical enhancement methods are developed to increase TD topical permeability and allow patient-controlled administration. Iontophoresis is a currently available enhancement mechanism and various others (e.g., electroporation, sonophoresis, and microneedles) are being investigated (295–297).

Unfortunately, several factors limit TD and topical analgesic medication application. For one, significant contact dermatitis can occur among selected individuals. TD delivery and topical application are also not suitable for all medications. The efficacy of individual analgesic medications is further complicated by variables including humidity of the skin, ambient temperature, a drug’s thermodynamic properties, and the target tissue’s properties. The major disadvantage remains that most TD preparations (especially that of topical NSAIDs) have relatively short shelf lives, which makes mass production difficult.

The physiatrist should become familiar with TD delivery and topical application because they may encounter some of these medications in daily practice. A partial list of available TD medications classes includes angesics, antibiotics, anticholinergics, antiinfectives, and hormones. Aforementioned medications can also be delivered across mucous membranes using intranasal sprays, ophthalmologic solutions, and suppository.

Topical agents are widely available OTC but many TD medications are only available using a specialized pharmacy known as a compounding pharmacy.

**Individual Agents**

EMLA, lidocaine, and Fentanyl patches have been discussed previously in this chapter.

**Diclofenac Epolamine (Flector Patch)**

**Relevance to Physiatry**

Although the international community has had significant experience with topical NSAIDs and they are reported to be well tolerated, particularly with respect to their lack of GI side effects associated (297), as of early 2008 Flector was the only topical NSAID approved in the United States.

The active ingredient in Flector patch is 1.3% diclofenac epolamine. It is indicated for minor sprains, strains, and contusions. Randomized controlled trials have also shown that it is effective in short-term treatment of symptomatic knee OA and epicondylitis (298,299).

An investigation of a combination preparation (heparin and lecithin with diclofenac epolamine) is underway and preliminary results suggest that this combination has superior anti-inflammatory, hemorheologic, and antiedema effects for treating local trauma (300,301).

The FDA is also evaluating another topical NSAID (Pennsaid) which is available in Canada and several European countries for treatment of osteoarthritic pain (302).

**Mechanism of Action and Pharmacokinetics**

The anti-inflammatory, analgesic, and antipyretic effects of diclofenac epolamine are similar to other NSAIDs. Flector patch reaches peak plasma level between 10 and 20 hours after application and the half-life of each patch is approximately 12 hours (303). Diclofenac avidly binds to plasma protein, is metabolized by the liver, and is then excreted along with urine and bile.

**Preparation and Dosing**

Each Flector patch contains 180 mg of diclofenac epolamine in an aqueous base. The recommended dose is one Flector patch to the intact skin over the affected area bid (303).

**Relevant Side Effects and Drug Interactions**

The most adverse effects of Flector patch are local skin reactions. Although diclofenac has similar side effect and drug-interaction profiles as other NSAIDs, the minimal serum medication levels associated with use in patch formulation lead to a lower risk of GI side effects (298).

**Capsaicin (Zostrix, Zostrix-Hp)**

**Relevance to Physiatry**

Capsaicin is a naturally occurring, reversible neurotoxin extracted from Solanaceae family plants (i.e., “hot” chili peppers) and is classified as a capsaicinoid. It is used by physiatrists to manage localized pain states (e.g., focal neuropathic pain) and joint arthralgias (e.g., knee and finger OA). Its use in OA is discussed in greater detail in Chapter 31. Topical capsaicin is FDA approved for the treatment of diabetic neuropathy, postherpetic neuralgia, along with pain due to OA and rheumatoid arthritis. Several studies found it to be beneficial for postoperative pain, trigeminal neuralgia, and cluster headache (95,304–308). There is also some scant literature on
management of complex regional pain syndrome type 1 and traumatic amputee neurogenic residual limb pain (309,310).

Capsaicin has recognizable efficacy in managing chronic musculoskeletal and neuropathic pain but is generally considered as an adjuvant analgesic. Study outcomes on the combined use of capsaicin and other topical medications have generally been favorable. For example, the combination of capsaicin (0.25%) and topical 3.3% doxepin was found to produce synergistic effects in neuropathic pain (20). A 2004 systematic review declared that capsaicin might be useful as an adjunct or sole therapy for patients with refractory pain (311).

Less direct physiatric applications of capsaicin include painful urological conditions, temporomandibular pain, oral mucositis, rhinitis, and psoriasis (16–20).

In addition to its clinical use, capsaicin has also become an integral part of pain management research (312,313). A human trigeminal sensitization pain model was developed to examine gender differences (314). In addition, basic science research suggests that an endogenous capsaicin-like substance is released in inflamed tissues and produces nociceptive neural impulses by acting on capsaicin receptors on sensory neurons (315).

Mechanism of Action and Pharmacokinetics
Analgesia is achieved by the binding to the vanilloid receptor-1 (VR1), which leads to the depletion of and inhibited accumulation of SP at the target site. SP is an endogenous neuropeptide produced by small-diameter, primary, sensory “pain” fibers; it is involved in the afferent transmission of pain impulse and stimulates immune cells (316,317). In addition to SP depletion, capsaicin also inhibits SP transport and de novo synthesis. The overall effect is a reversible sensory degeneration that leads to pain desensitization. Inflammation is also indirectly inhibited in the above process because inflammation has a neural component, which is referred to as neurogenic inflammation. The major proinflammatory players in this model are SP and related peptides (318). Capsaicin’s biochemistry is still under investigation (319,320).

Preparations and Dosing
Capsaicin is available under various trade names including 0.025% (Zostrix) and 0.075% (Zostrix-HP) topical cream preparations that are to be applied three to four times daily. Patients should be initiated on the lower-concentration preparation to minimize early, unfavorable effects.

Relevant Side Effects and Drug Interactions
Transient application site stinging and burning, caused by activation of C fibers, is experienced by up to 50% of patients at the onset of treatment, particularly with the higher potency formulation. Capsaicin studies have on average incurred 13% participant withdrawal rate due to intolerable burning (311). Capsaicin is not deemed appropriate for acute musculoskeletal pain for this reason though this side effect usually remits after the first few days (321). This phenomenon is attributed to depletion of SP with repeat application. Cough is another possible adverse effect. No adverse effect upon nerve function and no drug interactions have been documented to date (322).

ANTICOAGULANTS
Relevance to Physiatry
These medications are primarily used as prophylaxis against excessive clot formation. In general, they affect the function and synthesis of clotting factors, thus reducing the tendency toward clot propagation and the risk of developing or extending a deep venous thrombosis (DVT), or pulmonary/cerebral embolus. Patients at risk for developing thromboembolic phenomena are frequently encountered in the rehabilitative setting, and include those with SCI, stroke, multiple trauma, as well as hip and knee arthroplasties. Risk reductions, including anticoagulation and conservative means to stimulate blood flow in the legs, such as with external compression devices, are warranted in appropriate patients. Patients who have suffered a cerebral thromboembolic event, those with proximal DVT, and pulmonary emboli are generally treated similarly (323–330).

Anticoagulants can be classified into heparin and coumarin groups. The heparin group is administered parenterally and includes unfractionated heparin (UFH), low molecular weight heparin (LMWH), and heparinoids (including fondaparinux and danaparoid). IV full-dose UFH had been the mainstay of initial treatment of acute thromboembolic phenomena, such as DVT, pulmonary embolus (PE), as well as unstable angina and non-Q-wave myocardial infarction (MI), and requires inpatient therapy. Currently, it is acceptable to initiate therapeudic anticoagulation utilizing subcutaneous LMWH or certain heparinoids in appropriate patients, even on an outpatient basis in those with DVT, and even possibly PE (323–337). Although subcutaneous fixed low-dose UFH can be used for prophylaxis, LMWH and some heparinoids are more effective and are associated with a lower risk of hemorrhage (338–347). Heparinoids can also be used for thromboembolism prophylaxis or treatment in the presence of heparin-induced thrombocytopenia (348–351).

Warfarin is by far the most commonly used coumarin medication. Its use in DVT prophylaxis following hip or knee arthroplasty has been well established (352,353). Although its efficacy in preventing DVT in orthopedic patients may be somewhat less than that of LMWH, it is the mainstay of maintenance treatment of thromboembolism following initial heparin treatment. Another important indication for oral anticoagulation is stroke prophylaxis, especially in the presence of atrial fibrillation (AF) (353–355). Specifically, it has been shown that oral anticoagulation is associated with a significant reduction of stroke risk, and that it is more efficacious than aspirin in preventing stroke in the setting of AF (355).

Whenever an anticoagulated patient is undergoing physical and/or occupational therapy, it is prudent to include this information in the therapy prescription so that the therapist can take appropriate precautions to prevent falls and injuries.
This also alerts the therapist to monitor closely and report potential signs of bleeding, such as a newly swollen joint, as this may represent a hemarthrosis.

**Mechanism of Action and Pharmacokinetics**

Heparin and its derivatives affect key reactions involved with thrombosis and stable clot formation. Low-dose UFH and LMWH primarily prevent the conversion of prothrombin to thrombin (factor II) via inactivation of factor Xa. At higher doses, heparin can prevent fibrinogen-to-fibrin conversion by inactivation of thrombin and also prevents stable fibrin clot formation. After subcutaneous injection, activity onset is rapid, with peak plasma levels at about 4 hours. Metabolism occurs in the liver and the reticuloendothelial system.

Warfarin interferes with vitamin K, by inhibiting hepatic coagulation factor II, VII, IX, and X synthesis. Its effect is detectable once the baseline level of these factors already in circulation starts to be depleted by metabolic degradation. After oral administration, the maximal plasma concentration is between 1 and 9 hours. Almost all of the medication (97%) is bound to plasma albumin. The initial effect is apparent in 24 hours, but the peak effect occurs between 3 and 4 days and lasts for 4–5 days. It is metabolized in the liver and has a half-life of around 2½ days.

**Preparations and Dosing**

Both UFH and LMWH are mucopolysaccharides derived from porcine intestinal mucosa. LMWH is derived from UFH by depolymerization and possesses a higher ratio of antifactor Xa to antifactor IIa activity than pure UFH. Therefore, it is purported to have a lower incidence of bleed complications versus UFH. The molecular size of both heparin classes does not permit oral administration as they would not be readily absorbed from the GI tract. Therefore, they are currently administered either by deep subcutaneous injection or intravenously. Modified, orally active forms of heparin are being studied (356). Although it is necessary to monitor the activated partial thromboplastin time (aPTT) during full-dose IV UFH therapy, it is not necessary to do so during standard fixed low-dose UFH and LMWH therapy since this index is essentially unaffected. However, it is important to note that full therapeutic doses of LMWH can affect the aPTT.

These medications are generally utilized short term, in inpatient as well as outpatient settings (323–326,331,338–347). For prophylaxis, they are used based upon diagnosis or until risk of thromboembolism has diminished. In the event of acute PE or DVT, they are administered at therapeutic dosages concomitantly with warfarin, which is initiated within the first 3 days. Once a therapeutic INR is achieved, typically within 5 to 7 days, heparin is discontinued, while warfarin therapy is maintained for a variable time period depending on the underlying condition. In certain cerebrovascular accidents (CVAs) or MIs, antiplatelet agents are used in lieu of warfarin. Dosages and indications for subcutaneous administration of heparins and derivatives are shown in Table 65-18.

Warfarin sodium (Coumadin, Panwarfin) is the most commonly utilized oral anticoagulant. In those rare cases when oral administration is not feasible, warfarin may be given intravenously. Dosage is individualized and based on monitoring its efficacy by regular PT assessment. Because of significant variability in thromboplastin reagents used in making these assessments among laboratories, a common standardized scale, the (INR), was developed to allow for more comparable monitoring of efficacy regardless of the reagents used. An INR of 2.0 to 3.0 is generally recommended for both prophylaxis and thromboembolism treatment. The effectiveness of

**TABLE 65.18 Heparin and Heparin Derivatives Administered Subcutaneously**

<table>
<thead>
<tr>
<th>Generic (Trade) Name</th>
<th>Type</th>
<th>Typical Dosage and Indications (FDA Approved and Unapproved)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heparin sodium</td>
<td>UFH</td>
<td>DVT prophylaxis: 5,000 units q8–12 h; lower doses in geriatric patients</td>
</tr>
<tr>
<td>Dalteparin (Fragmin)</td>
<td>LMWH</td>
<td>DVT prophylaxis: 5,000 units qd following orthopedic surgery (FDA approved only for THA) or immobilization caused by acute illness; 2,500 units qd following abdominal surgery or for other conditions (not FDA approved).</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DVT/PE treatment: 200 units/kg qd</td>
</tr>
<tr>
<td>Enoxaparin (Lovenox)</td>
<td>LMWH</td>
<td>DVT prophylaxis: 30 mg q12h (approved for THA/TKA; unapproved for SCI and MT) or: 40 mg qd (approved alternate dosing following THA, immobilization because of medical illness or following abdominal surgery).</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DVT/PE treatment: 1 mg/kg q12h, or 1.5 mg/kg qd</td>
</tr>
<tr>
<td>Tinzaparin (Innohep)</td>
<td>LMWH</td>
<td>DVT/PE treatment: 175 anti-Xa units/kg qd until therapeutic anticoagulation with warfarin. Also unapproved indication for DVT prophylaxis: 4,500 units qd following orthopedic surgery or 3,500 units qd following general surgery</td>
</tr>
<tr>
<td>Fondaparinux (Arixtra)</td>
<td>Synthetic factor Xa inhibitor</td>
<td>DVT prophylaxis: 2.5 mg qd (approved for THA/TKA and hip fracture surgeries, as well as following abdominal surgery)</td>
</tr>
<tr>
<td>Desirudin (Iprivask)</td>
<td>Selective thrombin inhibitor</td>
<td>DVT prophylaxis following THA: 15 mg q12h</td>
</tr>
</tbody>
</table>
anticoagulation within this range is essentially equivalent to that of higher dosages, yet with fewer hemorrhagic events (357).

In contrast to biological heart valve replacement, a higher INR (2.5 to 3.5) is suggested for mechanical heart valves. For the clinician, it can sometimes be challenging to determine the specific dosage of warfarin that will maintain the INR of a given patient within the therapeutic range. However, there has been recent association of specific genotypes, particularly CYP2C9 and CYP2C9 and VKORC1, with variability in warfarin dosing efficacy in certain individuals, which may lead to the development of more “personalized” dosing regimens with a reduction of unpredictability in achieving therapeutic INRs (338,358–366).

Warfarin's usual starting dose is from 5 to 10 mg qd. Larger loading doses should be avoided, given the increased risk for hemorrhage. It is also prudent to start with lower loading doses in the elderly or debilitated or in those who are known to be sensitive to warfarin. Oral anticoagulation is usually initiated at the time of heparin treatment, is overlapped with it for about 4 to 5 days in order for the INR to increase to 2.0–3.0, and is then continued alone. For instance, the common practice in treating orthopedic patients is to discontinue the anticoagulant before discharge home or when the patient is fully ambulatory. However, there is evidence that supports longer treatment in order to further reduce DVT risk (367,368).

Relevant Side Effects and Drug Interactions

Adverse effects of greatest concern with anticoagulants are hemorrhage, most commonly presenting as bruising, petechiae, epistaxis, GI, or urinary tract bleeding. Extreme caution should be exercised in those at increased risk for bleeding, including those who are already taking antiplatelet agents.

A potential problem specific to heparin is heparin-induced thrombocytopenia which occurs from either a direct effect on platelets or by an immunologic response. Thus, it is important to monitor CBCs, especially at the start but also at regular intervals during treatment. Also, there have been reports of spinal epidural hematomas with ensuing paralysis following use of LMWH in patients who have or very recently had epidural catheters or punctures. Therefore, the manufacturers of LMWH have issued warnings against administering these drugs within specific times before and following procedures involving epidural invasion. Other adverse effects that may occur include local irritation (especially with deeper injections; thus the need to avoid IM injection) and hypersensitivity reactions, especially with pork allergies.

A relatively rare but potentially hazardous reaction associated with warfarin use is skin necrosis, which may develop in susceptible individuals such as those patients with protein C deficiency because there is normally local thrombosis that occurs within the first few days of initiating the drug. The concurrent administration of heparin for the first 5 to 7 days of anticoagulation reduces the risk of this reaction.

Many drugs can alter the therapeutic efficacy of warfarin. Those that may increase the PT include APAP, amiodarone, ASA, NSAIDs, phenytoin, sulfonamides, and thyroid supple-ments. Those that can decrease the PT include adrenocorticoids, antacids, antihistamines, carbamazepine, haloperidol, and vitamin C. Others, including diuretics and H2 blockers, may cause either an increase or decrease in the PT.

ANTITHROMBOTICS

Relevance to Physiatry

Antithrombotics (i.e., antiplatelet drugs, because of their primary effect of inhibiting platelet aggregation) have widespread use in patients with cardiovascular illness. The most widely used agent in this group is ASA, which, in appropriate patients, has been associated with significant reductions stroke and MI risks (369–371). It has also been suggested that ASA given early after ischemic CVA onset can help to reduce morbidity and mortality (372,373). However, other antithrombotic drugs such as clopidogrel, cilostazol, ticlopidine, and dipyridamole play an important role. In fact, there is evidence that they can be more effective than ASA in stroke prevention. This had been demonstrated with ticlopidine, but the risk of potentially serious adverse effects has limited its use (374). In contrast, clopidogrel, a drug similar to ticlopidine, has been reported to be as effective yet with a better safety profile (375–380). The combined use of dipyridamole and ASA is superior for stroke prevention compared to ASA alone (381). Ongoing trials are comparing these agents to one another as well as in combination to determine greatest efficacy in cardiovascular disease prevention (382–386).

Cilostazol is a phosphodiesterase III inhibitor that improves walking distance and exercise tolerance in patients with peripheral vascular disease and may also improve lipid profiles (387–390).

MECHANISM OF ACTION AND PHARMACOKINETICS; PREPARATIONS AND DOSING; RELEVANT SIDE EFFECTS AND DRUG INTERACTIONS

See Table 65-19.

ANTHYPERLIPIDEMIC MEDICATIONS

Relevance to Physiatry

The importance of antihyperlipidemic agents in therapy and risk reduction of cardiovascular disease has been well established (391–394). The reduction of serum lipids, including cholesterol and triglycerides, is associated with decreased cardiac morbidity from atherosclerotic cardiovascular disease. More recent evidence suggests that these medications can reduce ischemic stroke risk in the presence of cardiovascular risk factors and may also play a neuroprotective role during CVA (391–398).

Mechanism of Action and Pharmacokinetics

These drugs can be classified into fibrates and HMG-CoA reductase inhibitors, known popularly as the statins. The latter
include bile acid sequestrants as well as ezetimibe, an inhibitor of GI absorption of cholesterol. Other agents

primarily lower serum low-density lipoprotein (LDL) cholesterol levels, whereas the fibrates lower serum triglyceride and very low density lipoprotein (VLDL) cholesterol levels. All these medications have the beneficial effect of elevating serum high-density lipoprotein (HDL) cholesterol. Other agents

Preparations and Dosing
See Table 65-20.
Relevant Side Effects and Drug Interactions
Potential concerns associated with these drugs include myopathy and rhabdomyolysis, especially when both classes are used together, and liver dysfunction. The fibrates in particular can also elevate PT results in patients receiving warfarin.

CARDIOVASCULAR MEDICATIONS
Medications that affect the cardiovascular system are frequently encountered in the rehabilitative setting, especially in the elderly. It is not uncommon for the physiatrist to adjust these medications, especially while the patient is on an inpatient rehabilitation unit. Because these medications can have such a profound impact on the patient’s medical status and general well being, it is important for the physiatrist to have a basic knowledge of the types of medications used, their indications, and their effects.

Alpha-blockers: α-1 Adrenergic and α-2 Adrenergic Antagonists
Relevance to Physiatry
The primary use of these medications, especially α-1 adrenergic blockers, is for the treatment of HTN and prostate-related bladder outlet obstruction, especially α-1 adrenergic blockers. They are also used for the control and prevention of vascular manifestations of autonomic dysreflexia (399). A nonselective α-blocker such as phenoxybenzamine can also be effective in RSD because it affects both types of α receptors, thus leading to a chemical sympathectomy. It is usually best to initiate these drugs at bedtime given the frequent occurrences of orthostatic hypotension and reflex tachycardia during the first few days of initiating treatment (i.e., the first-dose phenomenon) and the fact that postural hypotension can also occur following rapid increases in dosage, during the established use of these medications, or following the addition of a second antihypertensive.

Alpha-2 selective blockers (e.g., yohimbine [Yocon]) might help with erectile dysfunction, diabetic neuropathy, and postural hypotension (400,401).

Mechanism of Action and Pharmacokinetics
These medications block α-1 and/or α-2 adrenergic receptors. Selective α-1 receptor blockers inhibit vasoconstriction, resulting in decreased peripheral vascular resistance and a subsequent decrease in BP. They also reduce resistance to urinary outflow by inhibiting smooth muscle contractions at the base of the urinary bladder. In contrast, selective α-2 blockers and nonselective blocking agents indirectly cause the release of NE from peripheral sympathetic nerve endings, thus increasing overall sympathetic outflow, and leading to subsequent elevations in heart rate (HR) and BP.

Preparations and Dosing; Relevant Side Effects; and Drug Interactions
See Table 65-21.

Alpha-2 Agonists
Relevance to Physiatry; Mechanism of Action; and Pharmacokinetics
These medications are indicated for HTN, either alone or in combination with other medications. They may be used in acute autonomic dysreflexia refractory to other measures and also in the prevention of its recurrence (399). Because of their centrally acting mechanism (stimulation of α-2 adrenergic receptors), they may cause drowsiness, a definitely untoward

### TABLE 65.21 Cardiovascular Medications Acting on the Sympathetic Nervous System

<table>
<thead>
<tr>
<th>[Alpha agonists]</th>
<th>Typical Dose (mg)</th>
<th>Other Properties and Side Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clonidine hydrochloride (Catapres, Catapres-TTS)</td>
<td>Oral: 0.1 bid–2.4 mg/d patch: 0.1–0.3/d–change q wk</td>
<td>Renally metabolized; also used to blunt autonomic symptoms from narcotic withdrawal</td>
</tr>
<tr>
<td>Guanfacine hydrochloride (Tenex)</td>
<td>0.5–1.0 PO/d–3.0/d</td>
<td>Taken H/S to avoid daytime somnolence; renal metabolism</td>
</tr>
<tr>
<td>Guanabenz acetate (Wytensin)</td>
<td>4 PO bid and increased prn by 4–8/d up to 32/d</td>
<td>Hepatically metabolized</td>
</tr>
<tr>
<td>Methyldopa (Aldomet)</td>
<td>250 PO bid up to 3,000/d</td>
<td>First metabolized in the brain to methyl-NE, which activates α-2 receptors; parenteral form also available</td>
</tr>
<tr>
<td>Tizanidine (Zanaflex)</td>
<td>4–8 PO q6–8 h, max of 38 for every 24 h for spasticity control</td>
<td>Reported to have a much less tendency than clonidine for lowering of BP, thus not used for HTN</td>
</tr>
</tbody>
</table>

### [α-1 antagonists]

<table>
<thead>
<tr>
<th>[α-1 antagonists]</th>
<th>Typical Dose (mg)</th>
<th>Other Properties and Side Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Doxazosin (Cardura)</td>
<td>Start 1 PO qd; maximum 16 PO qd</td>
<td>If peripheral edema, can switch to Minizide (prazosin/ polythiazide), 1 tab PO bid</td>
</tr>
<tr>
<td>Prazosin (Minipress)</td>
<td>Start 1 PO bid-tid; maximum 40 PO qd</td>
<td>Reportedly less postural hypotension compared to the other α-blockers</td>
</tr>
<tr>
<td>Tamsulosin (Flomax)</td>
<td>BPH: 0.4 PO qd; maximum 0.8 PO qd</td>
<td></td>
</tr>
<tr>
<td>Alfuzosin (Uroxatral)</td>
<td>BPH: 10 qd</td>
<td></td>
</tr>
<tr>
<td>Terazosin (Hytrin)</td>
<td>Start 1 PO qH/S; maximum 20 PO qd</td>
<td>Not available in combination with a diuretic</td>
</tr>
</tbody>
</table>
effect when attempting to objectively monitor the progress of a patient with altered mental status (e.g., early TBI or CVA) and therefore can interfere with rehabilitation. Clonidine and especially the newer α-2 agonist tizanidine (Zanaflex) are also useful in treating spasticity in various patients, including those with SCI, TBI, multiple sclerosis, and CVA (402–408). TD or intrathecal clonidine may be of benefit in RSD pain (409,410).

Preparations and Dosing
See Table 65-21.

Relevant Side Effects and Drug Interactions
Dry mouth, sedation, various GI symptoms, and orthostatic hypotension are the most common side effects. Rebound HTN may occur with their sudden discontinuation. Major drug interactions include potentiation of the action of CNS depressants, attenuation of the hypotensive effect of TCAs, and additive cardiovascular side effects with other cardiovascular medications.

Angiotensin-Converting Enzyme Inhibitors (ACEIs) and Angiotensin II Receptor Blockers
Relevance to Physiatry
These medications are commonly encountered in hypertensive and/or cardiac rehab patients. They exert a renal-protective effect in diabetics and are useful in reducing the risk of recurrent CVA (411–416). An advantage of these classes of medications is their relative lack of cardiovascular side effects compared to other agents.

Mechanism of Action and Pharmacokinetics
ACEIs prevent the conversion of angiotensin I to angiotensin II, a potent vasoconstrictor and adrenal stimulator. The overall effect is a reduction in systemic vascular resistance with subsequent lowering of BP and cardiac output improvement. The hypotensive effects of angiotensin II receptor blockers (ARBs) occur via selective blocking of the binding of angiotensin II to specific AT receptors especially in vascular smooth muscles and adrenal glands. Medications of this class are well absorbed from the GI tract and are well tolerated. They vary in onsets, peaks, and durations of action, which is reflected in their recommended dosage.

Preparations and Dosing
See Table 65-22.

Relevant Side Effects and Drug Interactions
These medications are generally well tolerated with rare serious adverse reactions. Side effects can include skin rash, neutropenia, impaired renal function (especially if there is preexisting impairment), GI discomfort, and a nonproductive cough due to an alteration in bradykinin production. Renal function should be monitored, especially in the elderly with poor oral fluid intake. ACE inhibitors can act synergistically with diuretics or other antihypertensives. Hyperkalemia can occur when drugs that can elevate serum potassium concentrations (such as potassium supplements or potassium-sparing diuretics) are used with ACEIs and ARBs. Adverse effects of ARBs are generally minor and rare.

ANTIARRHYTHMIC MEDICATIONS
Relevance to Physiatry
This group is likely to be encountered in patients undergoing cardiac rehabilitation or those who have a pertinent cardiac history and undergoing inpatient rehabilitation.

Preparations and Dosing
See Table 65-23.

Relevant Side Effects and Drug Interactions
The most common antiarrhythmic medication side effects include cardiac rhythm disturbances, hypotension, and GI disturbance. Since various drug interactions are possible, it is important for the physiatrist to consider this whenever a new medication is added.

Beta-Blockers
Relevance to Physiatry
The physiatrist should understand the reason(s) for which a patient is on a β-blocker. This is especially important following MI or with cardiac arrhythmias. During the second phase of a three-phase cardiac rehab program, a therapeutic exercise program of appropriate intensity for a patient on β-blockers is based on a percentage of the symptom-limited HR or on the maximum work load performed on an exercise stress test rather than being based on absolute HR (417). Beta-blockers tend to impair a patient’s exercise tolerance.

Beta-blockers are especially used for HTN, angina pectoris, and arrhythmias. More recently, some of the β-blockers have been used in selected patients with compensated CHF (418,419). In addition, β-blockers are also used for migraine headaches, heightened metabolic turnover states, such as thyrotoxicosis, tremors of various etiologies, and may be of benefit in treating aggression such as that associated with TBI (420–425).

Mechanism of Action
These medications bind with varying affinities to β-1 and β-2 adrenergic receptors. Those having a predominant β-1 receptor affinity are referred to as β-1 selective or cardio selective as this receptor is primarily cardiac. The therapeutic effect therefore is to decrease HR and myocardial contractility, which contributes to decreasing myocardial oxygen demand in coronary artery disease and leads to BP reduction in certain hypertensive patients. Others β-blockers are nonselective since they demonstrate essentially equal affinities for both β-1 and β-2 receptors. There are also β-2 selective compounds; however, they are not clinically relevant because blockade of the β-2 receptors (mostly located in bronchiolar smooth muscles) causes bronchoconstriction. Some also possess mild β-mimetic or
NSAIDs, phenytoin, and phenobarbital can alter the effects of β-blockers. Antihypertensives and antiarrhythmics should be used cautiously if given with β-blockers.

**Calcium Channel Blockers**

**Relevance to Physiatry**

These medications are commonly encountered in the physiatric setting given their efficacy and convenient dosing (especially the controlled- and sustained-release preparations) for HTN and angina. The IR forms, although useful in various clinical situations, have become less popular because of adverse outcomes reportedly related to their use (426). This especially applies to SCI patients with autonomic dysreflexia for which nifedipine had been commonly given.

### TABLE 65.22 ACEIs and Angiotensin II Receptor Blockers

<table>
<thead>
<tr>
<th>Generic (Trade) Name</th>
<th>Oral Dose (mg)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Angiotensin-Converting Enzyme Inhibitors</strong></td>
<td></td>
</tr>
<tr>
<td>Benazepril (Lotensin)</td>
<td>HTN: 10 qd, up to 80 daily</td>
</tr>
<tr>
<td>Captopril (Capoten)</td>
<td>HTN: 25 bid-tid, up to 450 daily</td>
</tr>
<tr>
<td></td>
<td>CHF: 6.25–12.5 tid</td>
</tr>
<tr>
<td></td>
<td>LV dysfunction post MI: 12.5 tid to 50 tid</td>
</tr>
<tr>
<td></td>
<td>Diabetic nephropathy: 25 tid</td>
</tr>
<tr>
<td>Enalapril (Vasotec)</td>
<td>HTN: 5 qd/bid, up to 40 daily</td>
</tr>
<tr>
<td></td>
<td>CHF: 2.5 qd/bid, up to 40 daily</td>
</tr>
<tr>
<td>Fosinopril (Monopril)</td>
<td>HTN: 10 qd, up to 80 daily</td>
</tr>
<tr>
<td>Lisinopril (Prinivil, Zestril)</td>
<td>HTN: 10 qd, up to 40 daily</td>
</tr>
<tr>
<td></td>
<td>CHF: 5 qd, up to 20 daily</td>
</tr>
<tr>
<td>Moexipril (Univasc)</td>
<td>HTN: 7.5 qd, up to 30 daily</td>
</tr>
<tr>
<td>Perindopril (Aceon)</td>
<td>4 qd up to 16 daily</td>
</tr>
<tr>
<td>Quinapril (Accupril)</td>
<td>HTN: 10 qd, up to 80 daily</td>
</tr>
<tr>
<td></td>
<td>CHF: 5 bid, up to 20–40 daily</td>
</tr>
<tr>
<td>Ramipril (Altace)</td>
<td>HTN: 2.5 qd, up to 20 daily</td>
</tr>
<tr>
<td></td>
<td>Reduction in risk of myocardial infarction, stroke, and death from cardiovascular causes: initial 2.5 qd for 1 wk; 5 qd for the next 3 wk, then up to 10 qd</td>
</tr>
<tr>
<td></td>
<td>CHF post-MI: 2.5 bid up to 5 bid</td>
</tr>
<tr>
<td>Trandolapril (Mavik)</td>
<td>HTN: 1 qd, up to 4 daily</td>
</tr>
<tr>
<td></td>
<td>CHF post-MI or LV dysfunction post-MI: initially 1 qd to 4 qd</td>
</tr>
<tr>
<td><strong>Angiotensin II Receptor Blockers</strong></td>
<td></td>
</tr>
<tr>
<td>Candesartan (Atacand)</td>
<td>HTN: 16 qd, max 32 qd</td>
</tr>
<tr>
<td></td>
<td>CHF: 4 qd to 32 qd</td>
</tr>
<tr>
<td>Eprosartan (Teveten)</td>
<td>HTN: 600 qd, max 800 qd (in single or divided bid)</td>
</tr>
<tr>
<td>Irbesartan (Avapro)</td>
<td>HTN:150 qd, max 300 qd</td>
</tr>
<tr>
<td></td>
<td>Diabetic nephropathy: 300 qd</td>
</tr>
<tr>
<td>Losartan (Cozaar)</td>
<td>HTN: 50 qd, max 100 qd</td>
</tr>
<tr>
<td></td>
<td>Diabetic nephropathy: 50 qd to 100 qd</td>
</tr>
<tr>
<td>Olmesartan (Benicar)</td>
<td>HTN: 20 qd, max 40 daily # 30</td>
</tr>
<tr>
<td>Telmisartan (Micardis)</td>
<td>HTN: 40 qd, max 80 qd</td>
</tr>
<tr>
<td>Valsartan (Diovan)</td>
<td>HTN: 80 qd, max 320 qd</td>
</tr>
<tr>
<td></td>
<td>CHF: 40 bid, max 320 qd (in divided doses)</td>
</tr>
<tr>
<td><strong>Direct Renin Inhibitor</strong></td>
<td>HTN: 150 to 300 qd</td>
</tr>
<tr>
<td>Aliskiren (Tekturna)</td>
<td></td>
</tr>
<tr>
<td><strong>Selective Aldosterone Receptor Antagonist</strong></td>
<td></td>
</tr>
<tr>
<td>Eplerenone (Inspra)</td>
<td>HTN: 50 qd, max 50 bid</td>
</tr>
<tr>
<td></td>
<td>CHF post-MI: start 25 qd, target 50 qd</td>
</tr>
</tbody>
</table>

*a* Starting dosage should be ½ in volume-depletion.

α-1 antagonistic activity, which facilitates cardiac stimulation while blocking systemic effects of excess catecholamines.

**Preparations and Dosing**

See Table 65.24.

**Relevant Side Effects and Drug Interactions**

Adverse effects that can interfere with rehabilitation include orthostatic hypotension, bradycardia, and chronic obstructive pulmonary disease (COPD), asthma, or CHF exacerbation. Fatigue, sleep disturbance, and depression may also occur. Sudden medication discontinuation may exacerbate angina because of increased catecholamine sensitivity that may develop during prolonged treatment.
and edema. Coughing, GI disturbances, and bradyarrhythmias (especially in diseases affecting sinoatrial and AV nodes) can occur. Concomitant use with β-blockers is usually well tolerated, but sometimes leads to hypotension, CHF, or cardiac dysrhythmias. CaCBs may elevate serum digoxin and levels of other medications metabolized via the cytochrome P-450 system.

**Cardiac Glycosides**

**Relevance to Physiatry**

Cardiac glycosides include digitalis, digoxin, digitoxin, and deslanoside, and are primarily used for mild to moderate CHF and ventricular rate control in AF. Although digoxin is most commonly used, the group is often referred to as digitalis, the prototype agent. As potentially lethal side effects can occur, Concomitant use with β-blockers is usually well tolerated, but sometimes leads to hypotension, CHF, or cardiac dysrhythmias. CaCBs may elevate serum digoxin and levels of other medications metabolized via the cytochrome P-450 system.

**Mechanism of Action and Pharmacokinetics**

These drugs inhibit the entry of Ca²⁺ ions into the myocardial, sinoatrial, AV, and vascular smooth muscle cells by binding to specific voltage-sensitive channels in the cell membrane. This results in multiple cardiovascular effects. One of these is to reduce the strength of myocardial contraction. Another is to relax the arterial smooth muscles in the coronary arteries and peripheral arterioles. This leads to improved coronary blood flow and decreased peripheral vascular resistance with systemic BP reduction. Finally, pacemaker cell depolarization is impaired with a resultant slowing of cardiac conduction.

**Preparations and Dosing**

See Table 65-25.

**Relevant Side Effects and Drug Interactions**

Side effects most commonly relate to vasodilatation and include dizziness, orthostatic hypotension, reflex tachycardia, headache, and edema. Coughing, GI disturbances, and bradyarrhythmias (especially in diseases affecting sinoatrial and AV nodes) can occur. Concomitant use with β-blockers is usually well tolerated, but sometimes leads to hypotension, CHF, or cardiac dysrhythmias. CaCBs may elevate serum digoxin and levels of other medications metabolized via the cytochrome P-450 system.

**Cardiac Glycosides**

**Relevance to Physiatry**

Cardiac glycosides include digitalis, digoxin, digitoxin, and deslanoside, and are primarily used for mild to moderate CHF and ventricular rate control in AF. Although digoxin is most commonly used, the group is often referred to as digitalis, the prototype agent. As potentially lethal side effects can occur, one should have a high index of suspicion for early signs of digoxin toxicity and the lowest effective dose should be used. Periodic monitoring of its efficacy and serum medication and electrolyte levels is wise.

**Mechanism of Action and Pharmacokinetics**

Digitalis increases myocardial contractile force via increased intracellular Ca²⁺ concentration, which enhances myofilament cross-bridging. Digitalis also slows cardiac impulse conduction. This is related to changes in cellular ion transport and concentrations that affect the rates and degree of transmembrane potentials.

GI absorption of digoxin varies depending on the preparation; thus, it is important to familiarize oneself with a specific preparation. After oral dosage, peak plasma concentration occurs in 2 to 3 hours and maximum effect in 4 to 6 hours. The half-life is 1 to 2 days, and steady-state plasma concentration

---

**TABLE 65.23 Antiarrhythmic Medications**

| Class IA | Electrophysiologic effect: blockage of sodium channels | Prevention of AF/flut and PVCs, 324–648 q8–12 h; AF/flut, 200–400 q6–8 h (IR) or 300–600 q8–12 h (ER); ventricular dysrhythmias, 250–625 q3h (IR) or 500–1,250 q6h (SR) or 500–1,000 q12h (Procanbid); ventricular dysrhythmias 400–800 per 24 h if IR or q12h if ER; 200 q8h, max of 900 per 24 h
| Quinidine gluconate (Quinaglute, Quinalan, Quinate); quinidine sulfate (Quinidex, Quinora); procainamide (Procanbid, Pronestyl); disopyramide (Norpace, Norpace CR); moricizine (Ethmozine) |
| Class IB | Electrophysiologic effect: blockage of sodium channels | VT/VF, monitored IV dosage; ventricular dysrhythmias, 200 q8h, max 400 per 24 h; ventricular dysrhythmias, 400 q8h, max 2,400 per 24 h
| Lidocaine (Xylocaine, Xylocard); mexiletine (Mexitil); tocainide (Tonocard) |
| Class IC | Electrophysiologic effect: blockage of β adrenergic receptors | Venticular dysrhythmias, 200 bid, max 1,200 bid; 10–30 tid–qid; ventricular dysrhythmias (Betapace) or AF/flut (Betapace AF), 80 bid, max 640 per 24 h
| Acebutolol (Sectral); propranolol (Indral, Inderal LA); sotalol (Betapace, Betapace AF) |
| Class II | Electrophysiologic effect: blockage of calcium channels | Venticular dysrhythmias, 800–1,600 qd for 1–3 wk, then 600–800 qd for 1 month, then 400 qd; monitored IV dosage; AF/flut, 0.5 bid; see above
| Flecainide (Tambocor); propafenone (Rhythmol) |
| Class III | Electrophysiologic effect: prolongation of membrane depolarization | PSVT or PAF: 50 q12h, max 300 per 24 h or sustained VT: 100 q12h, max 400 per 24 h; PSVT, PAF, sustained VT: 150 q8h, max 300 q8h
| Amiodarone (Cordarone, Pacerone) brytium (Brytyl); dofetilide (Tikosyn); sotalol (see earlier) |
| Class IV | Electrophysiologic effect: blockade of calcium channels | PSVT or RVR in AF, 240–480 per 24 h divided tid–qid
| Verapamil (Calan, Calan SR, Covera-HS, Isoptin) |

---

sublingually, a practice that should be avoided. Predominant side effects include orthostatic hypotension, peripheral edema, and headache. The physiatrist should be careful in ordering physical therapy modalities (e.g., whirlpool or Hubbard tank) that can accentuate the vasodilation in a patient taking a CaCB or other vasodilator. Concomitant use of other antihypertensives or antiarrythmic should be done cautiously. Although grapefruit juice can enhance the effect of CaCBs, this is not usually of clinical significance unless the patient is on the previously popular grapefruit fad diet.
### TABLE 65.24 Beta-Blockers

<table>
<thead>
<tr>
<th>[Beta-Blocker Class]</th>
<th>Generic (Trade) Name</th>
<th>Oral Dose (mg)</th>
<th>Indications and Unique Properties</th>
</tr>
</thead>
<tbody>
<tr>
<td>[Beta-1 Selective]</td>
<td><strong>Acebutolol (Sectral)</strong></td>
<td>200 bid–1,200 qd</td>
<td>HTN and ventricular arrhythmia, also angina (off-label); mild intrinsic β-mimetic properties</td>
</tr>
<tr>
<td></td>
<td><strong>Atenolol (Tenormin)</strong></td>
<td>50 qd–100 qd</td>
<td>IV dosing available for acute MI</td>
</tr>
<tr>
<td></td>
<td><strong>Betaxolol (Kerlone)</strong></td>
<td>10 qd–20 qd</td>
<td>HTN, angina. Also unapproved for migraine prophylaxis</td>
</tr>
<tr>
<td></td>
<td><strong>Bisoprolol (Zebeta)</strong></td>
<td>5 qd–20 qd</td>
<td>HTN. Also unapproved for migraine prophylaxis</td>
</tr>
<tr>
<td></td>
<td><strong>Metoprolol (Lopressor)</strong> (Toprol XL)</td>
<td>50 bid–400 daily</td>
<td>Angina, HTN, compensated CHF; IV dosing available for acute MI</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50–100 qd–400 daily</td>
<td>Also unapproved for migraine prophylaxis, HR control in AF, akathisia</td>
</tr>
<tr>
<td>[Beta (Nonselective)]</td>
<td><strong>Carteolol (Cartrol)</strong></td>
<td>2.5 qd up to 10 qd</td>
<td>HTN; mild intrinsic β-mimetic activity</td>
</tr>
<tr>
<td></td>
<td><strong>Carvedilol (Coreg, Coreg CR)</strong></td>
<td>CHF: 3.125 bid–50 bid, if over 85 kg</td>
<td>Compensated CHF, HTN, and LV dysfunction post-MI; exerts α-1 antagonistic effect</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HTN and LV dysfunction: up to 50 total daily dose</td>
<td>Also unapproved for chronic stable angina and idiopathic cardiomyopathy</td>
</tr>
<tr>
<td></td>
<td><strong>Labetalol (Trandate, Normodyne)</strong></td>
<td>100 bid–2,400 daily</td>
<td>HTN; α-1 antagonistic properties; IV dosing available for hypertensive emergencies</td>
</tr>
<tr>
<td></td>
<td><strong>Nadolol (Corgard)</strong></td>
<td>40 qd–320 daily</td>
<td>HTN and angina; also unapproved for migraine prophylaxis, ventricular arrhythmias, and Parkinsonian tremor</td>
</tr>
<tr>
<td></td>
<td><strong>Penbutolol (Levatol)</strong></td>
<td>20 qd–80 daily</td>
<td>HTN; mild intrinsic β-mimetic activity</td>
</tr>
<tr>
<td></td>
<td><strong>Pindolol (Visken)</strong></td>
<td>5 bid–60 daily</td>
<td>HTN; also angina (off-label); mild intrinsic β-mimetic activity</td>
</tr>
<tr>
<td></td>
<td><strong>Propranolol (Inderal)</strong> (Inderal-LA, Inno-Pran XL)</td>
<td>40 bid–640 daily 80 h to 120 h (ext. rel.)</td>
<td>HTN, angina, cardiac arrhythmias, MI, migraine prophylaxis, and hypertrophic subaortic stenosis (ext. rel.). Also unapproved for generalized anxiety and Parkinsonian tremors (ext. rel.)</td>
</tr>
<tr>
<td></td>
<td><strong>Sotalol (Betapace, Betapace AF)</strong></td>
<td>80 bid–640 daily</td>
<td>Ventricular arrhythmias, AF/flutter</td>
</tr>
<tr>
<td></td>
<td><strong>Timolol (Blocadren)</strong></td>
<td>10 bid–60 daily</td>
<td>HTN, post-MI, migraine headaches; angina (off-label)</td>
</tr>
</tbody>
</table>

### TABLE 65.25 Calcium Channel Blockers

<table>
<thead>
<tr>
<th>Generic (Trade) Name</th>
<th>Typical Oral Dose (mg)</th>
<th>Indications</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Amlodipine (Norvasc)</strong></td>
<td>2.5–5 qd–10 qd</td>
<td>HTN, stable and variant angina</td>
</tr>
<tr>
<td><strong>Diltiazem (Cardizem)</strong> (Cardizem CD, Dilacor, Dilacor CR, Tiazac)</td>
<td>20 qd–360 qd</td>
<td>HTN (extended preparations), stable and variant angina</td>
</tr>
<tr>
<td></td>
<td>60–120 bid–360 qd</td>
<td>Can be use IV for rapid atrial fibrillation</td>
</tr>
<tr>
<td></td>
<td>180–240 qd–540 qd</td>
<td></td>
</tr>
<tr>
<td><strong>Felodipine (Plendil)</strong></td>
<td>5 qd–10 qd</td>
<td>HTN</td>
</tr>
<tr>
<td><strong>Isradipine (DynaCirc, DynaCirc CR)</strong></td>
<td>2.5 bid (or 5 qd of CR)–20 qd</td>
<td>HTN; IV form for HTN emergencies</td>
</tr>
<tr>
<td><strong>Nicardipine (Cardene)</strong> (Cardene SR)</td>
<td>20 tid–120 qd</td>
<td>HTN (SR form), stable angina; IV form available for hypertensive emergencies</td>
</tr>
<tr>
<td></td>
<td>30 bid–120 qd (SR)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10 tid–120 qd (IR)</td>
<td>Angina</td>
</tr>
<tr>
<td><strong>Nifedipine (Procardia, Adalat)</strong> (Procardia XL, Adalat CC)</td>
<td>30–60 qd–120 qd (ER)</td>
<td>HTN (excessive decline in BP or even stroke may occur if IR form taken SL)</td>
</tr>
<tr>
<td><strong>Nimodipine (Nimotop)</strong></td>
<td>60 q4h for 21 d</td>
<td>Only indicated for reduction of cerebral vasospasm following subarachnoid hemorrhage; must start within 96 h of the bleed</td>
</tr>
<tr>
<td><strong>Nisoldipine (Sular)</strong></td>
<td>10–40 qd</td>
<td>HTN</td>
</tr>
<tr>
<td><strong>Verapamil (Isoptin, Calan)</strong> (Isoptin SR, Calan SR, Verelan, Covera HS)</td>
<td>80 tid–360 qd</td>
<td>HTN, can also be used IV for treatment of SVT, although adenosine is preferred for this indication</td>
</tr>
<tr>
<td></td>
<td>240 qd–480 qd</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Covera HS: 180–480 H/S</td>
<td></td>
</tr>
</tbody>
</table>
may require 1 week of regular dosage, unless a loading dose is given initially. Increased extracellular potassium (K+) concentrations decrease its tissue binding. Elimination is renal; thus, dose adjustments should be made if there is renal dysfunction.

**Preparations and dosing**

Maintenance digoxin (Lanoxin, Lanoxicaps, Digitek) doses are generally 0.125 to 0.25 mg qd.

**Relevant Side Effects and Drug Interactions**

Digitalis toxicity is a frequent occurrence, especially with unmonitored dosing and lack of a high enough index of suspicion for early signs and symptoms of toxicity. These signs often include anorexia, GI disturbances (nausea, vomiting, or diarrhea), fatigue, headache, drowsiness, and visual disturbances. Cardiac abnormalities can also occur, and with increasing toxicity, the risk for severe or even fatal rhythm disturbances increases. In fact, digitalis toxicity can mimic essentially any rhythm disturbance, particularly premature atrial or ventricular beats, paroxysmal atrial tachycardia, ventricular tachycardia, or high degree AV block.

Many drugs may affect the efficacy of digitalis. Certain diuretics and other medications may cause hypokalemia, which increases toxicity risk. Quinidine, verapamil, diltiazem, and amiodarone cause increased plasma digitalis concentrations. Beta adrenergic agonists may increase the risk of cardiac arrhythmias. In addition, drugs that induce hepatic microsomal enzyme activity may increase digitalis metabolism.

**Diuretics**

**Relevance to Physiatry**

Although diuretics are primarily used for HTN and certain forms of edema, other physiologic uses have included the acute treatment of immobilization hypercalcemia (e.g., in SCI) (427). Diuretics may also facilitate fluid mobilization in early lymphedema but should not be used for chronic treatment of this condition (428). Diuretic-associated orthostatic hypotension is a potential problem, particularly for transfer and gait training.

**Mechanism of Action**

Diuretics in general act by decreasing the volume of extracellular fluid by enhancing urinary excretion through sodium resorption. Because this increases the tubular osmotic gradient, water also is resorbed into the tubules and subsequent diuresis occurs. Each of the three most frequently utilized diuretic classes specifically act in different areas of the tubules to achieve this effect. The thiazides and high-ceiling (or “loop diuretic”) classes are more vigorous in their diuretic, natriuretic, and potassium-wasting effects. Potassium-sparing diuretics are less vigorous in action and because they interfere with the normal sodium-potassium exchange in the distal tubule, potassium tends to be spared from increased secretion into the urine. A less commonly used class is the carbonic anhydrase inhibitors. These alkalinate the urine through increased bicarbonate concentration and, as a result, cause sodium and potassium ions to be excreted. Additional agents with diuretic properties include osmotic agents such as mannitol and methylxanthines (e.g., theophylline).

**Nitrates**

**Relevance to Physiatry**

Organic nitrates are commonly used in inpatient rehabilitative settings where there are many patients with coronary artery disease. In addition to their use in angina and in certain patients with CHF, oral or TD nitrates may help to reduce acute BP elevations, such as in autonomic dysreflexia. The physiatrist should also be well aware of possible deleterious effects in those undergoing intensive rehabilitation. In particular, it is important to prevent orthostatic hypotension and tachycardia through measures such as gradual position changes, especially in the morning, and having patients avoid prolonged standing if lower extremity strength is diminished.

**Mechanism of Action and Pharmacokinetics**

After being converted into nitrous oxide within the body, organic nitrates cause smooth muscle relaxation. This occurs by stimulation of intracellular synthesis of cyclic guanosine monophosphate (GMP), which leads to myosin dephosphorylation and loss of smooth muscle contractility.

After oral ingestion, organic nitrates undergo first-pass hepatic inactivation via hydrolysis. Sublingual administration of organic nitrates leads to an onset of action in 1 to 2 minutes and a rapid decrease in effects thereafter. When given orally, higher doses must be used to saturate the hepatic metabolism and prevent their degradation. Subsequently, they have a slow onset of action, with peak effects occurring at 60 to 90 minutes, and a duration lasting from 3 to 6 hours. In contrast, TD application, either by ointment or patch, allows for gradual absorption and prolonged delivery. With the former, the initial effects occur within 60 minutes and last 4 to 8 hours. In dosing by patch, peak effects are achieved after 1 to 2 hours and can last up to 24 hours. It is important to note, however, that long-term continuous use of organic nitrates can lead to the development of tolerance to them. This can be avoided by ensuring a nitrate-free interval for short periods of time (e.g. removing a nitrate patch daily at bedtime and reapplying a new one in the morning).

The efficacy of nitrates in relieving angina pectoris lies in their ability to decrease venous return to the heart (preload) and systemic peripheral resistance (afterload), thereby reducing myocardial oxygen demand. Because they also affect nonvascular smooth muscle, they may relieve atypical chest pain, which can occur in spasmodic conditions of the esophagus or bile ducts.

**Dosage and Preparations**

Multiple organic nitrate preparations exist. Sublingual, IV, and inhalant preparations are used in acute anginal episodes, whereas buccal and SR oral tablets, ointments, and patches are preferred for prophylaxis against anginal episodes.

**Relevant Side Effects and Drug Interactions**

See Table 65-26.

Potential medication interactions should be considered. These include cardiac toxicity from digitalis due to diuretic-induced hypokalemia, decreased renal lithium clearance, interference with oral hypoglycemics, and increased risk of NSAID-induced renal failure.
Preparations and Dosing

See Table 65-27.

GI MEDICATIONS

Antiemetics

Relevance to Physiatry

Nausea is a troublesome symptom that frequently prevents patients from meaningfully participating in a rehabilita-

Relevant Side Effects and Drug Interactions

The most common side effect is vascular headache that can occasionally be severe. It usually improves, however, with reduction of the dosage and continued use. Postural hypotension may occur but can be relieved by maneuvers that increase central venous return, such as lying supine with leg elevation or with true Trendelenburg positioning. Concomitant use of nitrates with alcohol or antihypertensives can lead to significant hypotension.

### TABLE 65.26 Commonly Prescribed Diuretics

<table>
<thead>
<tr>
<th>[Class]</th>
<th>Generic (Trade) Name</th>
<th>Dose (mg)</th>
<th>Other Characteristics and Side Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thiazide</td>
<td>Chlorthiazide (Diurol)</td>
<td>250–500 PO qd/bid</td>
<td>Avoid in patients with allergies to sulfonamide-containing drugs</td>
</tr>
<tr>
<td></td>
<td>Chlorthalidone (Hygroton)</td>
<td>25–100 PO qd</td>
<td>Available in a suspension form</td>
</tr>
<tr>
<td></td>
<td>Hydrochlorothiazide (Aquaize, E-zide, HCTZ, Esidrix, Microzide, Oretic, Hydrodiuril)</td>
<td>25–200 PO qd</td>
<td>Rarely used</td>
</tr>
<tr>
<td></td>
<td>Indapamide (Lozol)</td>
<td>1.25–5.0 PO qd</td>
<td>Lower doses effective in antihypertensive combinations; can be used less frequently than daily for maintenance edema management</td>
</tr>
<tr>
<td></td>
<td>Methylcloothiazide (Enduron, Aquatensen)</td>
<td>2.5–10 PO qd</td>
<td>An indoline; main use is for HTN</td>
</tr>
<tr>
<td></td>
<td>Metolazone (Zaroxlyn)</td>
<td>5–20 PO qd</td>
<td>Maintenance dosage for edema may be q.o.d. or 3–5 d/wk</td>
</tr>
<tr>
<td>Loop</td>
<td>Bumetanide (Bumex)</td>
<td>0.5–2.0 PO qd, 5–1 IV/IM</td>
<td>Used with loop diuretics in edema refractory to loop diuretics alone</td>
</tr>
<tr>
<td></td>
<td>Ethacrynic acid (Edecrin)</td>
<td>25–100 PO qd/bid, 5–1 mg/kg IV up to 50</td>
<td>Used for rigorous diuresis and in edema with decreased renal function. 1 mg equivalent to 40 mg Lasix</td>
</tr>
<tr>
<td></td>
<td>Furosemide (Lasix)</td>
<td>20–80 PO qd/bid, 1 mg/kg IV up to 20–40</td>
<td>Most commonly used loop diuretic</td>
</tr>
<tr>
<td></td>
<td>Torsemide (Demadex)</td>
<td>5–20 PO/IV qd</td>
<td>Equal efficacy with oral or IV doses</td>
</tr>
<tr>
<td></td>
<td>Potassium sparing</td>
<td></td>
<td>Possible hyperkalemia, especially if used with ACEIs or ARBs</td>
</tr>
<tr>
<td></td>
<td>Amiloride (Midamor)</td>
<td>5–10 PO qd</td>
<td>Rarely used; lack mineralocorticoid side effects</td>
</tr>
<tr>
<td></td>
<td>Spironolactone (Aldactone)</td>
<td>25–50 PO qd/bid</td>
<td>Aldosterone antagonist</td>
</tr>
<tr>
<td></td>
<td>Triamterene (Dyrenium)</td>
<td>100 PO bid</td>
<td>Rarely used; lack mineralocorticoid side effects</td>
</tr>
</tbody>
</table>

### TABLE 65.27 Nitrates

<table>
<thead>
<tr>
<th>Generic (Trade) Name</th>
<th>Dose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Isosorbide dinitrate</td>
<td>10–40 mg PO at least 6 h apart</td>
</tr>
<tr>
<td>Tablets (Isordil, Sorbitrate)</td>
<td>1 tab (2.5–10 mg) SL prn</td>
</tr>
<tr>
<td>Sublingual tablets (Isordil, Sorbitrate)</td>
<td>5–40 mg tid; max of 80 mg PO bid</td>
</tr>
<tr>
<td>Sustained release (Isochron, Dilatrate SR)</td>
<td></td>
</tr>
<tr>
<td>Isosorbide mononitrate</td>
<td>20 mg PO bid at least 7 h apart</td>
</tr>
<tr>
<td>Tablets (ISMO, Monoket)</td>
<td>30–240 mg PO qd (start with 30–60 PO qd)</td>
</tr>
<tr>
<td>Extended release (Imdur)</td>
<td>(15 mg/inch)</td>
</tr>
<tr>
<td>Nitroglycerin</td>
<td>0.5 inch q8h, maintenance 1–2 inch q8h; max. 4 inch q4–6h</td>
</tr>
<tr>
<td>Ointment 2% (Nitro-Bid, Nitrol)</td>
<td>1–2 SL sprays prn</td>
</tr>
<tr>
<td>Spray (Nitroloinguin, Nitro Mist)</td>
<td>0.4 mg SL prn</td>
</tr>
<tr>
<td>Sublingual (Nitrostat, NitroQuick)</td>
<td>1 patch (doses of 0.1 to 0.8 mg/h) 12–14 h daily</td>
</tr>
<tr>
<td>TD (Depontit, Minitrans, Nitro-Dur, Nitrodisc, Transderm-Nitro)</td>
<td>1 cap inhaled (2 or more times) for rapid relief of angina</td>
</tr>
<tr>
<td>Amyl nitrite</td>
<td></td>
</tr>
</tbody>
</table>
 CHAPTER 65 | PHARMACOTHERAPY OF DISABILITY

|TABLE 65.28 Antiemetics |

<table>
<thead>
<tr>
<th>Medication Class and Mechanism of Action</th>
<th>Generic (Trade) Name</th>
<th>Typical Dose (mg)</th>
<th>Other Properties and Side Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Phenothiazine derivatives:</strong> act at the CTZ and the vomiting center</td>
<td>Prochlorperazine (Compazine)</td>
<td>5–10 PO/IM tid–qid 25 PR q12h IV dosing possible</td>
<td>Antidopaminergic: extrapyramidal reactions and neuroleptic malignant syndrome possible</td>
</tr>
<tr>
<td></td>
<td>Promethazine (Phenergan)</td>
<td>1/kg up to max 25–50 PO/IM/PR q4–6</td>
<td>H-1 receptor blocker, therefore antihistaminergic; much less effect on DA; IV form also available</td>
</tr>
<tr>
<td></td>
<td>Thiethylperazine (Torecan)</td>
<td>10 PO/IM q8</td>
<td>Same as for prochlorperazine</td>
</tr>
<tr>
<td><strong>Antihistamines:</strong> act at the CTZ and the vomiting center</td>
<td>Dimenhydrinate (Dramamine)</td>
<td>10 PO/IM 50 PO/IM/IV q4–6</td>
<td>Sedating; TD preparation no longer commercially available</td>
</tr>
<tr>
<td></td>
<td>Trimethobenzamide (Tigan)</td>
<td>250 PO q6–8 200 IM/PR q6–8</td>
<td>Hypersensitivity and Parkinson-like symptoms have been reported</td>
</tr>
<tr>
<td></td>
<td>Scopolamine (Transderm Scop)</td>
<td>One patch behind ear, change after 3 d prn</td>
<td>Belladonna alkaloid, therefore anticholinergic side effects</td>
</tr>
<tr>
<td><strong>Other:</strong> act at the CTZ and vomiting center</td>
<td>Dolasetron (Anzemet)</td>
<td>Up to 100 PO single dose (alternate IV dosing)</td>
<td>Prevention of chemotherapy-related or postop nausea</td>
</tr>
<tr>
<td></td>
<td>Granisetron (Kytril)</td>
<td>1 PO bid × 1 d 2 PO single dose (alternate IV dosing)</td>
<td>Prevention of chemotherapy-related or radiation therapy-related nausea</td>
</tr>
<tr>
<td></td>
<td>Ondansetron (Zofran)</td>
<td>8 PO before and 8 h after 8 PO tid (alternate IV dosing)</td>
<td>Prevention of chemotherapy-related, RT-related and postop nausea</td>
</tr>
<tr>
<td><strong>Cannabinoids:</strong> complex CNS effects, including sympathomimetic</td>
<td>Dronabinol (Marinol)</td>
<td>Dosage varies</td>
<td>Chemotherapy-related nausea or anorexia associated with AIDS</td>
</tr>
<tr>
<td><strong>Selective serotonin 5-HT3 receptor inhibitors:</strong> blocks receptors in peripheral vagus nerve terminals and CTZ</td>
<td>Dolasetron (Anzemet)</td>
<td>Up to 100 PO single dose (alternate IV dosing)</td>
<td>Prevention of chemotherapy-related or postop nausea</td>
</tr>
<tr>
<td></td>
<td>Granisetron (Kytril)</td>
<td>1 PO bid × 1 d 2 PO single dose (alternate IV dosing)</td>
<td>Prevention of chemotherapy-related or radiation therapy-related nausea</td>
</tr>
<tr>
<td></td>
<td>Ondansetron (Zofran)</td>
<td>8 PO before and 8 h after 8 PO tid (alternate IV dosing)</td>
<td>Prevention of chemotherapy-related, RT-related and postop nausea</td>
</tr>
</tbody>
</table>

Medications that Act by Affecting GI Motility

Relevance to Physiatry

Abnormally slow GI transit can especially occur in diabetic patients. Since physiatrists commonly manage patients who also have diabetes, knowledge of promotility agents is important. The two most commonly used ones are metoclopramide (Reglan) and cisapride (Propulsid). Use of the latter is limited because of many drug interactions and potentially fatal cardiac dysrhythmias. Metoclopramide is dosed for gastroesophageal reflux and diabetic gastroparesis at 10 to 15 mg orally 30 minutes before meals and sleep, and should be reduced to 5 mg in elderly patients. It can be used for prevention of chemotherapy-induced emesis and is available for IV or IM dosing. Because it is a DA receptor antagonist, CNS side effects and extrapyramidal reactions can occur.

Mechanism of Action and Pharmacokinetics

These drugs enhance acetylcholine (Ach) release and/or action at the myenteric plexus, thereby stimulating upper GI peristalsis. This leads to reflux reduction and gastric emptying. Metoclopramide undergoes hepatic conjugation, but otherwise little hepatic metabolism.

Antidiarrheal Drugs

Relevance to Physiatry

Hospitalized patients commonly complain of diarrhea. This condition obviously can interfere with rehabilitation. Therefore, medications are often used for short-term management of acute, nonspecific diarrhea and may also be used for chronic inflammatory bowel disease–associated diarrhea. They should be avoided, however, when diarrhea is suspected to be the result of obstructive jaundice, fecal impaction, or infection with toxin production. If diarrhea is not controlled within 48 hours, further workup is warranted.

The two most commonly used antidiarrheal agents for hospitalized patients are loperamide (Imodium) and diphenoxylate or its principal metabolite difenoxin, combined with atropine (Lomotil and Motofen, respectively).

Mechanism of Action and Pharmacokinetics

As they are opioid derived, antidiarrheal drugs directly inhibit intestinal smooth muscle peristalsis.
Preparations and Dosing
See Table 65-29.

Side Effects and Drug Interactions
As is expected from medications that impair intestinal motility, abdominal discomfort, and nausea can occur. In addition, opioid-like side effects such as fatigue, drowsiness, or dry mouth can occur. Lomotil and Motofen are chemically related to the narcotic analgesic meperidine (Demerol) and are combined with atropine. They may therefore interact with MAO inhibitors and are likely to cause sedation when used with other CNS depressants. Lomotil and Motofen may also prolong the biological half-lives of drugs for which the elimination rate is dependent upon microsomal drug metabolizing enzyme systems. Although Imodium is somewhat chemically similar to meperidine, it is not combined with atropine and therefore has a safer side-effect profile and less drug interactions.

Medications that Reduce Gastric Acid Secretion
Relevance to Physiatry
Besides their use in the prevention and treatment of NSAID-induced GI pathology (see NSAID section of this chapter), these medications are also used for idiopathic gastric and duodenal ulcers, as well as for gastroesophageal reflux disease. They are also often used to prevent stress ulcers and are thus encountered in patients undergoing rehabilitation for stroke, TBI, or SCI. Other than cimetidine that can cause CNS changes in the elderly, they are usually well tolerated and do not interfere with rehabilitation.

Mechanism of Action and Pharmacokinetics; Preparations and Dosing; Relevant Side Effects and Drug Interactions
See Table 65-30.

### TABLE 65.29 Antidiarrheal Agents

<table>
<thead>
<tr>
<th>Agent</th>
<th>Dosage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kapectate®</td>
<td>2 tabs or 30 mL pm, up to 8 doses in 24 h</td>
</tr>
<tr>
<td>Imodium®</td>
<td>4 mg initially, then 2 mg s/p loose bowel movement to a maximum of 16 mg/24 h</td>
</tr>
<tr>
<td>Lomotil</td>
<td>2 tablets qid or 10 cc qid</td>
</tr>
<tr>
<td>Motofen</td>
<td>2 tabs initially, then 1 tab prn to a maximum of 8 tabs/24 h</td>
</tr>
</tbody>
</table>

*a Also available OTC.

### TABLE 65.30 Medications that Act by Reducing Gastric Secretion

<table>
<thead>
<tr>
<th>Medication Class: Mechanism</th>
<th>Generic (Trade) Name</th>
<th>Usual Oral Dose (mg)</th>
<th>Unique Characteristics and Side Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>H-2 blockers: Block histamine-2 receptors of gastric acid-producing parietal cells</td>
<td>Cimetidine (Tagamet)</td>
<td>800 h or 400 bid</td>
<td>Because of the change in gastric fluid acidity due to H₂ blockers, GI absorption of various drugs can be altered; IV form also available; replaced by newer agents due to side effects (esp. in elderly and in impaired renal function) and medical interactions</td>
</tr>
<tr>
<td></td>
<td>Famotidine (Pepcid)</td>
<td>40 h or 20 bid</td>
<td>IV form also available</td>
</tr>
<tr>
<td></td>
<td>Nizatidine (Axid)</td>
<td>300 h or 150 bid</td>
<td>IV form also available</td>
</tr>
<tr>
<td></td>
<td>Ranitidine (Zantac)</td>
<td>300 h or 150 bid</td>
<td></td>
</tr>
<tr>
<td>Proton pump inhibitors: decrease gastric acid secretion by inhibiting the parietal cell membrane enzyme that actively transports hydrogen ions out of the cell</td>
<td>Esomeprazole (Nexium)</td>
<td>20 to 40 qd</td>
<td>Numerous indications, including gastroesophageal reflux disease, gastritis, peptic ulcer disease, erosive esophagitis, and prevention/treatment of NSAID or ASA-induced ulcers. Well tolerated but sometimes nausea, diarrhea, headache, dizziness, and possible gastric bacterial colonization from elevation of gastric fluid PH; inhibition of metabolism of medications processed by cytochrome P-450 system.</td>
</tr>
<tr>
<td></td>
<td>Lansoprazole (Prevacid)</td>
<td>15 to 30 qd</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Omeprazole (Prilosec, Zegerid, Rapinex)</td>
<td>20 to 40 qd</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pantoprazole (Protonix)</td>
<td>40 qd</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Rabeprazole (Aciphex)</td>
<td>20 qd</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Misoprostol (Cytotec)</td>
<td>100–200 ug qid (depending on indication)</td>
<td>Diarrhea very common; not to be used in women of child-bearing age due to abortifacient properties from smooth muscle contraction; no known drug interactions</td>
</tr>
</tbody>
</table>
CHAPTER 65 | PHARMACOTHERAPY OF DISABILITY

Laxatives

Relevance to Physiatry
As is true for diarrhea, constipation can also have a significant negative impact upon rehabilitation. Laxatives can facilitate or restore regular bowel movements in constipated patients. Some are used as part of a bowel program in neurologically impaired patients, especially those with SCI (323,324). Many preparations are available (Table 65-31) and their effects range from mild (bulking agents) to more aggressive (hyperosmolar preparations). Caution should be exercised in those with nausea, vomiting, or unexplained abdominal pain.

<table>
<thead>
<tr>
<th>Medication Class:</th>
<th>Generic (Trade) Name</th>
<th>Common Doses</th>
<th>Unique Characteristics and Side Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bulking agents:</td>
<td>Psyllium (Metamucil, Fiberall); methylcellulose (Citrucel); polycarbophil (FiberCon, Fiberall)</td>
<td>1 tbsp PO qd–tid; 1 tsp PO qd–tid; 1 gm PO qid prn</td>
<td>Gritty texture prohibits use in some patients</td>
</tr>
<tr>
<td>Colonic stimulants: (“irritant cathartics”)</td>
<td>Bisacodyl (Dulcolax); cascara sagrada (Cascara); castor oil (Purge); sodium biphosphate (Fleet enema, Fleet Phospho-soda, Fleet Bisacodyl supp., Fleet Bisacodyl tabs); senna (Senokot) tablets (8.6 mg/tab) or syrup (8.8 mg teaspoon)</td>
<td>5–15 mg PO/PR pm; 325 mg or 5 mL PO qd; 15–30 mL PO qd; 1 PR; 45 mL; 1 PR; 4 tabs; 2–4 tabs PO qd–bid; 10–15 mL PO qd</td>
<td>Dulcolax affects fluid and electrolyte absorption throughout the intestine; Cascara and Senokot act on the large intestine, so there is a delayed effect of approximately 8 h and may cause urine discoloration; Castor oil intestinally hydrolyzed to a cathartic (ricinoleic acid)</td>
</tr>
<tr>
<td>Combination products</td>
<td>Docusate/casanthranol (PeriColace) senna/docusate (Senokot-S tabs) Milk of Magnesia-Cascara</td>
<td>1–2 caps qd 2, max 4 PO qd 15–30 mL qd</td>
<td>Pericolace capsule, 100 mg docusate/30 mg casanthranol Pericolace liquid, 60 mg docusate/20 mg casanthranol per 15 mL</td>
</tr>
<tr>
<td>Hyperosmolar agents: (“bulk cathartics”)</td>
<td>Magnesium citrate soda; magnesium hydroxide (Milk of Magnesia); Polyethylene glycol (MiraLax, Glycolax) Lactulose (Cephulac, Enulose, Constulose)</td>
<td>200–300 mL PO; 15–60 mL PO 17 grams of powder qd in eight ounces of water 15 to 30 mL (10 to 20 g) up to 60 mL qd</td>
<td>Because of their osmotic concentration, hyperosmolar agents should be avoided in acute CHF and used with caution with impaired renal function, heart disease, and electrolyte imbalances; Fleet products usually used as bowel preps 3 h before a procedure</td>
</tr>
<tr>
<td>Miscellaneous agents</td>
<td>Glycerin suppositories; lactulose (Chronulac); mineral oil;</td>
<td>1 PR; 15–30 mL or 10–20 gm powder qd; 15–45 mL PO or 120 mL PR</td>
<td>Glycerin stimulates rectal contraction within 15–30 min via hyperosmotic and irritant actions; lactulose is a synthetic disaccharide converted by colonic bacteria to short-chain organic acids that cause colonic fluid accumulation; mineral oil softens and lubricates stool and is useful for fecal impaction; MiraLax</td>
</tr>
<tr>
<td>Stool softeners: surfactant, thus increases stool H2O absorption as it descends the lower GI tract</td>
<td>Docusate sodium (Colace) (capsules or syrup/liquid); docusate calcium (Surfak, Sulfolax)</td>
<td>100 mg PO qd–tid; 240 mg PO qd</td>
<td>May not see the effect until 1–3 d; begin with the highest dose and then reduce when the first bowel movement occurs</td>
</tr>
</tbody>
</table>

Mechanism of Action and Pharmacokinetics; Preparations and Dosing; Relevant Side Effects and Drug Interactions
See Table 65-31.

HYPNOTICS

Relevance to Physiatry
Insomnia is frequently associated with major stressors such as injury or illness or conditions such as fibromyalgia, myofascial pain syndromes, and chronic fatigue syndrome (429). It can
significantly interfere with rehabilitation due to resultant daytime lethargy, irritability, difficulty with concentration, etc. Hence, treatment of insomnia with hypnotics can play an important role in overall patient management. Hypnotics are commonly used in hospitalized patients in acute care, in rehabilitation units, and rehabilitation hospitals.

**Mechanism of Action and Pharmacokinetics**

BNZs exert hypnotic and anxiolytic effects by binding nonselectively to the GABA-benzodiazepine (B2) receptor complex, particularly in the limbic system, thalamus, and hypothalamus. In doing so, they reduce delta sleep. They are well absorbed, undergo very little first-pass metabolism, are highly metabolized by the liver, are excreted with short elimination half-lives, and their metabolites do not accumulate.

Zolpidem tartrate (Ambien) is structurally unrelated to other hypnotics. It preferentially binds to the omega-1 subunit of the GABA-BZ receptor complex, which is the purported reason for its selective hypnotic effect yet lack of anticonvulsant, anxiolytic, and muscle relaxant properties seen with BNZs. Sleep architecture is relatively preserved in both healthy patients and in those with acute or chronic insomnia. It is rapidly absorbed and has a mean elimination half-life of 2½ hours. It does not accumulate if used for a short term and thus is less likely to cause daytime sedation.

**Preparations and Dosing**

See Table 65-32.

**Relevant Side Effects and Drug Interactions**

Daytime drowsiness, headache, and fatigue are the most frequently reported side effects in controlled clinical studies involving bedtime BNZ use. Dependence and rebound insomnia can develop, particularly if used regularly for more than a few weeks. These drugs should be used with caution in patients taking other CNS depressants.

Ambien’s side-effect profile is superior but still includes daytime drowsiness usually during short-term use and dizziness or “drugged” feelings with long-term use. As with BNZs, it should be used cautiously in patients taking other CNS depressants.

**HYPOGLYCEMICS**

**Relevance to Physiatry**

A physiatrist often manages patients who have diabetes mellitus as a comorbid medical condition. Proper glucose control is important both acutely, in order to avoid uncontrolled blood sugar elevations or hypoglycemia, and chronically in order to avoid long-term complications (430, 431). The effects of exercise on blood glucose levels must be taken into account especially in patients who have not previously exercised regularly.

**Insulin**

**Mechanism of Action and Pharmacokinetics**

Insulin is a large polypeptide secreted by pancreatic β cells that lowers blood glucose concentrations by inhibiting hepatic gluconeogenesis and by increasing peripheral glucose uptake and metabolism. It also affects lipid, ketone, and protein metabolism by reducing glycerol and free fatty acid blood concentrations, reducing ketone body formation, and increasing protein synthesis in muscles and other tissues. Insulin’s various effects are exerted via interaction with highly specific receptors various target cell membranes. This interaction leads to intracellular mechanisms that produce target cell-specific end organ effects. Exogenous insulin preparations can be of human, porcine, or bovine origin. The latter two differ from human insulin by one and three amino acids, respectively, within the peptide chain.

**Preparations and Dosing**

As shown in Table 65-33, insulin preparations can be classified as short acting, intermediate acting, and long acting. Most dosing regimens involve a mixture of short and intermediate-acting insulin preparations given at fixed dosages prior to breakfast and dinner. Another method of dosing insulin is frequent monitoring of blood glucose levels and adjusting
Preparations and Dosing
It is given prior to major meals. In type I diabetes, it is initiated at 15 mcg subcutaneously and titrated to a maintenance dose of 30 or 60 mcg as tolerated. In type II diabetes, the starting dose is 60 mcg and increased to a dose of 120 mcg as tolerated. When initiating therapy with pramlintide, it is crucial to decrease insulin dosing initially to reduce the risk of insulin-induced hypoglycemia.

Side Effects and Drug Interactions
Although pramlintide alone does not cause hypoglycemia, it increases the risk of insulin-induced severe hypoglycemia, particularly in patients with type I diabetes.

Exenatide (Byetta)
Mechanism of Action and Pharmacokinetics
Sharing a similar amino acid sequence with glucagon-like peptide-1 (GLP-1), exenatide binds to the GLP-1 receptor, leading to an increase in both synthesis and secretion of insulin from pancreatic β cells in the presence of elevated blood glucose concentrations. It also lowers serum glucagon levels in the presence of hyperglycemia, slows gastric emptying, and reduces food intake. It is used as adjunctive therapy to improve glycemic control in patients with type II diabetes mellitus who are taking either or a combination of metformin, a sulfonylurea, or a thiazolidinedione but have not achieved adequate glycemic control.

Preparations and Dosing
Initially 5 mcg subcutaneously before breakfast and dinner. Can be increased to 10 mcg twice daily.

Side Effects and Drug Interactions
Exenatide, in combination with oral hypoglycemic medications, particularly a sulfonylurea, may have an increased risk of hypoglycemia. Its use is commonly associated with GI adverse reactions, including nausea, vomiting, and diarrhea. It is not recommended for use in patients with severe renal impairment. It can reduce the efficacy of acetaminophen, digoxin,
lovastatin, and certain oral antibiotics. It can lead to an increase in INRs in patients taking warfarin.

**Oral Hypoglycemics**

**Mechanism of Action and Pharmacokinetics**

Although the exact mechanisms of action vary somewhat according to the chemical class as shown in Table 65-34, all depend upon a certain degree of endogenous insulin production; thus, they are not effective in treating type 1 diabetes. All are well absorbed from the GI tract.

### Preparations and Dosing, Relevant Side Effects, and Drug Interactions

See Table 65-34.

---

**TABLE 65.34 Oral Hypoglycemic Agents**

<table>
<thead>
<tr>
<th>Class: Mechanism of Action</th>
<th>Generic (Trade) Name</th>
<th>Usual Oral Dose (mg) (Starting Dose/Maximum)</th>
<th>Unique Characteristics and Side Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sulfonylureas (first generation): stimulate insulin release from pancreatic β cells and increase the insulin sensitivity of various peripheral tissues</td>
<td>Chlorpropamide (Diabenese); tolbutamide (Oramide, Orinase); tolazamide (Tolamide, Tolinase)</td>
<td>100–250/1,000 qd; 250–500/3,000 qd (daily or divided into two doses); 100–250/1,000</td>
<td>Chlorpropamide has a very long duration - up to 24 h, making it relatively dangerous, esp. in the elderly; tolbutamide has a duration up to 10 h</td>
</tr>
<tr>
<td>Sulfonylureas (second generation): Same mechanism as earlier</td>
<td>Glimepiride (Amaryl); glipizide (Glucotrol, Glucotrol XL); glyburide (DiaBeta, Micronase, Glynase)</td>
<td>1–2/8 qd; 5/40 qd (divided into two daily doses if &gt;15/d is required, max dose of XL 20 qd); 2.5–5/20 qd</td>
<td>Starting doses of glipizide and glyburide 2.5 and 1.25, respectively, recommended in elderly; duration up to 10 h</td>
</tr>
<tr>
<td>Biguanides: improve glucose tolerance in patients with NIDDM without inducing hypoglycemia: lowers both basal and postprandial glucose levels by decreasing hepatic glucose production, decreasing intestinal glucose absorption, and improving insulin sensitivity</td>
<td>Metformin (Glucophage, Glucovance, Glucophage XL)</td>
<td>500 qd or bid/2,560 qd (higher doses usually divided into bid or tid)</td>
<td>May cause potentially serious lactic acidosis in those with abnormal renal or hepatic function; thus, it is contraindicated in these patients and should be temporarily held in patients undergoing radiologic studies involving injectable contrast agents and in those who are dehydrated</td>
</tr>
<tr>
<td>Glucosidase inhibitors: decrease the digestion of ingested carbohydrates, thereby reducing postprandial hyperglycemia, possess no systemic effects</td>
<td>Miglitol (Glyset); acarbose (Precose)</td>
<td>Both started at 25 tid; individualized dosing to max of 100 tid</td>
<td>Metabolized in the GI tract; contraindicated in patients with inflammatory bowel disease or intestinal obstruction</td>
</tr>
<tr>
<td>Meglitinide: stimulates insulin release</td>
<td>Repaglinide (Prandin)</td>
<td>0.5–2 tid/max of 16 qd in divided doses (tid)</td>
<td>Most common adverse effect is hypoglycemia</td>
</tr>
<tr>
<td>Other agent, amino acid derivative having similar action to meglitinide</td>
<td>Nateglinide (Starlix)</td>
<td>120 tid</td>
<td>Most common adverse effect is hypoglycemia; to be used with caution in presence of liver dysfunction</td>
</tr>
<tr>
<td>Thiazolidinediones: decrease insulin resistance in peripheral tissues by stimulating the peroxisome proliferator-activated receptor</td>
<td>Pioglitazone (Actos); rosiglitazone (Avandia)</td>
<td>15–30/45 qd; 4/8 qd</td>
<td>May cause fluid retention, thus to be used cautiously in patients with CHF or edema; liver function should be monitored</td>
</tr>
<tr>
<td>Dipeptidyl peptidase-4 inhibitor: believed to act on the slowing the inactivation of incretin hormones such as glucagon-like peptide-1 and glucose-dependent insulino-tropic polypeptide</td>
<td>Sitagliptin (Januvia)</td>
<td>100 qd</td>
<td>Used in type II diabetes alone or as in combination with metformin or a peroxisome proliferator-activated-receptor agonist. Dosage reduction is necessary in moderate to severe renal impairment</td>
</tr>
</tbody>
</table>
RESPIRATORY MEDICATIONS

Decongestants, Expectorants, and Mucolytics

Relevance to Physiatry

Respiratory tract disorders and infections are commonly encountered in rehabilitation patients. Various medications are used in managing these conditions (Table 65-35). Decongestants treat upper respiratory congestion and the increased mucosal secretion that occurs in colds, seasonal allergies, or infections. Particularly when combined with postural drainage and percussion, expectorants and mucolytics are useful in facilitating pulmonary toilet through improving the quality and expulsion of mucus. Selective β-2-adrenergic agonists are useful in the treatment of reversible airway obstruction in bronchial asthma or COPD. Their use in exercise-induced asthma also has direct physiatric applications. Anticholinergic medications can be of benefit in relieving bronchoconstriction, particularly in COPD. Finally, inhaled glucocorticoids and leukotriene inhibitors, a newer medication class, are effective in maintenance therapy of asthma.

TBI MEDICATIONS

Each year more than 1.5 million Americans survive a TBI, among which 80,000 to 90,000 are left with permanent disability each year. An estimated 5.3 million people in the United States live with long-term disability from TBI (432). Currently, there are no medications that are FDA approved to treat the early or late sequelae of TBI. The complex nature of the injury and recovery, along with previous research failures, has created pharmaceutical industry apathy. Medications used to treat the behavioral and cognitive consequences of TBI are often based on success in other disorders with similar symptoms. Most TBI population neuropharmacology studies are just case reports.

The TBI manifestations for which medications are most commonly used include disorders of memory, arousal, attention, concentration, initiation, mood, and behavior. Before drug therapy is started, a careful review of the patient’s current medications is warranted since medications commonly used for other medical conditions may cause or exacerbate the TBI manifestation (Table 65-36). In general, non-centrally acting medications should be used since these are less likely to have cognitive side effects. Only after eliminating medications with these unwanted side effects and potential medication interactions should initiation of new drugs be considered.

Cognitive and behavioral impairment in the patient with TBI results from neurochemical alterations. These alterations are caused by shear injury of the ascending and descending cortical pathways to the brainstem, along with hypoxic lesions, contusions, and hemorrhages (433). More specific understanding of the pathophysiology of TBI on a cellular and neurochemical level will allow clinicians to tailor the pharmacologic treatment to individual patients. The following will describe those agents more commonly utilized in the patient with TBI categorized by predominant neurochemical activity.

Agents that Act on Serotonin

In addition to their use in neuropathic pain and depression discussed elsewhere in this text, SSRIs are also used in the setting of TBI to treat sleep disturbance, agitation, affective disorders, emotional incontinence, aggression, and impaired arousal (434,435). SSRIs are preferred to TCAs in the TBI population since they have fewer cognitive side effects. The use of SSRIs in TBI is directly related serotonin’s multiple roles in brain functioning including sleep initiation, sexual response, movement, mood, aggression, anxiety, appetite, and addictive behaviors (436).

Trazodone

Trazodone is specifically used in TBI patients for agitation (dosed 25 to 100 mg q6h around the clock or prn) and for sleep disturbance (initiated at 25 mg H/S and titrated up to 300 mg to achieve 6 to 8 hours of sleep) (437). Its role in sleep disorders is particularly important as insomnia occurs in more than 50% of TBI patients both early and late after injury and may exacerbate disability by further impairing arousal, attention and increasing one’s susceptibility to confusion and agitation (438).

SSRIs

These SSRIs are commonly used for other cognitive/behavioral manifestations besides sleep disturbance. Sertraline shows the best evidence for depression treatment effectiveness and tolerability among the SSRIs. Dosages are the same as recommended for depression. Slow titration and clinical monitoring is necessary to avoid apathy, an SSRI side effect (436).

Agents that act on Catecholamines

Cerebral catecholamines are produced in the sympathetic ganglia from the amino acid tyrosine. DA is enzymatically converted to NE and then to epinephrine. An increase in any one of the catecholamines increases the others. DA cell bodies are primarily in the substantia nigra and hypothalamus. There are five types of DA receptors, each with specialization of action. DA neuronal connections have important effects on attention, arousal, and memory (439). NE CNS cell bodies are found in the locus ceruleus. Although this small nucleus contains only several hundred neurons, it sends axons to all CNS regions. Catecholamines are thought to play an important, although not fully defined, role in recovery after brain TBI. Catecholamine antagonists should therefore be minimized to avoid the risk of detrimental effects on functional outcomes. A detrimental effect from haloperidol and a beneficial effect of dextroamphetamine on motor recovery were reported in an animal model (440). Significantly poorer motor and functional outcomes in CVA patients exposed to catecholamine antagonists were then reported (441). Finally, a longer duration of posttraumatic amnesia in those treated with haloperidol for posttraumatic agitation was found in a retrospective review of rehabilitation outcomes in severe TBI (442).

In contrast, catecholamine agonists have a beneficial effect on motor control and executive function, with a
## TABLE 65.35  Respiratory Medications

<table>
<thead>
<tr>
<th>Medication Class (Mechanism)</th>
<th>Generic (Trade) Names</th>
<th>Dose</th>
<th>Unique Characteristics and Side Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Anticholinergics:</strong> relieve bronchoconstriction via antagonism of muscarinic cholinergic receptors</td>
<td>Ipratropium bromide (Atrovent) Tiotropium (Spiriva)</td>
<td>2 puffs (by metered-dose inhaler) qid, up to 12 puffs for every 24 h; maximal clinical effects in 30–90 min and last for 4 h; 18 mcg inhaled qd</td>
<td>Most cause significant systemic side effects because of easy absorption; fewer side effects with ipratropium and tiotropium because very little respiratory or GI absorption.</td>
</tr>
<tr>
<td><strong>Decongestants</strong> (α-1 adrenergic agonists): nasal mucosa blood vessel vasoconstriction, thus decreased secretion and congestion</td>
<td>Ephedrine, epinephrine, naphazoline, oxymetazoline, phenylephrine, phenylpropanolamine, pseudoephedrine (Sudafed), tetrahydrozoline, and xylometazoline</td>
<td>Depends upon the particular preparation</td>
<td>Decreases secretions and congestion, also used in combination with agents such as expectorants, antitussives, or antihistamines; side effects, H/A, dizziness, increased BP, and palpitations</td>
</tr>
<tr>
<td><strong>Expectorants</strong>: not clearly understood how they enhance respiratory tract secretion production, but make it easier to advance sputum upward</td>
<td>Guaifenesin (Robitussin, Humibid LA, Guiatuss, Fenesin, Guaifex LA, Organidin), terpin hydrate</td>
<td>100–400 mg PO q4h or 600–1,200 mg PO q12h (extended release)</td>
<td>Usually given orally, either alone or in combination with other respiratory agents; GI distress, especially in high doses; in extended usage, those containing iodide may cause iodism, hypothyroidism or hypersensitivity</td>
</tr>
<tr>
<td><strong>Mucolytics:</strong> decrease mucus viscosity by splitting mucoprotein disulfide bonds</td>
<td>Acetylcysteine (Mucomyst, Mucoisol)</td>
<td>6–10 cc 10% or 3–5 cc 20% solution; nebulized form q6–8h</td>
<td>Possible nausea and vomiting, stomatitis, or rhinorrhea</td>
</tr>
<tr>
<td><strong>Inhaled steroids:</strong> inhibit inflammatory cells and mediators</td>
<td>Beclomethasone (Beclovent, Vanceril); Budesonide (Pulmicort); Flunisolide (AeroBid); Fluticasone (Flovent); Triamcinolone (Azmacort)</td>
<td>2 puffs (84 mcg) tid–qid; max 20 puffs/24 h; 200 mcg/puff, 1–2 puffs bid, max 4 puffs bid; 250 mcg/puff, 2 puffs bid, max 4 puffs bid; 2–4 puffs bid, max of 880 mcg bid; 2 puffs (200 mcg) tid–qid or 4 puffs bid, max 16 puffs/24 h</td>
<td>Flunisolide is the most potent of these inhaled drugs; fluticasone inhaler available in strengths of 44, 110, 220 mcg</td>
</tr>
<tr>
<td><strong>Leukotriene inhibitors:</strong> block leukotriene receptors found in the lung airways (montelukast and zafirlukast) or inhibit the formation of certain leukotrienes (LTB4, LTC4, LTD4, LTE4) by specifically inhibiting 5-lipoxygenase (zileuton)</td>
<td>Montelukast (Singular); zafirlukast [Aczolatil, zileuton (Zyflo)]</td>
<td>10 mg qPM; 20 bid, 600 qid</td>
<td>Common side effects are headache, abdominal discomfort, and cough; also, zafirlukast and zileuton may elevate proxtimes in patients on warfarin; also, zileuton may augment the effects of propranolol and may cause elevated liver enzymes</td>
</tr>
<tr>
<td><strong>Selective β-2-adrenergic agonists:</strong> affinity for β-2 adrenergic receptors in bronchial smooth muscles</td>
<td>Albuterol (Proventil), Ventolin, Velmax), Arformoterol (Brovana), bitolterol (Tornalate), formoterol (Foradil), isoetharine (Bronkosal), nebulized; Brokometer, inhaled); levalbuterol (Xopenex), metaproterenol (Alupent, Metaprel, ProMeta); pirbuterol (Maxair), salmeterol (Serevent), terbutaline (Brethine, Bricanyl)</td>
<td>Response rate varies with administration mode: rapid response to aerosol inhalation or subcutaneous injection; delayed but longer if given orally</td>
<td>Tachycardia, arrhythmias, and myocardial ischemia, especially with underlying cardiac disease</td>
</tr>
</tbody>
</table>
Medications that Can Impair Cognition

| Aminophylline, anticholinergic agents (e.g., Benadryl), anticonvulsants, antiemetics, antipsychotics, barbiturates, \( \beta \)-blockers, \( \beta \)-blockers, BNZs, central-acting antihypertensive agents (e.g., Clonidine), gastric motility agents (e.g., metoclopramide), cardiac glycosides, \( \mathrm{H}_2 \) blockers, hypnotics, opiates |

resultant improvement in arousal, attention, processing speed, and depression (439). A beneficial effect of a single dose of amphetamine in the motor function of CVA patients was demonstrated (441). Catecholamine agonist side effects include hallucinations, GI upset, orthostasis, and dyskinesias.

Carbidopa/Levodopa
This presynaptic DA agonist is used in Parkinson’s disease and is reported to improve cognition in TBI patients with persistent deficits and to increase arousal after TBI and encephalitis (443–445). It is used at 25/100 mg tid up to 25/250 mg 8x/day.

Bromocriptine (Parlodel)
This DA agonist has been reported in case studies to enhance functional recovery in vegetative TBI patients (446). It is initiated at 2.5 mg q AM with slow titration up to 60 mg qd in divided doses with careful monitoring for nausea.

Amantadine (Symmetrel)
Amantadine facilitates DA action at presynaptic and postsynaptic receptors and is thought to have neuroprotective properties by acting as an NMDA receptor antagonist (447). It has been reported to improve arousal, attention, concentration, and decrease agitation but can lower the seizure threshold (447,448). Studies have recently suggested that amantadine may hasten recovery from diffuse axonal injury and improve recovery after severe TBI (449,450). Dosing is 100 to 400 mg divided qd to bid.

Methylphenidate (Ritalin)
Ritalin blocks presynaptic DA and NE reuptake. It is most commonly used for impaired initiation, speed of processing, concentration, and attention. More recent studies of Ritalin show positive effects on attention and speed of processing (451). It has been inferred from studies that methylphenidate’s action early after TBI is related to its effect on attention and concentration leading to increased therapy participation and thus faster functional recovery (452–454). To avoid insomnia, it should be given early in the day at 5 to 60 mg in divided doses. Although anorexia, arrhythmias, and HTN are potential side effects, studies have not supported that it increases seizure risk or has a significant effects on HR or BP (455,456).

Atomoxetine (Strattera)
Atomoxetine is a selective NE reuptake inhibitor currently indicated for attention-deficit hyperactivity disorder in children and adults. This agent increased noradrenergic activity in the locus ceruleus and prefrontal cortex. In the TBI population, this novel stimulant may be useful for attention, hypoarousal, fatigue, and as an adjunct for depression. This agent is well tolerated and thought to have lower abuse potential than traditional stimulants. Dosage ranges from 25 mg in the morning to 100 mg. Dosages greater than 40 mg are generally divided to bid early in the day. No published studies to date have been done in the TBI population, but trials on fatigue and cognitive impairment are warranted (457).

Agents that Affect Ach
Ach is primarily produced in the nucleus basalis of Meynert and is the neurotransmitter most involved with memory (458). Cholinergic fiber loss is thought to play a key role in attention and memory deficits in Alzheimer’s disease and TBI (459). The Ach system has therefore been targeted for pharmacologic modulation to enhance memory. Older cholinergic agents, such as physostigmine and tacrine, had limited use due to systemic side effects and short half-lives. More recently, a new class of acetylcholinesterase inhibitors has been FDA approved for dementia and has been studied in TBI. Smaller studies have behavioral and memory improvement in subjects taking donepezil (Aricept) (458) without significant side effects. More recently, a multicenter placebo-controlled trial of rivastigmine (Exelon) to treat cognitive deficits showed promising results for patients with moderate to severe cognitive deficits (459).

Agitation is a common occurrence in the TBI recovery period (435). Before considering an agent to treat this, any medication that may potentiate agitation by impairing cognition should be minimized (see Table 65-36), environmental modifications should be instituted to minimize unnecessary stimuli and assist with orientation and disordered sleep-wake cycles should be treated as needed. Pharmacologic treatment for agitation spans a broad range of medications including: anticonvulsants, antidepressants, antihypertensives, antipsychotics, BNZs, buspirone, stimulants, and amantadine (435). Treatment choice is based on type and frequency of target behaviors. Optimally, the medication should minimize agitation without impairing arousal and cognition. When possible, one should be chosen that also treats a coexisting condition. Careful monitoring of target behaviors and side effects is critical for successful treatment.

Miscellaneous Agents: Modafinil
Hypoarousal and fatigue are common complaints after TBI and can lead to poor cognitive and physical functioning. Stimulants can be effective in some patients. Modafinil (Provigil) is approved to treat narcolepsy. Modafinil was reported to reduce fatigue in multiple sclerosis and was well tolerated at 200 mg qd (460). For the TBI patient, dosing should start at 100 mg and then can be titrated up to 400 mg. No published studies to date have been done in the TBI population, but trials on fatigue are currently being done. Although Modafinil’s mechanism of action is still illusive it works in the hypothalamus to promote awakefullness.


Analgesia is the relief of pain and is achieved by removing the painful stimulus, blocking the generation of nerve impulses at sensory nerve endings, or reducing signal strength in pain pathways. Analgesia is achieved through the systemic administration of either narcotic or nonnarcotic analgesic agents. Anesthesia is the loss of all sensation with or without the loss of consciousness and can be achieved by blocking the generation of nerve impulses at sensory nerve endings or reducing signal strength in pain pathways. Anesthesia is achieved by either local application or systemic administration of anesthetic agents. Each of these groups of agents has as a therapeutic goal to reduce pain perception, which is the process whereby the brain integrates both sensory and emotional components of nociceptive or neuropathic input.

The process of nociception involves activation of specialized sensory receptors (nociceptors) in nerve endings by noxious insults such as inflammation or injury. A diverse array of chemical mediators, including prostaglandins, leukotrienes, bradykinin, adenosine triphosphate (ATP), histamine, acetylcholine, substance P, and serotonin, either directly or indirectly lower the threshold for activation of nociceptors on primary afferent neurons that terminate in the dorsal horn of the spinal cord. Synapses are formed with spinal neurons that project to supraspinal structures involved in processing and integrating nociceptive inputs resulting in the perception of pain.

A number of different neurotransmitters and neuropeptides, as well as ion channels, have been implicated in playing both direct and indirect roles in modulating signal strength in spinal and supraspinal nociceptive relay circuits (1,2). For example, glutamate released from primary sensory terminals activates both N-methyl-D-aspartate (NMDA) and non-NMDA subtypes of receptors; substance P and calcitonin gene–related peptide (CGRP) are colocalized and presumably co-released from terminals of primary afferents; adenosine and protein kinase C appear to play a role in nociceptive signaling; endogenous opioid peptides modulate both ascending (pain perception) and descending (pain control) signaling in nociceptive relay circuits; norepinephrine and serotonin mediate one component of descending control of nociceptive relay circuits; glutamate is also responsible for excitatory conduction in brain pathways mediating nociception and awareness; gamma-aminobutyric acid (GABA) decreases neuronal signal strength by increasing inhibitory conduction in nociceptive pathways; and ion channels (Ca\(^{2+}\), Na\(^{+}\), K\(^{+}\)) mediate axonal as well as synaptic conduction in nociceptive pathways resulting in both excitatory and inhibitory transmission. Given the complexity of nociceptive signaling circuits, antidepressants, anticonvulsants, neuroleptics, corticosteroids, NMDA receptor antagonists, central nervous system (CNS) stimulants, and \(\alpha_2\) receptor agonists are employed in the management of pain syndromes with varying results. Frequently, these agents are used in combination, especially with the nonnarcotic and narcotic analgesics.

Based on the actions of these various mediators, analgesic agents either block or augment signaling, resulting in a decrease in pain perception. Of the large group of agents noted above, this chapter focuses on nonnarcotic and narcotic analgesic agents, which have as their primary clinical use the management of moderate and severe pain, respectively. Typically, these agents are the first line of treatment for pain syndromes and as such provide a backbone for combination therapy with secondary agents. As such, this chapter is intended to provide an overview of these two groups, rather than a detailed account of individual agents. Dosages and summary of usage are found in Chapters 83 and 65.

**NONNARCOTIC ANALGESIC AGENTS**

Nonnarcotic analgesics (aspirin, acetaminophen, propionic acid derivatives, and cyclooxygenase II [COX II] inhibitors) provide relief of low- to moderate-intensity pain such as headache, myalgia, arthralgia, and other pains arising from integumental structures. Antipyretic (fever reduction) is also a prominent action of the nonnarcotic analgesics. In contrast to acetylsalicylic acid, the salicylates, propionic acid derivatives, and selective COX II inhibitors all exhibit significant anti-inflammatory effects. In general, this group of agents is available over the counter and thus widely available for self-medication and also abuse (3). Consumption is driven by the public who have become increasingly aware of managing their own health status as well as aggressive marketing campaigns by manufacturers of these products. As a result adverse reactions and interactions are occurring with increasing frequency, especially in the elderly population.

**Salicylates and Propionic Acid Derivatives (Ibuprofen, Naproxen, Ketoprofen)**

Acetylsalicylic acid (ASA), also known as aspirin, is absorbed rapidly following oral administration, partly from the stomach,
PART VI | MANAGEMENT METHODS

but mostly from the upper small intestine. Appreciable concentrations are found in plasma in less than 30 minutes, with peak blood levels occurring within 1 to 2 hours. The rate and extent of absorption are determined by a number of factors, including the disintegration and dissolution rates of tablets, the pH at the mucosal surface (acid conditions favor salicylate absorption due to increasing the nonionized or uncharged fraction), and gastric emptying time.

Aspirin is rapidly hydrolyzed by tissue and blood esterases to acetic acid and salicylic acid (salicylate), with the salicylate being distributed throughout most body tissues. Eighty to ninety percent of the salicylate is bound to plasma proteins, especially albumin. In the liver, salicylate is either conjugated with glycine to form salicyluric acid, conjugated with glucuronic acid to form ester or ether glucuronides, or metabolized (1%) to gentisic acid. The plasma half-life of aspirin is 15 minutes and that of salicylate 2 to 3 hours in low doses and 15 to 30 hours at high or toxic doses (4). This dose-dependent elimination is a result of the limited availability of hepatic glycine to form salicyluric acid and the availability of glucuronic acid to form glucuronides. As a result, at these high concentrations, there is increased unchanged (free) salicylate in the urine. Renal excretion of free salicylate is highly variable and is pH dependent; as little as 5% free salicylate is excreted in acid urine, and up to 85% free salicylate is excreted in alkaline urine (5).

After oral administration, the propionic acid derivatives (the “profens,” e.g., ibuprofen, naproxen, ketoprofen) are rapidly absorbed; peak concentrations vary from 1 to 2 hours (half-life, 2 hours) for ibuprofen, 2 to 4 hours (half-life, 14 hours) for naproxen, and 1 to 2 hours (half-life, 2 hours) for ketoprofen. The propionic acid derivatives are extensively bound (99%) to plasma proteins, usually albumin. The concentration of the propionic acid derivatives may be higher in synovial fluid than in plasma. These agents are extensively metabolized in liver, with more than 90% of an ingested dose of the propionic acid derivatives excreted in the urine as water-soluble metabolites or their conjugates.

The analgesic and anti-inflammatory actions of both the salicylates and propionic acid derivatives are a direct result of their ability to inhibit prostaglandin synthesis. Intermediates in prostaglandin biosynthesis or prostaglandins themselves sensitize pain receptors to mediators such as kinins, histamine, 5-hydroxytryptamine (5-HT), and so forth, which are released as a result of tissue damage. Salicylates and the propionic acid derivatives decrease prostaglandin synthesis through inhibition of prostaglandin synthase I and II (COX I and II) at both central and peripheral sites (6). COX I is constitutive and expressed in most tissues, including those involved with kidney function and gastric cytoprotection. COX II is inducible and activated by cellular mediators released owing to tissue damage and in part is responsible for the inflammatory response. Although both COX I and COX II isoenzymes are inhibited by these agents, there is considerable variation in potency; naproxen is about 20 times more potent than aspirin or ibuprofen. In addition to inhibiting prostaglandin synthesis, both aspirin and the propionic acid derivatives inhibit the migration of polymorphonuclear leukocytes and macrophages into inflammation sites and stabilize lysosomal membranes (thus decreasing the release of mediators from granulocytes, basophils, and mast cells), furthering the anti-inflammatory properties of these agents. Aspirin and the propionic acid derivatives effectively penetrate joint fluid to maintain an anti-inflammatory action that exceeds the plasma half-life (7).

Inhibition of prostaglandin synthesis also mediates the antipyretic actions of aspirin and the propionic acid derivatives (8). Bacterial pyrogens stimulate the production of interleukin-1 (IL-1) in macrophages, which in turn stimulates prostaglandin synthesis through activation of COX in or near the hypothalamus (9). The increase in prostaglandin E-2 (PGE2) elevates body temperature by interfering with hypothalamic temperature control mechanisms, which regulate heat production versus loss. By blocking prostaglandin synthesis, aspirin and the propionic acid derivatives facilitate heat loss and reduce elevated body temperature.

The antiplatelet effects of aspirin have received significant notoriety with regard to reducing heart attacks and stroke. Aspirin as ASA before deacetylation in the portal circulation irreversibly acetylates COX in platelets (presystemic inhibition), thus blocking the synthesis of thromboxane A2 and hence the role in platelet aggregation (10). Platelets are particularly sensitive to ASA’s effects because they do not synthesize new COX; the effects of a single dose can last 8 to 10 days (11). The propionic acid derivatives and salicylic acid do not irreversibly inhibit the COX, and thus, it has not been established if these agents produce similar benefits.

The widespread use of aspirin and the propionic acid derivatives provides significant documentation for the adverse side effects of these agents. The most prominent side effect is gastric irritation, and up to 40% of patients taking aspirin have been demonstrated to experience this problem (12). The propionic acid derivatives have significantly less gastric irritation, but this varies from one agent to another and one patient to another. Prostaglandins protect the gastric mucosa by reducing gastric acid production and release. Inhibition of COX I by these agents releases the prostaglandin inhibitory effect on gastric acid production leading to gastric irritation and potential gastric and duodenal ulcers. Coadministration of a prostaglandin analog (misoprostol) can reduce these side effects (13).

Prostaglandins also influence salt and water excretion by regulating renal blood flow and by direct effects on the renal tubules. Because both COX I and COX II are expressed in the kidney, enzyme inhibition can lead to side effects such as fluid retention and electrolyte imbalance (14). Patients with renal disease appear to be most sensitive to the salicylates and propionic acid derivatives such that a reversible decrease in GFR, interstitial nephritis, and nephrotic syndrome may occur. The potential for renal side effects is greater with agents that have a longer half-life, including agents that have greater selectivity for the COX II isoenzyme.
Children with chicken pox or flulike symptoms who take aspirin are up to 25 times more likely to acquire Reye’s syndrome (15). Reye’s syndrome is a deadly disease of children and teenagers occurring in the aftermath of a viral infection (especially the influenza virus) (16). Aspirin administration, in conjunction with viral illnesses, appears to damage hepatic mitochondrial membranes in genetically predisposed individuals, leading to liver injury and encephalopathy. A warning now appears on products containing aspirin in order to avoid the possible complications associated with Reye’s syndrome.

Hypersensitivity reactions to aspirin include asthma (8% to 20% of adult asthmatics) (17), skin rashes, angioedema, and potentially anaphylactoid reactions. Patients with asthma and nasal polyps may experience further sensitivity to aspirin, including bronchospasm and hypotension (18). An additional consideration with both acute and chronic use of aspirin is the prolongation of bleeding time. Although alterations in bleeding time have not been reported to occur with the propionic acid derivatives, hypersensitivity reactions have been reported (19).

Toxicity to aspirin has been well documented, largely because of the general availability of aspirin in household medicine cabinets. Mild aspirin toxicity or salicylism includes headache, dizziness, mental confusion, tinnitus, nausea, and vomiting. These symptoms are dose dependent and resolve with reducing the dosage (20). Severe aspirin toxicity is marked by hyperventilation due to increased CO₂ production from salicylate-induced uncoupling of oxidative metabolism in skeletal muscle and direct stimulation of the medullary respiratory center. Respiratory alkalosis, which is compensated for by increased sodium and potassium bicarbonate excretion, leads to severe disturbances in acid-base balance proceeding to convulsions, coma, and, after a period of unconsciousness, death from respiratory failure. Treatment is symptomatic in mild cases of poisoning, but administration of intravenous fluids and correction of acid-base and electrolyte balance is required in more severe cases. Alkalization of the urine will increase the excretion of free salicylate (21).

Because of their widespread use, drug interactions are not unusual with administration of aspirin and the propionic acid derivatives (22,23). In particular, any other agent that produces gastric irritation potentiates the effects of these drugs on the gastric mucosa. Oral anticoagulants are displaced from plasma protein-binding sites by these agents, which results in an increase in anticoagulant action. Among a number of drugs, combined administration with aspirin can lead to an increase in free versus bound tolbutamide, chlorpropamide, and phenytoin as a result of displacement from plasma protein–binding sites.

**COX II Inhibitors**

The selective COX II inhibitors have garnered much press lately based on the withdrawal of rofecoxib (Vioxx) from clinical use and the precautions for using celecoxib (Celebrex) (24). Along with valdecoxib (Bextra) and newer compounds such as etoricoxib, these compounds are second-generation nonsteroidal anti-inflammatory agents that demonstrate selectivity for the COX II enzyme (25). Based on selectivity for the COX II enzyme, the therapeutic advantage of these agents is a reduction in gastrointestinal side effects while producing anti-inflammatory effects comparable to the propionic acid derivatives (26,27). The use of the COX II inhibitors is indicated primarily in clinical settings where gastrointestinal side effects preclude the use of aspirin or the propionic acid derivatives (28). However, although the gastrointestinal side effect profile is reduced, contraindications such as allergy, aspirin-induced asthma, and compromised renal function still should be considered. Of greater concern is the risk of cardiovascular and cerebrovascular events with COX II inhibitors which have resulted in the worldwide withdrawal of rofecoxib and significant restrictive labeling for other agents in this class. Such warnings provide even greater concern for potential reactions and interactions when combined with over-the-counter agents such as the propionic acid derivatives administered for long periods of time (29).

**Acetaminophen**

Acetaminophen has analgesic and antipyretic actions, which are comparable to therapeutic doses of aspirin and the propionic acid derivatives. The mechanism whereby acetaminophen produces these actions is unclear because it demonstrates weak inhibition of COX I and COX II and also does not suppress inflammation. Studies suggest that the weak inhibition of COX and hence lack of anti-inflammatory action may be due to the presence of high concentrations of peroxides found at inflammatory sites which inactivate the acetaminophen molecule (30). The presence of a COX III isoenzyme that acetaminophen selectively inhibits or the production of reactive metabolites by the peroxidase function of COX II may further explain the lack of anti-inflammatory actions of acetaminophen (31,32).

Acetaminophen is rapidly and almost completely absorbed from the gastrointestinal tract. The concentration in plasma reaches a peak in 30 to 60 minutes, with a half-life in plasma of 2 to 3 hours. Unlike aspirin and the propionic acid derivatives, acetaminophen is relatively uniformly distributed throughout most body fluids; binding of the drug to plasma proteins is 20% to 50% (33). Practically, no acetaminophen is excreted unchanged, and the bulk is excreted after hepatic conjugation with glucuronate (about 60%) or sulfate (35%).

Acetaminophen is usually well tolerated though skin rash and other allergic reactions exhibited by localized areas of erythema or urticaria occur occasionally. The most serious toxic effect of acute overdosage of acetaminophen is a dose-dependent, potentially fatal hepatic necrosis (34). Studies have demonstrated that as little as 4 g/d of acetaminophen (eight extra-strength tablets) can produce severe liver injury.

When toxic doses are ingested, the ability to glucuronidate acetaminophen is exceeded, shifting metabolism to the microsomal enzyme system (microsomal enzyme oxidizing system)-dependent N-hydroxylation to form N-acetyl-p-benzo-quinoneimine (NABQ) (30). NABQ usually reacts with
sulphydryl groups in glutathione to form an inactive product. Because of the limited amount of hepatic glutathione, excess NABQ reacts with sulphydryl groups in hepatic proteins, producing cellular damage and hepatic necrosis. Early signs of liver toxicity include nausea, vomiting, and abdominal pain. Clinical signs of liver damage may become apparent within 2 to 4 days of ingestion of toxic doses (35). Regular consumption of ethanol can induce hepatic cytochrome P-450, leading to increased oxidative metabolism of acetaminophen to NABQ and lowering the threshold for acetaminophen-induced liver damage (36). Alcoholics may also exhibit hepatic glutathione depletion, which can further increase the severity of liver injury.

Treatment of overdosage with acetaminophen requires vigorous supportive therapy, including procedures to limit continued absorption of the drug such as induction of vomiting or gastric lavage, followed by oral administration of activated charcoal. Hemodialysis, if it can be initiated within the first 12 hours, has been advocated for all patients. Administration of sulphydryl compounds, such as N-acetyl cysteine, increases the hepatic content of glutathione, potentially reducing the availability of free NABQ (37).

Renal tubular necrosis can result from long-term abuse of analgesic formulations containing acetaminophen (38). This nephropathy results from p-aminophenol, a metabolite of acetaminophen, which concentrates in the hypertonic renal papillae. Renal papillary necrosis occurs as a result of p-aminophenol--induced depletion of reduced glutathione.

**Ketorolac**

Ketorolac differs from the previously mentioned agents in that, although it is a potent inhibitor of COX I and COX II, it is used primarily as a potent analgesic, usually postoperatively as an alternative to narcotic analgesics (39). Ketorolac is rapidly absorbed orally or intramuscularly with peak blood concentrations within 30 to 50 minutes and with a half-life of 4 to 6 hours. It is highly bound to plasma proteins (98% to 99%) and is metabolized in the liver by conjugation with glucuronic acid; about 60% is excreted unchanged (40). Ketorolac is a potent inhibitor of COX I and COX II and retains the properties of being analgesic and antipyretic. Although these actions would predict a role in inflammation, the recommended short-term therapy with ketorolac precludes its use for this purpose.

Short-term therapy and administration of the lowest effective dose help to avoid adverse effects, which can be severe (41). Gastrointestinal bleeding and perforation and reduced hemo-stasis due to platelet inhibition have been reported. Although renal impairment has been associated with administration of ketorolac, recent studies indicate this is not significant when adhering to the above dosage guidelines. Elderly patients may exhibit a lower threshold for adverse effects. Dosage is not an issue with administration of ketorolac to asthmatic patients, patients with nasal polyposis, or those with aspirin sensitiv-ity. Such hypersensitivity reactions can lead to life-threatening bronchospasm (42).

**NARCOTIC ANALGESIC AGENTS**

Agents with morphinelike actions (opiates) are known as narcotic analgesics and are used for the treatment of all types of severe pain. The broad spectrum of actions mediated by opiate receptors predicts the use of these agents in multiple clinical settings. For example, selected natural and synthetic opiates are used not only in the management of pain but also in the treatment of cough, diarrhea, and dyspnea due to acute left ventricular failure and also as anesthetic and adjunctive anesthetic agents.

Narcotic analgesic and antagonist agents are classified based on their source or actions at the opiate receptor. The naturally occurring opiates are derived from the exudate of the seed pod of the opium poppy, which contains more than 20 alkaloids, including morphine and codeine. A second classification based on source is the synthetic opiates derived from commercial synthesis and accounts for more than 20 different compounds with varying therapeutic properties. The classification based on action at opiate receptor includes full agonists, partial agonists, agonist-antagonists, or full antagonists. The action at the opiate receptor in large part determines their clinical use.

**The Opiate Receptor**

In 1973, radioligand-binding techniques demonstrated the presence of an "opiate" receptor (43–45). Binding of ligands to these receptors explained the diverse pharmacologic effects of opiate agonists and antagonists (46). High concentrations of these receptors were demonstrated in areas of the CNS known to involve pain signal transmission, including the dorsal horn of the spinal cord, periaqueductal gray, the rostral ventral medulla, and several thalamic nuclei (47). Additional properties included structural specificity—small modifications of the drug molecule cause large changes in drug binding (and in drug effect *in vivo*); stereospecificity—only the 1-isomeric form of the agent binds with high affinity (and is active as an analgesic); competition between agonists and antagonists—drugs of partially similar structure can bind to the receptor and block binding of agonists such as morphine; reversibility—bound drug can be displaced from the receptors by an excess of other molecules that possess binding affinity; and binding affinity versus efficacy—a good correlation exists between the affinity for binding to the receptors and potency of agonist or antagonist *in vivo* (48).

During the series of studies characterizing the opiate receptor, it became clear that a single receptor could not account for the diverse pharmacologic effects of opiates such as morphine (49). Based on the use of selective agonists, the effects of narcotic analgesics could be differentiated into responses unique to three different opiate receptor subtypes: the mu (μ), delta (δ), and kappa (κ) opiate receptors (50,51). All three receptors are members of the G-protein–coupled family of receptors and show significant amino acid sequence homologies (52,53). Receptor selectivity profiles have subsequently revealed the presence of subtypes of each of these receptors including μ₁, μ₂, δ₁, δ₂; and κ₁, κ₂, κ₃ (51).
Endogenous Opiate Peptides

The discovery that opiate receptors could be demonstrated in “naïve” tissues suggested the existence of endogenous morphinelike compounds that interacted with the receptor. Three families of opiate peptides, or opiopeptins, were subsequently discovered that are derived from distinct precursor polypeptides and located in neuronal pathways that mediate nociceptive signaling (54–57). Beta endorphin is a large peptide split off enzymatically from proopiomelanocortin; it is concentrated in the pituitary and hypothalamus and is located in areas around the third and fourth ventricles. During painful episodes, beta endorphin is released into the cerebrospinal fluid (CSF). Enkephalins (methionine enkephalin, leucine enkephalin) are pentapeptides, breakdown products of proenkephalin; based on localization in primary afferent nerve endings, they appear to regulate transmitter release, thereby inhibiting transmission of pain stimuli. Dynorphins are formed from prodynorphin and share many of the properties of the enkephalins, especially being localized within nerve endings of CNS pain pathways.

These peptides bind with high affinity to opiate receptor subtypes (Table 66-1), and binding to the receptor is reversed by narcotic antagonists. In addition, these agents demonstrate cross-tolerance to narcotic agonists, and dependence is produced with chronic administration. Moreover, a withdrawal reaction is induced by the administration of narcotic antagonists in animals made dependent by chronic administration of the endogenous opiate peptides (57,58).

Clinically used opiates such as morphine produce their analgesic effects through activation of the opiate receptor subtypes. The pharmacologic profile of these agents indicates multiple actions at the receptor, including full agonist, agonist-antagonist, and partial agonist actions. The receptor activation profile is listed below. It is noteworthy that, in the presence of morphine, the agonist-antagonists or partial agonists can function as antagonists to the actions of morphine. The actions of these agents at the receptor subtypes in large part indicate how they are used in the clinical setting (Table 66-2).

Pharmacokinetics

Depending on the agent, oral absorption of opiates is variable owing to first-pass metabolism in the liver. For example, the bioavailability of morphine following oral administration is only 25% (59). There is also great variability among patients with regard to both blood levels achieved and the rate of elimination. Absorption following intramuscular administration is usually complete, with maximum blood levels achieved within 15 to 30 minutes. Transdermal and epidural administration of opiates are also quite effective (60). Codeine and oxycodone demonstrate good oral-to-parenteral potency ratios due to less first-pass metabolism. In large part, this is due to structural protection from conjugation with glucuronic acid. Meperidine is well absorbed from all routes of administration; about 50% escapes first-pass metabolism following oral administration.

Although distribution to tissues is based largely on blood flow, the levels in the CNS are determined by the blood-brain barrier and the lipid solubility of the agent. For example, whereas a relatively small amount of morphine penetrates into the CNS, the more lipid-soluble fentanyl and its congeners penetrate much more readily (61). Protein binding of the various agents also influences the amount that enters the CNS. For example about 30% of morphine is protein bound and thus unavailable for penetration into the CNS. This compares with about 10% protein binding for codeine, 60% for meperidine, and 90% for methadone.

Most of the opiates are metabolized in the liver to form water-soluble conjugates that are eliminated by the kidney. For example, very little morphine is excreted unchanged, with most appearing as the inactive morphine-3-glucuronide and the active metabolite morphine-6-glucuronide. The morphine-6-glucuronide has actions that are very similar to morphine and
may account for the prolongation of analgesia with morphine administration (62). The glucuronides are eliminated through glomerular filtration and excretion in the urine.

A potentially fatal interaction may occur as a result of coadministration of meperidine and monoamine oxidase inhibitors (MAOIs) (63). Meperidine is N-demethylated in the liver to normeperidine and then to normeperidine acid with subsequent conjugation. MAOIs inhibit a step in the metabolism, leading to normeperidine accumulation, which causes excitation, hyperpyrexia, and seizures (64).

**Pharmacologic Actions**

**Cellular Effects**

All the actions of opiates are mediated by binding to opiate receptors at presynaptic or postsynaptic sites in the CNS, or tissues in the periphery. These agonists produce their unique effects by binding to receptors in tissues that mediate therapeutic (or side effect) outcomes. For example, agents such as morphone produce their analgesic effects by binding to receptors located in nociceptive neuronal pathways in the brain and spinal cord; constipation, a side effect of morphine administration, results from binding to opiate receptors in gastrointestinal smooth muscle. It is binding to these receptors, which results in common therapeutic and side effects among the opiate-narcotic analgesic group of agents. Effects do vary quantitatively, however, among the various agents.

At the cellular level, opiate receptors are coupled to guanosine triphosphate (GTP)-binding proteins, leading to inhibition of adenylate cyclase at both presynaptic and postsynaptic sites (65). Presynaptically, a reduction in voltage-gated Ca\(^{2+}\) influx reduces excitatory neurotransmitter release (66). Postsynaptic actions include increased K\(^+\) efflux through activation of receptor-operated K\(^+\) currents, leading to hyperpolarization and a reduction in neuronal firing rate (67). The fact that these actions alter signaling in a number of different neurotransmitter systems accounts for the multiple effects of the narcotic analgesics at various central and peripheral sites.

**CNS Effects**

Analgesia results from the inhibition of nociceptive reflexes at both spinal and supraspinal sites, primarily by interacting with \(\mu\) receptors. The threshold for pain perception is raised, and the reaction to the pain is reduced (even though pain is perceived). Continuous dull pain is relieved more effectively than the reaction to the pain is reduced (even though pain is perceived). Pain is relieved more effectively than the reaction to the pain is reduced (even though pain is perceived).

Narcotic analgesics produce dose-dependent respiratory depression by decreasing the sensitivity of the respiratory center chemoreceptors to \(\text{CO}_2\). Decreases in rate, minute volume, and tidal exchange are all possible with administration of opiates (68). Separate central effects include miosis resulting from increased activity in the parasympathetic nerve innervating the pupil. *Nausea and vomiting* can occur with opiate administration and are a result of direct stimulation of the chemoreceptor trigger zone for emesis, in the area postrema of the medulla. All clinically useful \(\mu\) agonists produce some degree of nausea and vomiting and may involve a vestibular component because the incidence is much lower in recumbent patients. *Cough suppression* is a useful side effect and results from direct depression of the cough center in the medulla. Both codeine and dextromethorphan, the nonaddicting p-isomer of the codeine analog levallorphan, are used clinically as antitussives. Finally, truncal muscle rigidity may occur with high doses of opiates, especially the highly lipid-soluble agents such as fentanyl. This side effect appears to result from enhanced tone in large trunk muscles, probably owing to increased impulse activity in dopaminergic pathways (69).

**Peripheral Effects**

At therapeutic doses, the opiates produce minimal effects on the cardiovascular system (blood pressure, cardiac rate and rhythm) in the supine patient. However, morphine does reduce baroreceptor mechanisms, producing arteriolar and venous dilation, which results in *orthostatic hypotension*. This may be complicated by morphine-induced histamine release and resultant vasodilation. *Histamine release* may also lead to cutaneous flushing, itching, and loss of body heat (hypothermia). Depending on the agent, bronchoconstriction and hypotension may occur as a direct result of the histamine release (70).

The presence of opiate receptors in smooth muscle accounts for significant effects of agents such as morphine on tone and motility in intestinal, biliary tract, ureter, and urinary bladder smooth muscle. Large and small intestine smooth muscle tone is increased and propulsive contractility decreased. Although these actions produce constipation as the most common side effect, it is a desirable outcome in patients with diarrhea or irritable bowel syndrome. Constriction of biliary smooth muscle and the sphincter of Oddi in the biliary tract leads to increased biliary tract pressure (*biliary colic*) (71). In addition, opiate-induced increases in tone in the ureter, detrusor muscle of the urinary bladder, and vesicle sphincter can produce *urinary retention*. Combined with a lack of appreciation that the bladder is full, urinary retention is a particularly problematic side effect of opiate administration.

**Tolerance and Physical Dependence**

The repeated and continued administration of morphine and other opiates results in the development of tolerance to both therapeutic and side effects of these agents. As a result, increasing dosages are required to produce quantitatively the same effect. Tolerance is gained to analgesic, euphoric, sedative, and respiratory depressant effects, but not to the miotic and constipating effects (72). Less tolerance is gained to the agonist-antagonist agents; tolerance develops more slowly and to a lesser degree to methadone, which is why it is used in opiate abuse treatment programs. Chronic administration of all of the narcotic analgesics results in the development of tolerance; the degree to which tolerance is gained varies according to the percentage of action as an agonist (full agonist > agonist-antagonist > partial agonist).
Cross-tolerance occurs across the spectrum of agents that are full agonists.

Chronic agonist treatment leading to down-regulation of opiate receptors results in attenuation of opiate responses (73,74). It is unclear how this down-regulation occurs; however, constant availability of agonist has been demonstrated to induce phosphorylation of the opioid receptor, thereby inducing desensitization (75). Receptor desensitization most likely involves an interaction between multiple neurotransmitter systems, as suggested by studies that demonstrated inhibition of morphine tolerance and dependence by agents that, for example, block NMDA receptors (76).

Without the continued presence of agonist at the receptor site, the cellular processes linked to the receptor become hyperexcitable, leading to characteristic withdrawal or abstinence syndrome. Withdrawal can also be precipitated in a dependent individual by administering a narcotic antagonist that accomplishes the same outcome.

**Interactions**

Adverse reactions may occur with the administration of opiates in individuals with particular disease or injury states and those taking certain drugs. Because of the effects of opiates to depress respiration, caution should be used in patients with pulmonary disease and those with head injuries. In the latter case, after intravenous or intramuscular administration, opiates may lead to an increase in CO2, thereby increasing cerebral blood flow and intracranial pressure to exacerbate the injury (77). Because of the dependence on the liver for metabolism and termination of action, opiate dosage may need to be adjusted in patients with liver disease. Finally, patients with pancreatitis may experience a further increase in biliary tract pressure owing to effects of the agents on biliary smooth muscle tone.

Combined administration of narcotic analgesics along with CNS depressants (barbiturates, antianxiety agents, antipsychotics) results in potentiation of the sedative effects of opiates. As mentioned earlier, meperidine administered to patients taking MAOIs can produce excitation, convulsions, hyperpyrexia, respiratory depression, and hypotension.

**Toxicity**

Acute overdosing with opiates produces what has been described as the "triad" of symptoms, including depressed respiration, pinpoint pupils (except with meperidine), and coma (78). The first step in treatment is to establish an airway for ventilatory support and attempt to stabilize the cardiovascular system. Administration of a narcotic antagonist will reverse all of the effects of the opiate but may also precipitate withdrawal in dependent individuals.

**Narcotic Antagonists**

Opiate (narcotic) antagonists block or reverse all of the actions of the opiate agonists by displacing the agonists from the opiate receptors. A full antagonist binds to the μ, κ, and δ opiate receptors (affinity) but has no action or effect on the receptor (efficacy) (69). The three agents classified as full antagonists are naloxone (half-life, 1 to 2 hours), nalmefene (half-life, 8 hours), and naltrexone (half-life, 10 hours). Naloxone due to extensive first-pass metabolism is administered intravenously for reversal of agonist-induced adverse effects. Naltrexone is administered orally and effectively blocks the euphorogenic effects of opiates for up to 24 hours, thus indicating its use in opiate addiction programs (79).

In the presence of an agonist, antagonism at opiate receptors is also observed for mixed agonist-antagonists such as nalbuphine. Nalbuphine is an antagonist at μ receptors and an agonist at κ receptors (see Table 66-2). Pentazocine has weak antagonist actions at μ receptors and moderate agonist actions at κ receptors. Butorphanol is similar to pentazocine but more potent as an agonist at κ receptors. Partial agonists, such as buprenorphine, can also function as antagonists in the presence of morphine because they have greater affinity for μ receptors and do not activate the κ or δ receptors (80). Apparently, the partial agonist can sterically hinder morphine from binding to the full complement of receptor.

The actions of antagonists, agonist-antagonists, and partial agonists have the potential to reverse the beneficial effects of agonists such as morphine when attempting to antagonize adverse side effects such as respiratory depression (81). As a result, analgesia may be “unmasked” as well as sedation. An increased incidence of postoperative nausea and vomiting has been reported with the use of these agents. Precaution should be taken with regard to the half-life of the antagonist versus agonist, especially because the half-life of naloxone is only 1 to 2 hours versus the longer half-life of agonists. All of the agents with antagonist activity may precipitate withdrawal in dependent patients and in the fetus of narcotic-dependent pregnant patients (82).

**REFERENCES**

Injection of medications in proximity to nerves, muscles, and skeletal structures (bursae, joints, and tendons) provides an important intervention in the management of pain and dysfunction. These medications can provide analgesia, reduce inflammation of the affected structure, and promote tissue healing directly and indirectly through facilitation of rehabilitation. Commonly, local anesthetics are used for immediate anesthesia and analgesia, corticosteroids are used to control inflammation, radiopaque contrast is used to assist any fluoroscopic confirmation, and vasoconstrictors are used to prolong local anesthetic effects or warn of intravascular injection. Injection techniques constitute both diagnostic and therapeutic modalities in many settings. As in all patient care, the origin of pain or dysfunction involves careful history taking and a complete physical examination. The details of these techniques are addressed generally in Chapter 2 with a specific focus in Chapters 32, 33, 35, 49, and 68. This chapter first outlines the historical development and general principles of proper procedural techniques, including indications and contraindications. Significant side effects or complications and commonly used medications are reviewed. General techniques for nerve blocks and specific nerve blocks are then discussed.

HISTORICAL

Since antiquity, sharp objects have been used to inject various concoctions into the body as remedies for pain and dysfunction (1). The tools of modern-day injection procedures include the hollow needle developed by Rynd in 1845 (2) and the syringe developed by Pravaz in 1853 (3) and Wood in 1855 (4). Medications to inject for anesthesia were developed in the late 1800s, with extensive use of cocaine as a topical analgesic (5) and as an injected analgesic (6). The application of these devices and medications to achieve local anesthesia for relief of pain was described by Corning in 1894 (7). Historical information on neural blockade has been reviewed in detail by Brown and Fink (1).

Since the early 1900s, procaine had been injected into the synovium of inflamed joints for temporary relief of pain. In 1949, Hench (8) introduced systemic corticosteroids to suppress the inflammatory changes in rheumatoid arthritis. However, large doses of corticosteroids given systemically resulted in complications. In 1951, Hollander (9) introduced and reported on low-dose local (intraarticular and periarticular) injections of hydrocortisone acetate to control pain and inflammation caused by trauma and inflammatory joint disease. Historical information on intraarticular injection has been reviewed in detail by Hollander and colleagues (10).

Tender points in the muscle were identified in the 1800s by multiple investigators (11,12). Kraus in 1937 (13) and Travell in 1942 (14) introduced the treatment of myofascial pain by direct trigger point injection and use of vapocoolant spray and emphasized the importance of exercise in the treatment of patients with trigger point mediated pain. Historical information on myofascial pain has been reviewed in detail by Simons (11).

GENERAL

The general principles for use of neural blockade were described in detail by Bonica in 1953 (15). Injections of local anesthetic agents have been shown to be effective in managing patients with acute and chronic pain. The mechanism of action is the blockade of nociceptive input along the pathway of transmission (16). Intraarticular and periarticular injections of corticosteroids have been shown to reduce inflammation and pain as well as to facilitate mobility and function (17). Unfortunately, injection of medication is often used in isolation for the management of many pain problems. It is important to use injections as one component in a vast armamentarium of techniques that are often best used in concert with each other. An injection unaccompanied by physical therapy, exercise, stretching, behavior modification, and optimization of coexisting psychological disorders is often incomplete in alleviating pain or restoring function. This is best demonstrated by the “block clinics,” which were highly popular in the 1930s and 1940s but have since been superseded by multidisciplinary pain clinics. Injection procedures should be performed within the context of comprehensive rehabilitation.

Knowledge and Training

Using injection procedures requires sufficient knowledge to understand the diagnosis and treatment of pain syndromes. Limitations, complications, advantages, and disadvantages of each procedure and alternative treatments must be understood when choosing the best therapy or combination of therapies. The practitioner must be highly skilled in injection techniques based on education, training, and experience. This requires a
thorough knowledge of the anatomic basis of the procedure and the characteristics of the injectable medication, including expected side effects of the procedure as well as potential complications and their prevention and prompt treatment (16).

**Contraindications**

Injection procedures encompass a wide variety of techniques, including epidural, caudal, nerve, motor point, joint, and muscle. Although relative advantages and disadvantages exist, certain conditions comprise contraindications to perform any injection technique. Certain medical conditions also may preclude an injection. For example, a patient with severe lung disease should not have a rib block with its attendant risk for pneumothorax.

**Absolute Contraindications**

Absolute contraindications for injection procedures include patient refusal, localized infection, dermatologic conditions that preclude adequate skin preparation, existence of a tumor at the injection site, history of allergy to local anesthetic agents, the presence of severe hypovolemia, gross coagulation defects, increased intracranial pressure (for epidural procedures), and septicemia.

**Relative Contraindications**

Among relative contraindications are the lack of education, training, and skill on the physician’s part. In addition, relative contraindications are the patient’s minor coagulation abnormalities or concurrent use of certain anticoagulants such as minidose heparin. Diabetes is a relative contraindication for injections with corticosteroids because of the possibility of hyperglycemia, glycosuria, electrolyte imbalance, and increased risk for infection.

**Screening for Contraindications**

Early in the interview process, the patient should be asked specific questions about conditions that would preclude the injection procedure. Firstly, anticoagulation issues must be addressed, and the patient should have been able to discontinue antplatelet therapy (e.g., clopidogrel) or anticoagulation (e.g., unfractionated heparin or warfarin) for sufficient duration. The practitioners should be aware of current guidelines regarding anticoagulation and neuraxial procedures and have developed a standard practice ahead of time in regards to anti-coagulation and injections. Secondly, the patient should have transportation and a responsible adult who can care for the patient after the procedure, especially if the injection has the potential to cause weakness or sedation. Thirdly, the patient should be appropriately fasted to reduce the chance of aspiration, particularly if he or she is likely to need sedation. Any allergies to the expected injection medications or to material such as latex should be documented and alternatives be used for the procedure.

**Communication**

Communication with the patient is essential at all steps involved in patient care, including injection. Providing the patient with a complete explanation of the procedure will result in the individual having increased confidence and reduced anxiety. During the procedure, the physician should continually inform and reassure the patient as to the progress of the procedure.

**Informed Consent**

Although state laws vary in terms of the documentation required, informed consent involves providing enough information about the injection procedure so that the patient can intelligently decide whether or not to proceed (without external coercion or manipulation). This involves a thorough discussion of the procedure with the patient, including the possible benefits, alternative treatments, common side effects, and risks of complications. The patient should be given the opportunity to ask questions about details of the injection. The patient should not receive medications that could significantly impair responses or judgment before the consent is given. Depending on state regulations, the patient may or may not be required to sign an informed consent form. Significant complications of injection procedures are provided in Table 67-1.

**Universal Precautions**

Universal precautions are required for all injection techniques to reduce the incidence of transmission of infectious agents (18). These include the use of gloves and protective eye wear. Possible transmission of blood-borne pathogens during medical procedures is a common concern among patients and health care workers, owing to its potentially devastating consequences. Although this concern has largely focused on the human immunodeficiency virus (HIV), other agents, particularly hepatitis B, C, and G viruses, are of significantly greater risk. The data show that the risk to health care workers from patients is far greater than the risk to patients from health care workers (19). It is important that proper sharps disposal containers are maintained in all areas where needles are used. Hollow needle sticks pose the greatest risk for occupational transmission, with studies following health care workers after

<table>
<thead>
<tr>
<th>Significant Injection Complications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Systemic toxic reaction</td>
</tr>
<tr>
<td>Other systemic reactions</td>
</tr>
<tr>
<td>Epinephrine reaction</td>
</tr>
<tr>
<td>Vasovagal reaction</td>
</tr>
<tr>
<td>Allergic reaction</td>
</tr>
<tr>
<td>Accidental spinal block</td>
</tr>
<tr>
<td>Concurrent medical episode</td>
</tr>
<tr>
<td>Infection</td>
</tr>
<tr>
<td>Pneumothorax</td>
</tr>
<tr>
<td>Nerve injury</td>
</tr>
<tr>
<td>Other complications</td>
</tr>
<tr>
<td>Hypotension</td>
</tr>
<tr>
<td>Hematoma</td>
</tr>
</tbody>
</table>
such exposures finding the following seroconversion rates: hepatitis B, 5% to 37%; hepatitis C, 3% to 10%; and HIV, 0.2% to 0.8% (19–23). Research shows that recapping a needle increases the risk for needle stick, hence needles should either be laid down in the sterile field or disposed uncapped in an appropriate container.

**Positioning**

Optimal results of any injection depend on proper positioning of the patient before, during, and after the injection procedure. The proper position should facilitate access to the injection site, with the patient and provider being as comfortable as possible. Bony prominences should be cushioned where needed to avoid discomfort from pressure. When possible, the recumbent position is used because it is usually the most comfortable for the patient. It also minimizes the incidence of orthostatic hypotension caused by vasovagal reaction during the procedure. Positioning to avoid ergonomic stress on the patient and physician is essential. Consideration should be made for access to special equipment such as a fluoroscopic C-arm or ultrasound machine.

**Skin Preparation**

Strict aseptic preparation is required in all injection procedures, as well as intact skin at the site of injection. Preparation of the injection site is undertaken in a standard aseptic fashion over an area large enough to allow palpation of landmarks and any potential extension of the injection site, and sterile technique is used throughout the procedure to minimize the risks. Most injection techniques described in this chapter do not require full surgical draping; however, the phrase “sterile technique is used throughout the procedure” indicates an increased risk for major infectious complications, including epidural abscess, meningitis, and adhesive arachnoiditis.

The injection site is prepared with an antiseptic to reduce cutaneous microorganisms to the lowest level. Commonly used agents include chlorhexidine (Hibiclens, Zeneca Pharmaceuticals, Wilmington, DE), iodophors (Betadine, Purdue Frederick, Norwalk, CT), and alcohol (24). It is important for the clinician to wait about 2 minutes after the application of any of the above antiseptics to obtain maximal reduction in cutaneous bacteria. It is rarely justified to remove hair from the skin for an injection procedure (25). In addition to skin preparation, some procedures such as intradiscal procedures or prolonged catheter trial probably warrant systemic intravenous antibiotics prior to needle insertion (26). Many providers will wear a surgeon’s cap, filtration mask, and full surgical gown, although any differences in the rate of infections have been difficult to demonstrate.

**General Procedural Sequence**

Informed consent should be obtained from the patient, an intravenous catheter placed if sedation is to be used, and monitors applied (pulse oximeter, noninvasive blood pressure, and 3 to 5 lead EKG) if indicated based on use of sedation or on patient comorbidities. Resuscitation equipment, including oxygen source, mask and bag for ventilation, suction, code medications, and advanced airway equipment should be readily available. The patient should be properly positioned, and the injection site prepared and draped in standard aseptic fashion over an area large enough to allow palpation of landmarks. All medications, syringes, needles, and other equipment should be readily available. Syringes may vary from 3 to 12 mL. Needles may vary from as small as ½-in. (1-cm), 25-gauge needles to ¾-in. (9-cm), 21-gauge spinal needles. For convenience, an 18-gauge needle should be available to draw medications from the vial. All caregivers participating in the injection procedure should be gloved and practicing universal precautions. An audible “time-out” involving confirmation of the correct patient, injection site, and injection side (if applicable) should be performed. Then, as the procedure is underway, there should be periodic communication with the patient to screen for any problems, such as severe pain from the needle, signs of local anesthetic toxicity, paresthesias, etc. After the injection is complete, the patient should be monitored afterward for any side effects, and reexamined to evaluate for proper distribution and efficacy of analgesia. Discharge instructions should be given to the patient including symptoms to expect, symptoms which warrant calling a nurse or physician, activity restriction (if any), follow-up scheduling, and personal log to record duration and intensity of pain relief.

**Needle Insertion**

To reduce pain of the initial needle insertion, the skin may be stretched while rapidly piercing the skin. Once the skin is pierced, the tension is released and the needle is advanced slowly. Rapid infusion of medication may result in tissue distention, causing pain. Other methods of reducing pain with initial needle insertion include the use of topical anesthetics and vapocoolant sprays.

**Conscious Sedation**

The procedures described in this chapter rarely require conscious sedation. In circumstances in which conscious sedation is used, monitoring should be applied to the patient as previously described, and advanced airway management tools should be readily available. Resuscitation equipment, medications, and personnel must be readily available. Any prior sedation records should be reviewed for previous sedation dose requirements and any anesthetic complications.

**MEDICATIONS**

Injection procedures use three primary classes of medication: local anesthetic agents, neurolytic agents, and corticosteroids. All of these medications have multiple clinical applications and are documented to be safe and effective when used appropriately. It is incumbent on the practitioner to understand the efficacy, complications, and common side effects of these agents.
Anesthetic Agents

The mechanism of local anesthetic action is a reversible block of ion flux through the axon’s sodium channels. By blocking transmission from the peripheral nerves, no input is detected at the central nervous system (CNS). The degree of neural blockade depends on the properties, absorption, amount, location, and other characteristics of the drug injected (28,29) (Table 67-2). For purposes of attaining longer duration analgesia, agents such as bupivacaine or ropivacaine are acceptable and safe when used within the dosing guidelines. Anesthetic agents with fast onset, such as lidocaine, are useful particularly for skin infiltration prior to block insertion or for rapid diagnosis of the pain generator site. Local anesthetics should be chosen appropriately for both the purpose and the site of injection. For example, commercial preparations may come with preservatives, making them potentially neurotoxic if given intrathecally. Therefore, it is imperative to verify that the particular medication preparation is approved for the desired use. This information is available in many locations, including the package insert for the medication.

When higher doses of local anesthetics are used (i.e., near the upper limit of recommended dose), are injected near highly vascular structures (face, intercostals), or injected near the neuraxis (paravertebral, epidural, or intrathecal), monitoring of the patient is highly recommended. Verbal communication, constant screening for symptoms, and pulse oximetry can identify any CNS toxicity, while the electrocardiogram and blood pressure can identify any cardiovascular toxicity. If signs of toxicity or high spinal block are present, injection should be suspended and the patient closely evaluated and/or supported.

Neurolytic Agents

Neurolytic nerve blocks were popular in the early 1900s. However, improved use of analgesics, as well as the use of radiofrequency and cryoablative techniques, has replaced their use in most instances. Alcohol and phenol are the most widely used neurolytic agents in the United States. These agents indiscriminately affect motor and sensory nerves.

Phenol can be used intrathecally and epidurally, as well as for peripheral nerve and motor point blocks. It is poorly soluble in water and is often added to glycerin to achieve concentrations higher than 7%. It also can be added to radiographic contrast to allow fluoroscopic visualization of the spread during injection. Phenol has a local anesthetic effect, resulting in less pain after the injection. Because of this, long-term effects

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Procaine (Novocain)</th>
<th>Lidocaine (Xylocaine)</th>
<th>Prilocaine (Citanest)</th>
<th>Mepivacaine (Carbocaine)</th>
<th>Bupivacaine (Marcaine)</th>
<th>Tetracaine (Pontocaine)</th>
<th>Etidocaine (Duranest)</th>
<th>Ropivacaine (Naropin)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physicochemical</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relative potency</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>Relative toxicity</td>
<td>1</td>
<td>1.5</td>
<td>1.5</td>
<td>2.0</td>
<td>10</td>
<td>12</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>pH of solution</td>
<td>5–6.5</td>
<td>6.5</td>
<td>4.5</td>
<td>4.5</td>
<td>4.5–6</td>
<td>4.5–6.5</td>
<td>4.5</td>
<td>7.4</td>
</tr>
<tr>
<td>Clinical</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Onset</td>
<td>Moderate</td>
<td>Fast</td>
<td>Fast</td>
<td>Fast</td>
<td>Moderate</td>
<td>Very slow</td>
<td>Fast</td>
<td>Slow</td>
</tr>
<tr>
<td>Dispersion</td>
<td>Moderate</td>
<td>Marked</td>
<td>Marked</td>
<td>Marked</td>
<td>Moderate</td>
<td>Poor</td>
<td>Moderate</td>
<td>Moderate</td>
</tr>
<tr>
<td>Duration of action</td>
<td>Short</td>
<td>1–2</td>
<td>1.5–2</td>
<td>1.75–2</td>
<td>2–2.5</td>
<td>6–8</td>
<td>6–8</td>
<td>5–8</td>
</tr>
<tr>
<td>Relative duration</td>
<td>1</td>
<td>1–2</td>
<td>1–2</td>
<td>1–2</td>
<td>0.25–0.5</td>
<td>0.1–0.25</td>
<td>0.5–1</td>
<td>0.25–0.5</td>
</tr>
<tr>
<td>Concentration of solution (%)</td>
<td>1–2</td>
<td>1–2</td>
<td>1–2</td>
<td>1–2</td>
<td>0.25–0.5</td>
<td>0.1–0.25</td>
<td>0.5–1</td>
<td>0.25–0.5</td>
</tr>
<tr>
<td>Maximum recommended dose (mg/kg, adult)</td>
<td>10–14</td>
<td>6–10</td>
<td>6</td>
<td>6–10</td>
<td>2–3</td>
<td>2</td>
<td>4–5</td>
<td>3–4</td>
</tr>
<tr>
<td>Total dose (mg, adult)</td>
<td>500</td>
<td>300</td>
<td>—</td>
<td>400</td>
<td>150</td>
<td>—</td>
<td>300</td>
<td>200</td>
</tr>
<tr>
<td>Toxic blood levels (μg/mL)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CNS</td>
<td>—</td>
<td>18–21</td>
<td>20</td>
<td>22</td>
<td>4.5–5.5</td>
<td>—</td>
<td>4.3</td>
<td>4.3</td>
</tr>
<tr>
<td>CVS</td>
<td>—</td>
<td>35–50</td>
<td>—</td>
<td>—</td>
<td>6–10</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

aProcaine = 1.
CNS, central nervous system; CVS, cerebrovascular system.
of the block cannot be evaluated for 24 to 48 hours after the effects of the local anesthetic dissipate. Doses greater than 100 mg can result in serious toxicity (30).

Alcohol is used primarily intrathecally for nerve roots and locally for sympathetic blockade. Patient positioning with intrathecal use is extremely important, because alcohol is hypobaric in cerebrospinal fluid (CSF). It is readily soluble in body tissues and produces intense burning on injection. It requires 12 to 24 hours before block effects can be determined.

### Corticosteroids

Glucocorticosteroids are used in the injection treatment of inflammatory processes (31) (Table 67-3). Some of the commonly used forms of corticosteroid available for intraarticular injections are listed as follows:

- Betamethasone sodium phosphate and acetate (Celestone Soluspan), 6 mg/mL
- Methylprednisolone acetate (Depo-Medrol), 40 and 80 mg/mL
- Prednisolone sodium phosphate (Hydeltrasol), 20 mg/mL
- Prednisolone terbutate (Hydeltra-TBA), 20 mg/mL
- Triamcinolone acetonide (Kenalog), 40 mg/mL
- Triamcinolone hexacetonide (Aristospan) 20 mg/mL

These corticosteroids vary in strength, concentration, duration, and side effects, and all are effective. Triamcinolone hexacetonide has the longest duration of suppression of inflammatory activity. Fluorinated corticosteroids (e.g., triamcinolone) are rarely chosen for soft-tissue injection, because they are more likely to cause tissue atrophy. Prednisolone tebutate and methylprednisolone acetate are often used for soft-tissue injections because of their efficacy and cost.

### SIDE EFFECTS AND COMPLICATIONS

The injection procedures described here are associated with various side effects, although the risk for significant complications is very low. The physician who performs injection procedures should have the training and education to recognize and treat a wide range of potential complications (Table 67-4). Advanced cardiac life support (ACLS) protocols are used in airway management and cardiorespiratory resuscitation (27).

### Systemic Toxic Reaction

Various toxic reactions have been reported after use of local anesthetics, but with very low incidence. Local anesthetic agents are relatively lipid-soluble, low-molecular-weight compounds that readily cross the blood-brain barrier. As toxic levels are reached, disturbances of CNS function are observed initially, producing signs of CNS excitation. Early symptoms of overdose include headache, ringing in the ears, numbness in the tongue and mouth, twitching of facial muscles, and restlessness. As blood levels increase, generalized tonic-clonic seizures may occur. If sufficiently high blood levels are reached, the initial excitation is followed by generalized CNS depression. Respiratory depression and ultimately respiratory arrest may occur secondary to the toxic effect of the local anesthetic on the respiratory center in the medulla. Occasionally, the excitatory phase may not occur, and toxicity presents as CNS depression.

Cardiovascular system (CVS) effects either result indirectly from inhibition of autonomic pathways during regional anesthesia (as in high spinal or epidural block) or are directly due to depressant actions on the CVS. The CVS is generally

### TABLE 67.3 Corticosteroids

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Hydrocortisone (Cortisol)</th>
<th>Prednisolone (Hydeltra)</th>
<th>Methylprednisolone (Depo-Medrol)</th>
<th>Triamcinolone (Aristospan, Kenalog)</th>
<th>Betamethasone (Celestone)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physiochemical</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relative</td>
<td>1</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>25</td>
</tr>
<tr>
<td>anti-inflammatory potency&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relative mineralcorticoid potency&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1</td>
<td>0.8</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>pH of solution</td>
<td>5.0–7.0</td>
<td>6.0–8.0</td>
<td>4.5–6.5</td>
<td>6.8–7.2</td>
<td></td>
</tr>
<tr>
<td>Clinical</td>
<td>Fast</td>
<td>Fast</td>
<td>Slow</td>
<td>Moderate</td>
<td>Fast</td>
</tr>
<tr>
<td>Onset</td>
<td>Moderate</td>
<td>Poor</td>
<td>Intermediate</td>
<td>Moderate</td>
<td></td>
</tr>
<tr>
<td>Dispersion</td>
<td>Short</td>
<td>Intermediate</td>
<td>1+</td>
<td>Intermediate</td>
<td></td>
</tr>
<tr>
<td>Duration of action</td>
<td>2+</td>
<td>1+</td>
<td>180</td>
<td>300</td>
<td>100–300</td>
</tr>
<tr>
<td>Salt retention</td>
<td></td>
<td></td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Plasma half-life (min)</td>
<td>90</td>
<td>200</td>
<td>300</td>
<td>100–300</td>
<td></td>
</tr>
<tr>
<td>Concentration (mg/mL)</td>
<td>50</td>
<td>20</td>
<td>40–60</td>
<td>10–40</td>
<td>6</td>
</tr>
<tr>
<td>Range of usual dose (mg)</td>
<td>25–100</td>
<td>10–40</td>
<td>10–40</td>
<td>5–20</td>
<td>1.5–8</td>
</tr>
</tbody>
</table>

<sup>a</sup>Relative to hydrocortisone.
more resistant to toxicity than the CNS. The CVS-to-CNS toxicity ratio is lower for bupivacaine and etidocaine than for lidocaine. Convulsive activity may initially be associated with an increase in heart rate, blood pressure, and cardiac output. As the blood concentration of a local anesthetic agent increases further, CVS depression occurs, resulting in a decrease in blood pressure secondary to myocardial depression, impaired cardiac conduction, and eventual peripheral vasodilation. Ultimately, circulatory collapse and cardiac arrest may result. In addition, certain agents such as bupivacaine may cause ventricular arrhythmias and fatal ventricular fibrillation. The onset of CVS depression with bupivacaine may occur relatively early and be resistant to usual therapeutic modalities. The pregnant patient is more sensitive to the cardiotoxic effects of bupivacaine.

Systemic toxicity may be due to unintentional intravascular injection or drug overdose. Intravascular injection produces signs of toxicity (usually seizures) during the injection itself, especially if injected directly into blood vessels supplying to the brain (e.g., vertebral arteries). A relative overdose results in toxic reactions when peak blood levels are reached, about 20 to 30 minutes after the injection. Factors that affect the blood concentration (site of injection, drug, dosage, addition of vasoconstrictor, speed of injection) influence the potential for systemic toxic reactions to develop.

To minimize systemic reactions to local anesthetic agents, intravascular injection (the most common cause of seizures) should be avoided. Careful, intermittent aspiration should be used while injecting large quantities of local anesthetic agents. Patient complaints of metallic taste, numbness around the mouth, and ringing in the ears are suggestive of intravascular needle placement. Patients may be premedicated with midazolam or diazepam to raise the seizure threshold if necessary.

Systemic toxicity is treated with general supportive measures. If early signs of toxicity occur, constant verbal contact should be maintained, oxygen administered, breathing encouraged, and CVS function monitored. If seizure activity occurs, a clear airway should be maintained and oxygen administered by assisted or controlled ventilation. If seizures continue, thiopental (50 to 100 mg) or diazepam (5 to 10 mg) should be administered intravenously, avoiding large doses of thiopental, which may produce additional CVS or CNS depression. If airway maintenance is jeopardized, succinylcholine should be used to facilitate endotracheal intubation. Muscular convulsive activity is terminated with succinylcholine, but the seizure activity in the brain is not affected. If CVS depression occurs, hypotension should be treated by increasing intravenous fluids, positioning the patient properly (elevate the legs), and using vasopressors such as ephedrine or epinephrine (Table 67-5).

**Epinephrine Reaction**

Reaction to epinephrine may sometimes be confused with local anesthetic overdose. Systemic absorption of epinephrine

---

**TABLE 67.4 Differential Diagnosis of Local Anesthetic Reactions**

<table>
<thead>
<tr>
<th>Etiology</th>
<th>Major Clinical Feature</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Systemic toxic reaction</td>
<td>Immediate convulsion and/or cardiac toxicity</td>
<td>Injection into vertebral or a carotid artery may cause convulsion even with administration of small dose</td>
</tr>
<tr>
<td>Intravascular injection</td>
<td>Onset in 5–15 min with irritability, progressing to convulsions</td>
<td>May vary with vasopressor used Rapidly reversible by elevating the legs and discontinuing the noxious stimulus</td>
</tr>
<tr>
<td>Relative overdose</td>
<td>Tachycardia, hypertension, headache, apprehension Rapid onset</td>
<td>Allergy to amides extremely rare Cross-allergy, e.g., with preservatives in local anesthetics and food</td>
</tr>
<tr>
<td>Epinephrine reaction</td>
<td>Bradycardia</td>
<td>May lose consciousness with total spinal block; onset of cardiovrespirovascular effects more rapid than with high epidural or with subdural block</td>
</tr>
<tr>
<td>Vasovagal reaction</td>
<td>Hypotension</td>
<td>Medical history important</td>
</tr>
<tr>
<td>Allergy</td>
<td>Anaphylaxis (↓ blood pressure, bronchospasm, edema)</td>
<td></td>
</tr>
<tr>
<td>Immediate</td>
<td>Urticaria</td>
<td></td>
</tr>
<tr>
<td>Delayed</td>
<td>Gradual onset</td>
<td></td>
</tr>
<tr>
<td>High spinal or epidural</td>
<td>Bradycardia*</td>
<td></td>
</tr>
<tr>
<td>Concurrent medical episode</td>
<td>Hypotension</td>
<td></td>
</tr>
<tr>
<td>(e.g., asthma attack, myocardial infarct)</td>
<td>Possible respiratory arrest</td>
<td></td>
</tr>
<tr>
<td></td>
<td>May mimic local anesthetic reaction</td>
<td></td>
</tr>
</tbody>
</table>

* Sympathetic block above T-4 adds cardioaccelerator nerve blockade to the vasodilatation seen with blockade below T-4; total spinal block may have rapid onset.

Vasovagal Reaction

Vasovagal reaction is a frequent response to injection procedures that is attributable to physiologic and psychological factors. This response may result in bradycardia, hypotension, and loss of consciousness. This usually occurs in the initial portion of the procedure when skin is penetrated before any medication is injected. Vasovagal reaction is often mislabeled as an allergic reaction to the medication. This response is often preceded by dizziness, faintness, sweating, and pallor. Vasovagal reaction is rapidly reversible by removing painful stimuli and placing the patient in a head down and legs up position to improve venous blood return to the heart. If this fails to relieve the symptoms, then treatment with general supportive measures is indicated, including airway maintenance, oxygen, intravenous fluids, anticholinergics such as atropine, and vasopressors such as ephedrine (see Table 67-5).

Allergic Reaction

An allergic reaction to local anesthetic agents rarely occurs and in some instances may be confused with a vasovagal reaction or a reaction to epinephrine. Ester anesthetic agents (e.g., procaine, tetracaine) are more frequently associated with allergic reactions than amide anesthetic agents (e.g., lidocaine, ropivacaine), because esters are derivatives of paraaminobenzoic acid (PABA). However, the use of amides from multiple-dose vials may result in an allergic reaction secondary to the preservative methylparaben. Allergic reaction is treated with general supportive measures and the administration of fluids, antihistamine, steroids, and epinephrine (if warranted), as well as removal of any additional offending agent. The patient should be closely monitored for a clear airway.

Although this is a rare event, if there is a question of patient hypersensitivity to anesthetic agents, intradermal skin tests (injection with diluted [1:1,000] followed by undiluted local anesthetic) can be used successfully to diagnose adverse responses. However, false-positive results may occur. Anaphylactic shock is treated as systemic toxic reaction with attention to maintaining cardiovascular and ventilatory function (see Table 67-5).

Accidental Spinal Block

Inadvertent subarachnoid or epidural blockade can occur with any injection that is performed close to the spine. These injections include intercostal nerve blocks, sympathetic blocks, and nerve root injections. Proper equipment and staff should always be available to treat any possible complication. This includes the ability to administer fluids and vasopressors if the patient develops hypotension from sympathetic blockade and to maintain ventilation with oxygen if the patient has impaired respiratory function.

Concurrent Medical Episode

Concurrent medical problems may be exacerbated by injection procedures. Hypotension will reduce myocardial perfusion and may be a major factor for reinfarction in patients with ischemic heart disease. Patients with poorly treated hypertension risk developing hypertensive crisis, myocardial infarction, or stroke triggered by the stress of the procedure. Patients with chronic renal failure are more susceptible to toxicity from local anesthetics. Diabetic patients have an increased sensitivity to the effects of corticosteroid injection. Patients with liver disease may have reduced metabolism of local anesthetic medications resulting in an increased possibility of toxicity at standard doses. Other medical illnesses may decompensate clinically, owing to mild toxicity and changes in fluid and electrolyte balance. These risks are minimized
with appropriate monitoring and optimal medication regimes. Patients should be medically stable before undergoing elective injection procedures.

**Infection**

Although infectious complications rarely occur, cutaneous and joint infection, epidural abscess, bacterial meningitis, and adhesive arachnoiditis have been associated with injection procedures. Epidural abscesses may cause spinal cord compression. Signs and symptoms of epidural abscess are severe back pain, localized tenderness, fever, leukocytosis, cervical rigidity, increased protein and leukocytes in the cerebrospinal fluid, progressive neurologic symptoms, and abnormal imaging studies of the spine. Early diagnosis and prompt treatment are essential to avoid catastrophic complications. Meningitis and adhesive arachnoiditis are the result of the introduction of bacterial or irritating contaminants into the cerebrospinal fluid as well as trauma during the procedure.

**Pneumothorax**

Injections into the thoracic region have the potential to cause a pneumothorax. In procedures that put the patient at risk for needle penetration of the lung, less than 1% develop a pneumothorax (32, 33). Most of these patients can be easily treated with administration of 100% oxygen, close monitoring (e.g., $O_2$ saturation, vital signs) of the patient, and, when necessary, needle aspiration of air. Only those pneumothoraces that result in significant dyspnea or those under tension require chest tube thoracostomy and vacuum drainage. Bilateral thoracic procedures should be undertaken with caution due to potential risk of bilateral pneumothorax.

**Nerve Injury**

Three major factors contribute to nerve injury during injections: trauma, toxicity, and ischemia, with all three contributing to most nerve injuries. Nerve blocks are the result of infiltration of anesthetic agents around the nerve, not directly into the neural tissue. Intraneural injections directly injure nerve fibers and cause a breakdown in the blood-nerve barrier. The use of short, beveled needles has significantly reduced nerve injuries from injections. Intense pain in the nerve distribution and high resistance upon injection are often the result of intraneural needle placement and necessitate immediate cessation of the injection and repositioning of the needle.

**Other Complications**

Local anesthetic agents used in recommended clinical concentrations have minimal irritating effects on the nerves, skin, and subcutaneous tissue. Complete recovery of function occurs after regional blocks. The administration of large doses of prilocaine may lead to methemoglobinemia, owing to the accumulation of a metabolite (OH-toluidine) that can convert hemoglobin to methemoglobin. It may be treated with intravenous methylene blue. 

Hypotension may result from sympathetic blockade. This commonly occurs in patients who are hypovolemic and receive a spinal or epidural block involving a large portion of the sympathetic chain ganglia. Hypotension is treated with general supportive measures, including administration of intravenous fluids, vasopressors such as epinephrine to maintain blood pressure within normal limits, and proper positioning (elevating the legs). Bradycardia from blockade of sympathetic outflow from T1-4 may require prompt treatment.

Bleeding at the site of injection occurs commonly at the surface. Hematoma at the site of injection is a possibility but usually not clinically significant. If arterial puncture occurs, prolonged direct pressure is usually adequate to prevent the development of a hematoma.

**COMMON NERVE BLOCK TECHNIQUES**

**General**

The placement of local anesthetics at various sites along the neural axis is an important tool in the diagnosis and treatment of a variety of pain disorders, such as complex regional pain syndrome and postherpetic neuralgia. Peripheral nerve blocks also can provide muscle relaxation and pain relief to facilitate an active physical therapy program.

When the point of injection has been determined, it is best marked with the tip of a retracted ballpoint pen or a needle hub by pressing the skin to produce a temporary indentation to mark the point of entry. The skin is then prepared in a standard sterile fashion, and sterile technique is used throughout the procedure. The skin and subcutaneous tissue at the injection site may be anesthetized by injecting 1% lidocaine with no epinephrine using a 25- to 30-gauge needle. Alternatively, a vapocoolant spray or analgesic cream applied to the skin surface may be used to provide adequate anesthesia.

Before injecting the medication, an attempt to aspirate should always be made to avoid accidental intravascular injection. After ensuring that the needle is in the joint space, the medication should be injected in a slow, steady fashion.

**Indications for Nerve Blocks**

Neural blockade may be used for the diagnosis, prognosis, and treatment of pain. Selective nerve blocks are indicated to determine the etiology of pain by isolation of specific anatomic structures. Selected nerve blockade is used to determine specific nociceptive pathways and other mechanisms involved in pain generation. Diagnostic blocks assist in narrowing the differential diagnosis of the site and cause of pain. Prognostic neural blockade is used to evaluate the possible outcome from neurolytic procedures. Therapeutic nerve blocks are indicated to decrease morbidity in acute postoperative pain, posttraumatic pain, and pain resulting from self-limiting conditions. Nerve blockade may provide rapid relief of pain and facilitate the patient’s participation in a comprehensive rehabilitation program. Therapeutic nerve
blocks may interrupt the pain cycle sufficiently to provide prolonged pain relief.

**Contraindications for Nerve Blocks**

Absolute contraindications for regional anesthesia include patient refusal, localized infection, a skin condition that prevents adequate skin preparation, the existence of a tumor at the injection site, a history of allergy to local anesthetics, the presence of severe hypovolemia (for blocks that could result in significant sympathetic blockade), gross coagulation defects, septicemia, and increased intracranial pressure (spinal, caudal, and epidural).

Prilocaine should not be used in doses greater than 600 mg, because significant methemoglobinemia may result. The use of corticosteroids with preservatives is contraindicated in epidural and subarachnoid techniques because the preservative may result in seizures and permanent CNS damage.

Relative contraindications include general medical conditions that would put the patient at increased risk. These include aortic stenosis, severe lung disease, sickle cell anemia, and preexisting neurologic diseases such as multiple sclerosis or amyotrophic lateral sclerosis, which could be worsened during regional anesthesia.

**Complications**

Complications common to nerve blocks include hypotension from sympathetic blockade. This usually occurs in patients who are hypovolemic and receive a block covering a large portion of the sympathetic chain ganglia, for example, during spinal or epidural blockade. Local anesthetic overdose or intravascular injection can result in CNS toxicity and, in some cases, pulmonary and cardiac arrest. Nerve injury from contact with the needle may occur but is rare, especially when a short, beveled needle is used. Other complications are dependent on the location of the block and are discussed separately.

**Techniques**

Before the injection, the appropriate landmarks are located and marked. The skin is scrubbed with antiseptic and allowed to dry for 2 minutes. The wearing of sterile gloves is required so that the bony landmarks in the sterile field may be palpated throughout the procedure. The standard sterile technique is required to minimize the risk for infection. It is preferable to use single-dose vials of the local anesthetic because this further reduces the risk for infection. A 25- to 27-gauge needle is used with 1% lidocaine with no epinephrine to raise a small skin wheal for skin anesthesia. Routinely a 1½-in. (4-cm), 21- to 25-gauge needle transverses the skin, joint capsule, synovial lining and then slides smoothly into the joint cavity.

Aspiration is done to ensure there is no intravascular penetration. If penetration occurs, the needle should be repositioned and aspirated to ensure that blood vessels have been avoided; then the medication is slowly injected. After the medication has been injected, the needle should be cleared with a new syringe containing a small amount of lidocaine or saline. The needle is then withdrawn with pressure applied to minimize bleeding.

**Nerve Stimulator**

Peripheral nerves may be localized with a nerve stimulator using a small adjustable amount of electrical current to depolarize neural tissue in proximity to the needle. The cathode (negative) terminal is connected to the needle, and the anode (positive) terminal is connected to a grounding patch. The stimulator initially is set to deliver 10 to 20 mA of current to detect the general area of the nerve. The current is then reduced to further localize the nerve. The needle is positioned to produce the maximal twitch at the lowest stimulus. The needle is usually adjacent to the nerve when 0.5 to 0.1 mA produces motor stimulation with an insulated needle and 1 mA with an uninsulated needle (Fig. 67-1).

Nerve stimulators do not substitute for knowledge of anatomy and proper needle placement. Insulated needles increase the point of maximal current density at the needle tip and are used for precise localization of specific nerves. Uninsulated needles are often accurate enough for many nerve blocks; however, both the tip and shaft of the needle have sufficient current density to stimulate a nerve. Local muscle twitches from the shaft of the uninsulated needle should not be confused with the response from the nerve to be blocked (Fig. 67-2).

**SPECIFIC NERVE BLOCKS**

**Occipital Nerve Block**

**Indication**

The occipital nerve blockade is used both diagnostically and therapeutically in the treatment of occipital neuralgia.

**Techniques**

After informed consent is obtained, the patient is placed in the sitting position with the head flexed forward. The occipital nerve is located at the midpoint between the mastoid process and the greater occipital protuberance at the superior nuchal line. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. A 1-in. (2.5-cm), 21- to 25-gauge needle is inserted perpendicular to the superior nuchal line. Before reaching the periosteum of the skull, paresthesias in the occipital nerve distribution may be elicited. If not, the periosteum is contacted and the needle withdrawn slightly. After negative aspiration, 3 to 5 mL of local anesthetic is injected to block the occipital nerve (Fig. 67-3).

**Comments**

Localization of the nerve may be accomplished by palpation of the occipital artery just lateral to the nuchal ridge. The nerve runs with the occipital artery, innervating the posterior portion of the skull. A nerve stimulator may be used for precise needle placement. An alternate approach may be used with the
patient positioned as above with anatomic landmarks identified. A 2-in. (5-cm), 21- to 25-gauge needle is inserted subcutaneously along the middle third of the superior nuchal line. After negative aspiration, 5 mL of local anesthetic is injected to block the greater and lesser occipital nerves (see Fig. 67-3).

Complications
Intravascular injection can occur, resulting in systemic toxicity and seizures, especially if larger volumes are used. Bleeding due to vascular injury also may occur. Nerve injury secondary to injection into the nerve may result in persistent numbness over the posterior portion of the scalp.

Stellate Ganglion Block
Indication
Stellate ganglion blockade is useful for diagnosing and treating pain of sympathetic origin. This includes pain involving the face, head, neck, and upper extremities secondary to complex regional pain syndrome, acute herpes zoster, and phantom limb pain.

Techniques
After informed consent is obtained, the patient is placed in the supine position with a pillow under the shoulders and the neck extended. The patient is prepared in a standard aseptic fashion
over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. The transverse process of the C6 vertebral body is palpated between the cricoid cartilage and the carotid artery. A 1½-in. (4-cm), 22-gauge needle is inserted vertically and advanced to touch the periosteum of the C6 transverse process. The needle is withdrawn slightly, and after negative aspiration, a 1-mL test dose of local anesthetic is injected. After unremarkable test dose and repeated aspiration, 9 mL of local anesthetic is injected, in divided doses with continuous monitoring, to block the stellate ganglion (Fig. 67-4).

Comments
This technique is commonly used for differential diagnosis and is the preferred treatment of sympathetic mediated pain involving the upper extremity. The stellate ganglion is located between the anterolateral surface of the seventh cervical vertebral body and neck of the first rib. It is formed by the inferior cervical ganglion and first thoracic sympathetic ganglion. Autonomic mediated pain does not usually correspond to segmental or peripheral nerve distribution.

This procedure requires full monitoring capability to include blood pressure, EKG, heart rate, level of consciousness, and pulse oximeter. Temperature should be monitored and recorded for each hand before, during, and after the procedure. The patient is continuously monitored for change and level of consciousness or for adverse reaction. Successful stellate ganglion blockade is suggested by rising temperature on the block side as well as evidence of a Horner’s syndrome (miosis, ptosis, anhydrosis, and enophthalmos). Nasal congestion and hoarseness may occur with this injection. It is recommended that intravenous access be available before the block in the event of intravascular injection resulting in seizure activity or cardiovascular instability (34). Rarely are ablative nerve procedures (neurolysis) recommended in the management of sympathetic mediated pain of the upper extremity.

Complications
Performance of this procedure outside of a fully monitored environment is not recommended. Resuscitation equipment and personnel must be readily available. Although appearing technically simple, this block has multiple potential hazards, owing to the proximity of the common carotid artery, vagus and recurrent laryngeal nerves, jugular vein, vertebral artery, trachea, esophagus, lung, and dura. Intraarterial and intradural injection of local anesthetic may result in death, seizure, respiratory arrest, cardiac arrest, cerebral damage with multiple sequelae, and other lesser complications. The risk for intervascular injection may be reduced if a test dose is given, the total dose is injected incrementally, and aspiration is performed before each injection.
Cervical Epidural Steroid Injection

Indications
Cervical epidural steroid injection (ESI) is primarily used to treat pain arising from cervical herniated discs or spinal stenosis.

Techniques
After informed consent is obtained, the patient is placed in the prone, lateral decubitus, or sitting position with the neck flexed. The sitting position, with the head resting on the examination table, provides better stabilization of the neck during the procedure but may be a problem if the patient is prone to light-headedness. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. Local skin anesthesia is provided with 1% to 2% lidocaine at the C7-T1 interspace. A Tuohy epidural needle is advanced in a midline, horizontal fashion until well seated in the posterior ligaments. A winged needle is preferred because it allows two-handed control of the needle as it is directed toward the epidural space. A midline or paramedian approach may be used, although large epidural veins lie laterally. The stylet is then removed, and a “loss-of-resistance” syringe is attached to the hub of the needle. Two to three mL of air or normal saline should be in the syringe. The needle is slowly advanced 1 to 2 mm at a time, with constant checking for loss of resistance by tapping on the plunger of the syringe. Alternatively, the needle can be advanced slowly and continuously with constant light pressure on the plunger. Once a distinct loss of resistance is obtained, the needle is halted, and an attempt is made to aspirate blood or CSF. After negative aspiration, 6 mg of betamethasone sodium phosphate and acetate or equivalent is injected. Care should be taken to flush the needle with normal saline and to replace the stylet before the needle is removed to avoid depositing steroid in the needle track to the skin.

Comments
Cervical epidural injections are similar in technique to lumbar epidural injections in some respects. The cervical spinous processes at C7 and T1 are oriented almost horizontally, as in the lumbar region. Of note, caution must be taken because the cervical ligamentum flavum is thinner in this region than at any other spinal level, and the width of the epidural space is only 3 to 4 mm. A catheter may be placed through a nonhearing needle to achieve the desired level under fluoroscopic guidance.

Complications
The most frequent complication of a cervical epidural injection is subarachnoid penetration (wet tap) due to the thinner ligamentum flavum and the reduced width of the cervical epidural space. Cervical epidural injections should only be attempted by those physicians with a great deal of experience with lumbar epidural techniques, because the spinal cord lies in close proximity to the epidural space. These injections are normally completed with fluoroscopic guidance, as noted in Chapter 68. Injection of local anesthetics into the cervical epidural space can result in respiratory depression, particularly if the phrenic nerve roots are blocked (C3 to C5).
Infection or bleeding into the closed epidural space also can result in significant neurologic deficits, including quadriplegia. Any complaint of increasing pain or neurologic changes should be investigated immediately. Early recognition can prevent permanent injury.

**Suprascapular Nerve Block**

**Indication**
The suprascapular nerve blockade is useful in patients as a therapeutic procedure for pain in the shoulder region. This block is used as an adjuvant to physical therapy in patients with limited range of motion secondary to arthritic shoulder pain, shoulder-hand syndrome, and shoulder pain.

**Techniques**
After informed consent is obtained, the patient is placed in the sitting position. The spine of the scapula is divided by a line formed by the bisection of the scapular angle. The upper outer quadrant formed from the spine of the scapula and the vertical line bisecting the angle of the scapula is marked, and a point is marked 2 cm anteriorly along this line for needle insertion. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. A 3-in. (8-cm), 23- to 25-gauge needle is inserted perpendicular to the skin and, using a nerve stimulator, advanced until needle placement is confirmed by movements of the supraspinatus and infraspinatus muscles. After negative aspiration, 5 mL of local anesthetic is injected (Fig. 67-5).

**Comments**
The suprascapular nerve is a branch from the trunk of the brachial plexus, which enters the scapular region through the suprascapular notch on the cephalic border of the scapula. Confirmation of the block is determined when abduction of the arm is diminished over the first 15 degrees. If a nerve stimulator is not available, the same technique is used, with the needle advanced to the dorsal surface of the scapula. The needle is then walked along the edge of the scapula to the suprascapular notch.

**Complications**
Intraneural injection may result in nerve damage. Severe pain on injection suggests the possibility of an intraneural injection, and the needle should be repositioned immediately. Hematoma and intravascular injection are possible, owing to the close proximity of the suprascapular vessels. Pneumothorax is possible if the needle is advanced beyond the scapula and into the pleura. Most pneumothoraces can be treated easily with administration of supplemental oxygen and close observation and, when necessary, needle aspiration of air. Only those pneumothoraces that result in significant dyspnea or those under tension require chest tube thoracotomy and vacuum drainage.

**Intercostal Nerve Block**

**Indications**
Intercostal blockade is used to treat pain from herpes zoster, rib fractures, and intercostal neuropathies. It is also used to diagnose unusual abdominal or chest wall pain.

**Techniques**
After informed consent is obtained, the patient is placed in either the prone or lateral position. In the lateral position, the injection site is along the midaxillary line, which may result in incomplete blockade of the lateral cutaneous branch of the intercostal nerve. In the prone position, the injection site is along the angle of the rib posteriorly. The ribs to be injected are marked at the angle of the rib or along the midaxillary line. If the ribs are not easily palpated, an alternative injection
technique such as an epidural or root block may be considered. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and strict sterile technique is used throughout the procedure. The index finger of the nondominant hand is used to palpate the rib and identify the intercostal space. The tip of the finger is placed in the intercostal space and the skin slid over the superior rib. A ½-in. (1.5-cm), 25-gauge needle is inserted directly over the rib until contact is made with the rib. The long axis of the needle and syringe should have a slight cephalad tilt and be perpendicular to the long axis of the rib. The needle is then moved to the inferior edge of the rib by walking (i.e., repeatedly slightly withdrawing) the needle in the subcutaneous tissue and allowing the skin to move back slowly to its original position. The needle should retain its slight cephalad tilt. As the needle slips off the inferior ridge of the rib, the tip is advanced about 3 mm and then aspirated. If aspiration is positive for blood or air, the needle should be repositioned; otherwise, 2 to 5 mL of local anesthetic is injected to block the intercostal nerve (Fig. 67-6).

Comments
A new needle should be used for each nerve blocked. Intercostal blocks are a simple and effective method of providing analgesia for painful disorders of the chest and abdominal walls. Because of the wide distribution of intercostal nerve innervation, the intercostal nerves above and below the level of pain must be blocked to gain optimal pain relief.

Complications
Intercostal blockade is often underused because of an exaggerated fear of pneumothorax. In actuality, less than 1% of all patients having an intercostal block develop a pneumothorax. Most can be easily treated with administration of supplemental oxygen and close observation and, when necessary, needle aspiration of air. Only those pneumothoraces that result in significant dyspnea or those under tension require chest tube thoracotomy and vacuum drainage. Local anesthetic toxicity can occur because of the rapid absorption after intercostal injection. Toxicity can easily be avoided by limiting the total amount injected to a known, safe level (see Table 67-4).

Spinal Nerve Root Block
Indications
Nerve root blockade is useful in diagnosing and treating pain present in a dermatomal distribution.

Techniques
After informed consent is obtained, the patient is placed in the prone position, and a line perpendicular to the axis of the spine is drawn across the top of the spinous process. In the lumbar region, a line drawn along the inferior edges of the two transverse processes will intersect the spinous process of the same vertebra at its most cephalad point. In the thoracic region, this line can extend as much as two vertebral levels caudally, particularly in the midthoracic region. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and strict sterile technique is used throughout the procedure. A 4-in. (10-cm), 22-gauge needle is inserted 3 to 5 cm laterally from the midline, which should overlie the transverse process. The needle is advanced perpendicular to the skin until contact is made with the transverse process. The depth is noted at this point. The needle is pulled back to skin level and redirected to pass below the transverse process. Appropriate response with a nerve stimulator or paresthesia indicates correct placement. After negative aspiration, 2 to 5 mL of local anesthetic is injected to block the spinal nerve root (Fig. 67-7).

Comments
The technique of nerve root blockade is similar to that of the intercostal block. The transverse process serves as the depth marker for nerve roots. A sound knowledge of the relationship of the transverse process and the spinous process is necessary to locate precisely the selected nerve root. Any disorder that responds to intercostal blockade also should respond to selective nerve root blockade.

Complications
Complications usually occur from injection of local anesthetic agents into areas adjacent to the paravertebral space such as the epidural or subarachnoid space. Puncture of retroperitoneal organs and bleed also can occur if care is not taken to consider the anatomy fully.

Celiac Plexus Block
Indications
Celiac plexus blockade is useful for diagnosing and treating pain of sympathetic origin. This includes pain involving the viscera, abdomen, and pelvis, secondary to cancer, complex regional pain syndrome, and vasospastic disorders.

Techniques
After informed consent is obtained, the patient is placed in the prone position with a pillow under the abdomen. The inferior edge of the spinous process of the first lumbar vertebra and the lower border of each 12th rib at 7 cm from the spinous process of the first lumbar vertebra is identified and marked. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. A 6-in. (15-cm), 22-gauge needle is inserted at the mark on the 12th rib toward the spinous process of L1 at 60 degrees from perpendicular. The needle is advanced until it contacts the lateral side of the L1 vertebra. The depth is noted at this point. The needle is pulled back to a subcutaneous level and reinserted at 45 degrees toward the spinous process of L1 and slightly cephalad until it slips off the edge of the vertebra. This is about 2 to 3 cm deeper than the original depth. After negative aspiration, 20 mL of local anesthetic is injected (Fig. 67-8).
FIGURE 67-6. Intercostal nerve block. Approach for intercostal nerve injection and neural blockade with injection sites marked at the angle of the rib. A: The tip of the finger is placed in the intercostal space and the slid (arrow) slid over the superior rib. B: The needle is inserted directly over the rib until contact is made. C: The needle is walked (arrow) to the inferior edge of the rib while maintaining a slight cephalad tilt. D: As the needle slips off the inferior edge of the rib, advance the needle approximately 3 mm (arrow), where it is adjacent to the intercostal nerve. (From Thompson GE, Moore DC. Celiac plexus, intercostal, and minor peripheral blockade. In: Cousins MJ, Bridenbaugh PO, eds. Neural Blockade in Clinical Anesthesia and Management of Pain. 3rd ed. Philadelphia, PA: JB Lippincott; 1998:459, with permission.)
Comments
This technique is commonly used for differential diagnosis and is the preferred treatment of sympathetic mediated pain involving the viscera and pelvis. The celiac plexus is located in the prevertebral region at the level of the L1 vertebral body. It is formed by the right and left celiacs, superior mesenteric, and aorticorenal ganglia. Autonomic mediated pain does not usually correspond to segmental or peripheral nerve distribution. Lateral and anteroposterior fluoroscopic views are recommended to ensure that the needle is properly positioned. It is recommended that intravenous access be available before the block in the event of hypotension from sympathectomy or toxicity from intravascular injection. It is necessary to perform this procedure bilaterally for a complete celiac plexus block. Rarely are ablative nerve procedures (neurolysis) required in the management of sympathetic mediated pain.

Complications
These injections are normally completed with fluoroscopic guidance, as noted in Chapter 68. Performance of this procedure out of a fully monitored environment is not recommended. Resuscitation equipment and personnel must be readily available. Although appearing technically simple, this block has multiple hazards, owing to the proximity of the aorta, kidney, pancreas, diaphragm, thoracic duct, and other vascular structures. Intraarterial or intradural injection of local anesthetic may result in death, seizure, respiratory arrest, cardiac arrest, cerebral damage with multiple sequelae, and other lesser complications. The risk for intravascular injection may be reduced if a test dose is given, the total dose is injected incrementally, and aspiration is performed before each injection.

Lumbar Epidural Steroid Injection
Indications
ESI is most effective for lumbosacral radiculopathy associated with intervertebral disc herniation, bulging, or degeneration. The main criterion for success is the presence of nerve root inflammation that can be relieved by the steroid. ESI also has been used to treat pain from degenerative joint disease, scoliosis, spondylolysis, spondylolisthesis, postlaminectomy syndrome, facet abnormalities, herpes zoster, and postherpetic neuralgia.
Techniques
After informed consent is obtained, the patient is placed prone, lateral decubitus, or sitting with the back, hips, and knees flexed. The injection should be performed as close as possible to the bony level of the nerve root irritation. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. Local skin anesthesia is provided with 1% to 2% lidocaine. A Tuohy epidural needle is then advanced between the spinous processes in a midline (or paramedian) approach until the needle is well seated in the posterior ligaments. The stylet is then removed and a loss-of-resistance syringe is attached to the hub of the needle. Two to 3 mL of normal saline should be in the syringe. The needle is slowly advanced 1 to 2 mm at a time, with constant checking for loss of resistance by tapping on the plunger of the syringe, or advanced slowly and continuously with light pressure on the plunger. Once a distinct loss of resistance is obtained, the needle is halted, and an attempt is made to aspirate blood or CSF. After negative aspiration, either 40 to 80 mg of methylprednisolone acetate or 6 mg mixture of betamethasone sodium phosphate and betamethasone acetate is injected. The steroids can be injected as is or diluted in 5 to 10 mL of preservative-free normal saline. If multiple levels are involved, dilution will ensure better spread of the steroid. The needle should be flushed with normal saline and the stylet replaced before it is removed from the skin to avoid tracking steroid to the skin. If blood is obtained on aspiration, the needle should be repositioned. If CSF is obtained on aspiration, the procedure can be continued; however, a spinal headache may result from dural puncture, and intrathecal injection is a possibility at this location (Fig. 67-9). The needle can be reintroduced at an adjacent level.

Comments
A complete evaluation of the pain should be undertaken before ESI to rule out serious neurologic dysfunction or malignancy. Local infection, sepsis, coagulation abnormalities, and patient refusal are contraindications to ESI. Success is higher if the patient has had no previous back surgeries and the pain has been present for less than 6 to 12 months. Methylprednisolone acetate is best used for localized nerve root irritation. Triamcinolone diacetate is water soluble and results in optimal outcome in generalized nerve root irritation such as arachnoiditis.

The paraspinous (paramedian) approach may be used if difficulty is encountered with the midline approach. A Tuohy epidural needle is inserted alongside the caudal edge of the inferior spinous process to the bony level of nerve root irritation. The needle is inserted with a 45-degree angulation to the long axis of the spine below and a 10-degree offset from midline and then advanced between the spinous processes until it is well seated in the posterior ligaments.

Complications
Inadvertent dural puncture may result in a postdural puncture headache. Risk of dural puncture is increased at sites of prior laminectomy. Intrathecal injection of steroid may result in aseptic meningitis, adhesive arachnoiditis, or conus medullaris syndrome. Each milliliter of methylprednisolone acetate contains about 30 mg of polyethylene glycol, which has been associated with nerve damage in experimental models (35). Epidural steroids can suppress plasma cortisol levels for about 3 to 5 weeks. Iatrogenic Cushing’s syndrome, fluid retention, and elevated serum glucose levels also can occur (36). Infection or bleeding into the closed epidural space also can result in significant neurologic deficits, including quadriplegia. Any complaint of increasing pain or neurologic changes...
should be investigated immediately. Early recognition can prevent permanent injury.

**Lumbar Sympathetic Block**

_Indications_

Lumbar sympathetic blockade is useful for diagnosing and treating pain of sympathetic origin. This includes pain involving the pelvis and lower extremity, secondary to complex regional pain syndrome, vasospastic disorders, and phantom pain.

_Techniques_

The spinous process of the second or third lumbar vertebra is palpated, and a point is marked 4 cm lateral to the middle of the spinous process. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. A 4-in. (10-cm), 22-gauge needle is inserted and advanced cephalad until contact is made with the transverse process. The needle is pulled back to skin level and redirected to pass between the transverse processes and alongside the anterolateral aspect of the vertebral body. After negative aspiration, 20 mL of local anesthetic is injected (Fig. 67-10).

_Comments_

This technique is commonly used for differential diagnosis and is the preferred treatment of sympathetic mediated pain involving the lower extremity. The lumbar sympathetic ganglia are located along the anterior lateral surface of the lumbar vertebral bodies and anteromedial to the psoas muscle. Autonomic mediated pain does not usually correspond to segmental or peripheral nerve distribution. Lateral and anteroposterior fluoroscopic views are recommended to ensure that the needle is properly positioned. It is recommended that intravenous access be available before the block in the event of local anesthetic toxicity and resulting seizure activity. Rarely are ablative nerve procedures (neurolysis) required in the management of sympathetic mediated pain.


**FIGURE 67-10.** Lumbar sympathetic approach. Fluoroscopic approach for lumbar sympathetic injection and neural blockade (arrow).
Complications
These injections are normally completed with fluoroscopic
guidance, as noted in Chapter 68. Performance of this pro-
cedure outside of a fully monitored environment is not rec-
commended. Resuscitation equipment and personnel must
be readily available. Although appearing technically simple,
this block has multiple hazards, owing to the proximity to
the aorta, inferior vena cava, kidney, pancreas, and intestines.
Intraarterial or intradural injection of local anesthetic may
result in death, seizure, respiratory arrest, cardiac arrest, cere-
bral damage with multiple sequelae, and other lesser complica-
tions. The risk for intravascular injection may be reduced if
a test dose is given, the total dose is injected incrementally,
and aspiration is performed before each injection. Significant
hypotension from the sympathetic block is possible but usually
does not occur with unilateral block.

Caudal Injection
Indications
The caudal approach to the epidural space is used to treat pain
in the lower back and pelvis.

Techniques
After informed consent is obtained, the patient may be placed
in a variety of positions, with patient comfort probably as
the prime concern. The preferred position is the lateral Sims’
position with the left side down for right-handed clinicians.
With the upper leg flexed, the buttocks are separated, allowing
easy access to the sacral-coccygeal junction. The patient
is prepared in a standard aseptic fashion over an area large
enough to allow palpation of landmarks, and sterile technique
is used throughout the procedure. The midline is identified
by palpating the tip of the coccyx with a finger and moving
cephalad about 4 to 5 cm in an adult, until the fingertip lies
over the sacral hiatus with the sacral cornua palpable on each
side. The palpating hand is kept in position, and a 2-in. (5-cm),
18-gauge short, beveled needle is inserted. The initial angle of
insertion is about 120 degrees to the coccyx. A “pop” is felt as
the sacrococcygeal ligament is penetrated. The needle is then
depressed to align with the long axis of the canal and inserted
1 cm. Once the caudal space has been entered, epidural posi-
tioning is confirmed by negative aspiration blood or CSF, then
40 to 80 mg of methylprednisolone acetate or 6 mg mixture
of betamethasone sodium phosphate and acetate is injected.
The steroids can be injected as is or diluted in 5 to 10 mL of
preservative-free normal saline (Fig. 67-11).

Comments
Epiduroscopy, a technique used to visualize the lumbar epidural
space, depends on this approach because the fiberoptic catheter

---

**FIGURE 67-11.** Caudal injection. A:Landmarks and approach for needle insertion.
B: Needle insertion through sacral-coccygeal membrane for injection. (From Willis RJ,
Caudal epidural blockade. In: Cousins MJ, Bridenbaugh PO, eds. Neural Blockade in
1998:333, with permission.)
cannot tolerate bending. Advantages of this approach include minimal risk for inadvertent dural puncture. Continuous catheter techniques can be used, but maintenance of site cleanliness is more difficult when compared with the lumbar approach to the epidural space. A Tuohy needle is not used for catheter placement because it will direct the catheter against the wall of the caudal canal and make catheter advancement difficult. Caudal epidurals and epidural lysis of adhesions can be performed via radiopaque catheter through a nonshearing needle under fluoroscopic guidance. Betamethasone sodium phosphate and acetate mixture is best used for localized nerve root irritation. Triamcinolone diacetate is water soluble and results in optimal outcome in generalized nerve root irritation such as arachnoiditis.

Complications
Improper needle placement can result in inadequate or absent block. This is due to variability in anatomy and inexperience. Rapid injection of large volumes of fluid is not recommended because this may result in large increases in CSF pressures, with the risk for cerebral hemorrhage, visual disturbances, headache, or compromised spinal cord blood flow. Pain at the injection site is a common complaint. Urinary retention can result from local anesthetic injection and should last only as long as the block.

Lateral Femoral Cutaneous Nerve Block
Indications
Lateral femoral cutaneous nerve blockade is useful for diagnosing and treating pain in the lateral thigh, thought to be from irritation of this nerve.

Techniques
After informed consent is obtained, the patient is placed in a supine position, and the anterosuperior iliac spine is palpated. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. A 1½-in. (4-cm), 21- to 23-gauge needle is inserted 1 cm medially and below the anterosuperior iliac spine. The needle is advanced deeply into the fascia lata toward the shelving of the iliac crest. After negative aspiration, 5 mL of local anesthetic is injected in a fanwise manner (Fig. 67-12).

This nerve can be blocked through an alternative approach by directing the needle superiorly beneath the inguinal ligament into the fascial compartment containing the nerve above the level of the inguinal ligament. This fascial compartment can be identified by directing a short, beveled needle medially to the anterosuperior iliac spine and advancing through the external oblique aponeurosis, the internal oblique muscle, and the fascia iliaca. The short, beveled needle allows the physician to feel a distinct loss of resistance or characteristic pop as the two fascial layers are penetrated. After negative aspiration, 5 mL of local anesthetic is injected to block the lateral femoral cutaneous nerve.

Comments
The lateral femoral cutaneous nerve emerges along the lateral border of the psoas muscle below the ilioinguinal nerve. It runs obliquely under the iliac fascia across the iliacus muscle and enters the thigh by passing posteriorly to the inguinal ligament, just medial to the anterosuperior iliac spine. It provides cutaneous innervation to the lateral aspect of the thigh to the knee. A large area over the lateral aspect of the thigh can be easily blocked with this technique.

Complications
The lateral femoral cutaneous nerve block has no significant complications, with the rare exception of a dysesthesia if the nerve is injured during the injection. Severe pain on injection suggests the possibility of an intraneural injection, and the needle should be immediately repositioned. It is possible to block the femoral nerve inadvertently when large amounts of local anesthetic is injected, resulting in a temporary weakness of knee extension and impaired ambulation. This occurs secondary to the medial spread of local anesthetic beneath the fascia iliaca.

Femoral Nerve Block
Indications
Femoral nerve blockade is useful in conjunction with other lower extremity blocks in treating complex regional pain...
syndrome and as an aid to decrease knee and ankle pain during physical therapy.

**Techniques**

After informed consent is obtained, the patient is placed in the supine position, and the femoral artery is located. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. The femoral artery is palpated below the inguinal ligament. A 1½-in. (4-cm), 22-gauge needle is inserted 1 to 2 cm below the inguinal ligament and lateral to the femoral artery. The needle is advanced in a lateral and posterior direction just distal to the inguinal ligament. A characteristic pop, when using a short, beveled needle, can be used to identify penetration of the fascia lata and the fascia iliaca, remembering that the femoral nerve lies deeper than both. When a nerve stimulator is used, contraction of the quadriceps muscle confirms correct placement of the needle. After negative aspiration, 10 mL of local anesthetic is injected to block the femoral nerve (Fig. 67-13).

**Comments**

At the level of the inguinal ligament, the femoral nerve lies anterior to the iliopsoas muscle and slightly lateral to the femoral artery. It does not lie within the femoral sheath. The nerve lies underneath the fascia lata and fascia iliaca within its own sheath. At the level of the inguinal ligament, the femoral nerve divides into anterior (superficial) and posterior (deep) bundles. The anterior bundle provides cutaneous innervation of the skin overlying the anterior surface of the thigh as well as providing motor innervation to the sartorius muscle. The posterior bundle provides innervation to the quadriceps muscles and the knee joint. It also gives off the saphenous nerve, which supplies cutaneous innervation to the medial aspect of the calf to the level of the medial malleolus. A catheter also can be placed within the femoral nerve sheath for continuous infusion of local anesthetics.

It is important to remember that the upper portion of the anterior thigh is innervated by the ilioinguinal and genitofemoral nerves and is not blocked when performing a femoral nerve block.

**Complications**

Significant complications associated with femoral nerve blockade are uncommon. Dysesthesia may result if the nerve is injured during the injection. Hematoma at the site is a possibility but is usually not clinically significant. If an arterial puncture occurs, prolonged direct pressure is usually adequate to prevent the development of a hematoma. The presence of a femoral artery vascular graft is a relative contraindication to femoral nerve blockade.

**Obturator Nerve Block**

**Indications**

Obturator nerve blockade is extremely useful as a diagnostic, prognostic, or therapeutic procedure in patients with adductor spasm that interferes with rehabilitation or personal hygiene.

**Techniques**

After informed consent is obtained, the patient is placed in the supine position with the leg to be blocked placed in slight abduction. It is not necessary to shave the pubic area. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. A 3-in. (8-cm), 22-gauge needle is inserted perpendicular to the skin at a point 1.5 cm lateral and inferior to the pubic tubercle. The needle depth at which the bone is contacted should be noted. The needle is withdrawn to skin level and redirected in a lateral and slightly superior direction, parallel to the superior ramus of the pubis. The needle is advanced until the inferior ramus of the pubis is contacted. The needle depth at which the bone is contacted should be noted. The needle is withdrawn to skin level and redirected in a lateral and slightly superior direction, parallel to the superior ramus of the pubis. The needle is advanced 2 to 3 cm beyond the previously noted depth until a paresthesia is elicited. A nerve stimulator makes it relatively easy to identify the obturator nerve by adductor muscle contraction. After negative aspiration for blood, 10 mL of local anesthetic is injected to block the obturator nerve. This traditional approach was first described by Labat (37) (Fig. 67-14).

Using the above techniques, Wassef has described an alternative approach using the femoral artery and adductor longus tendon as landmarks (38). A mark is made on the skin 1 to 2 cm medial to the femoral artery just below the inguinal ligament. This mark is used to indicate the direction of the.
needle toward the obturator canal. The adductor longus tendon is then identified near its insertion site at the pubis. A 3-in. (8-cm), 22-gauge insulated needle is introduced behind the adductor longus tendon and directed laterally, with a slight posterior and superior inclination toward the skin mark. The needle is advanced until adductor muscle contraction is elicited with a nerve stimulator (39).

Comments
The obturator nerve is formed by the union of the ventral branches of the anterior primary rami of L2, L3, and L4 within the substance of the psoas muscle. It emerges from the medial border of the psoas muscle at the brim of the pelvis. The nerve runs caudad and anteriorly along the lateral wall of the pelvis, along the obturator vessels to the obturator foramen. There it enters the thigh, supplying the adductor muscles and providing innervation to the hip and knee joints.

As the nerve passes through the obturator canal, it divides into anterior and posterior branches. The anterior branch supplies the hip joint, the anterior adductor muscles, and cutaneous branches to the medial aspect of the thigh. The cutaneous innervation of the obturator nerve can be extremely variable and can be nonexistent in some people. The posterior branch supplies the deep adductor muscles and frequently sends a branch to the knee joint.

This procedure is often performed on rehabilitation patients with spasticity or contractures that result in positioning difficulty. Confirmation of a successful obturator nerve block is demonstrated by paresis of the adductor muscles because the cutaneous contribution of the obturator nerve is inconsistent. An alternative to this procedure is selective root blockade at levels L2, L3, and L4 using a nerve stimulator to establish muscle innervation.

Complications
Hematoma and intravascular injection are possible due to the close proximity of the obturator vessels. If an arterial puncture occurs, prolonged direct pressure is usually adequate to prevent the development of a hematoma.

Sciatic Nerve Block
Indications
Sciatic nerve blockade is typically used to treat painful conditions of the lower leg such as complex regional pain syndrome and to facilitate physical therapy by decreasing pain in the lower extremity.

Techniques
A regional block of the sciatic nerve can be achieved anywhere along the course of the nerve. Most of the approaches have been developed mainly to avoid positioning problems that may be present in trauma patients or elderly people. The nerve can be blocked at the sciatic notch, at the level of the ischial tuberosity, greater trochanter, or superior aspect of the popliteal fossa.

Classic Approach
The classic technique described by Labat (37) blocks the nerve at the level of the greater sciatic notch, using the piriformis muscle as a landmark. After informed consent is obtained, the patient is placed in the lateral Sims’ position with the side to be blocked uppermost. The upper knee is flexed, and the patient’s back is rotated slightly forward. Some patients may find this position uncomfortable, particularly those with orthopedic problems.

The landmarks are the cephalad portion of the greater trochanter and the posterosuperior iliac spine. A line is drawn connecting these two points, corresponding to the superior border of the piriformis muscle and the upper border of the sciatic notch. A perpendicular line is drawn distally from the midpoint of the first line. The point of injection is 3 to 5 cm distal on the perpendicular line. Verification of the insertion point can be made by drawing a third line connecting the cephalad portion of the greater trochanter and the sacrococcygeal joint. This third line is used to compensate for the height of the patient. The intersection of lines 2 and 3 is the point of needle insertion (Fig. 67-15).

The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. A 4- to 5-in. (10- to 12-cm), 22-gauge spinal needle is...
introduced at right angles to the skin and advanced to a depth of 6 to 10 cm until a paresthesia is reported in the distribution of the sciatic nerve, preferably involving the foot. If periosteum is contacted, the needle is then redirected medially or superiorly. Touching the periosteum may produce a local paresthesia, which could be mistaken for a true sciatic nerve paresthesia. A nerve stimulator is extremely helpful in locating the nerve (39).

Doppler ultrasound also can be used to locate the dominant arterial structure within the sciatic notch (40). The needle is then advanced in the same orientation as the probe until a paresthesia is obtained. Successful blockade has been reported after one or two attempts in 70% of patients. After negative aspiration for blood, 20 to 30 mL of local anesthetic is injected to block the sciatic nerve.

A continuous sciatic nerve block can be performed by using a standard 16-gauge intravenous infusion cannula attached to a nerve stimulator. After obtaining muscle contraction in the lower leg, preferably dorsal or plantar flexion of the foot, an epidural catheter is advanced about 6 cm into the neurovascular space. Continuous infusion of a local anesthetic using an infusion pump can then be used to provide continuous analgesia (41). With the classic approach, both the posterior femoral cutaneous and pudendal nerves are usually blocked with the sciatic nerve.

**Posterior Approach**

An alternate approach may be used, with the patient positioned as above or prone. The ischial tuberosity and the greater trochanter are identified and a line drawn connecting these two points. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. A 3- to 4-in. (8- to 10-cm), 22-gauge spinal needle is inserted at the midpoint of the line until a paresthesia is elicited in the lower leg. After negative aspiration, 20 to 30 mL of local anesthetic is injected to block the sural nerve. The posterior femoral cutaneous nerve is often blocked at this level, but the pudendal nerve is frequently spared.

**Anterior Approach**

The anterior approach allows the sciatic nerve to be blocked without moving the patient, enabling the patient to remain in the supine position (37,42). This approach is especially helpful in trauma patients with a painful leg, but it is quite painful, and sedation is often necessary. The nerve is very deep at this point and can be difficult to locate. In adults, the sciatic neurovascular compartment is usually 4.5 to 6 cm below the surface of the femur. In children, however, the distance varies according to age and size of the child (43). The use of a nerve stimulator is advised in identifying the nerve. The posterior cutaneous nerve of the thigh may not be blocked with this approach, as tourniquet pain could result if a thigh tourniquet is applied (Fig. 67-16).

The patient is placed in the supine position with the leg in a neutral position. The anterosuperior iliac spine and the pubic tubercle are identified and marked. A line is then drawn connecting these two points, overlying the inguinal ligament, and trisected into equal parts. A perpendicular line is drawn distally from the junction of the medial and middle thirds.
A third line is drawn parallel to the first, starting from the cephalad aspect of the greater trochanter. The point of intersection of this third line and the perpendicular line is the insertion point of the needle.

The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. A 6-in. (15-cm), 22-gauge spinal needle is inserted and directed slightly laterally from a plane perpendicular to the skin. The needle is advanced until periosteum is contacted (usually the lesser trochanter). The needle is then partially withdrawn and redirected posteriorly and medially to slide beneath the femoral shaft until a paresthesia, or contraction of the calf or the anterior compartment muscles, occurs with the use of a nerve stimulator. After negative aspiration, 20 to 30 mL of local anesthetic is injected to block the sural nerve.

**Lateral Approach**

The lateral approach, initially described by Ichiyanagi (44), was found to be very difficult and never became popular. A new lateral approach described by Guardini et al. (45) is easier. It blocks the sciatic nerve just posterior to the quadratus femoris muscle in the subgluteal space.

The greater trochanter is identified, and the patient is prepared in a standard sterile fashion. A 5- to 6-in. (12- to 15-cm), 22-gauge spinal needle is advanced 3 cm distal to the maximum lateral prominence of the trochanter, close to its posterior margin. The needle is inserted until the periosteum is contacted. The needle is then partially withdrawn and redirected posteriorly and medially to slide beneath the femoral shaft until a paresthesia is elicited. After negative aspiration, 20 to 25 mL of local anesthetic agent is injected to block the sural nerve.

The main advantage of this technique is that the patient can remain in the supine position and the leg need not be manipulated. When using a nerve stimulator, it is important to make sure that the muscle contractions occur in the calf muscles or in the muscles of the anterior compartment. It is possible with this technique to stimulate inadvertently the nerve branch supplying the two heads of the biceps femoris muscle, producing thigh muscle contraction and misplacement of the local anesthetic.

**Comments**

The sciatic nerve is the largest in the body. It arises from both the lumbar and sacral plexuses. Anatomically, the sciatic nerve consists of two major nerve trunks: the tibial and common peroneal components. The tibial nerve is derived from the anterior rami of L4 to S3 nerve roots. The common peroneal nerve is derived from the dorsal branches of the anterior rami of the same roots. It leaves the pelvis along with the posterior cutaneous nerve of the thigh through the sciatic foramen beneath the inferior margin of the piriformis muscle. It passes halfway between the greater trochanter and the ischial tuberosity. It becomes superficial at the inferior border of the gluteus maximus muscle and travels downward the posterior aspect of the thigh. At the superior aspect of the popliteal fossa, the sciatic nerve physically separates into the tibial and common peroneal nerves.

In the past, the sciatic nerve block was considered unreliable, technically difficult, and uncomfortable for the patient. Sedation was often required, and this interfered with the patient’s ability to provide accurate verbal feedback. This was especially the case if a paresthesia was used to identify the nerve. Reported rates of success ranged between 33% and 95% using various techniques. Today’s insulated needles and nerve stimulators have made it easier to perform this block safely in sedated or even anesthetized patients with a higher rate of success.
NERVE BLOCKS AT THE KNEE

Nerve blockade at the knee is primarily used to treat pain disorders in specific nerve distributions. Diagnostic tibial nerve blocks also can be helpful in evaluating patients with spastic hemiparesis or myotonic disorders (46). Many early textbooks of regional blockade discouraged individual nerve blocks at the knee because they were thought to be difficult to perform and there was a possibility of a postanesthetic neuritis (47). Recently, studies have shown these nerve blocks can be safely and successfully performed at the knee, even in children (48,49).

The common peroneal and tibial nerves are extensions of the sciatic nerve. The sciatic nerve bifurcates at the superior aspect of the popliteal fossa, bordered by the biceps femoris muscle laterally and the semimembranosus and semitendinosus muscles medially. The two heads of the gastrocnemius muscle border the lower half of the popliteal fossa. Techniques have been described in which the tibial and common peroneal nerves are blocked with one injection, but it is possible to miss one of the branches (50). Identifying the two nerves separately and performing individual nerve blocks increases the likelihood of success.

Tibial Nerve Block

Indications
Tibial nerve blockade is useful as a diagnostic, prognostic, or therapeutic procedure in painful disorders involving the ankle and foot.

Techniques
After informed consent is obtained, the patient is placed in the prone position. The knee is flexed to allow palpation of the superior popliteal fossa borders and identification of the skin crease behind the knee joint. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. A 1½-in. (3– to 4-cm), 21- to 23-gauge needle is inserted just above the crease line in the middle of the popliteal fossa. A nerve stimulator is used to identify the tibial nerve by eliciting plantar flexion of the foot. The average depth from skin to nerve in adults is 1.5 to 2 cm. After negative aspiration, 5 mL of local anesthetic is injected to block the tibial nerve (Fig. 67-17).

Comments
The tibial nerve is the larger of the two branches of the sciatic nerve and supplies motor innervation to the flexor muscles at the back of the knee joint and calf. The cutaneous innervation supplies the skin overlying the popliteal fossa and down the back of the leg to the ankle. It travels through the center of the popliteal fossa as it proceeds distally down the leg.

Complications
Although the sciatic nerve is composed of mostly somatic nerves, it has a sympathetic component. The resulting sympathetic block may allow some mild venous pooling, but this is usually insufficient to cause clinically significant hypotension. Residual dysesthesias have been reported but usually improve in 1 to 3 days. This may be the result of nerve injury from the use of long beveled needles. Using short beveled needles for regional blocks may decrease the incidence of nerve injury.

Common Peroneal Nerve Block

Indications
Common peroneal nerve blockade is useful as a diagnostic, prognostic, or therapeutic procedure in painful disorders involving the ankle and foot.

Techniques
After informed consent is obtained, the patient is placed in the supine or lateral position. The common peroneal nerve can be easily palpated as it crosses the neck of the fibula. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. A 1-in. (2.5-cm), 25-gauge needle is inserted next to the nerve and advanced to contact the periosteum, with care being taken to avoid an intraneural injection. A sudden, intense pain on injection suggests intraneural injection. If this occurs, the needle should be repositioned before proceeding. A nerve stimulator may be used to identify the nerve by eliciting contraction of the anterior compartment muscles. The needle is withdrawn slightly, and after negative aspiration, 5 mL of local anesthetic is injected to block the peroneal nerve (see Fig. 67-17).

Comments
The common peroneal nerve is about half the size of the tibial nerve and contains articular branches to the knee joint. It provides motor innervation to the extensor muscles of the foot and cutaneous nerves to the lateral aspect of the leg, heel, and ankle. It separates from the tibial nerve at the superior aspect of the popliteal fossa and courses laterally around the fibular head where it divides into the deep and superficial peroneal nerves.

Complications
Complications from the common peroneal nerve block are rare, especially when care is taken to avoid an intraneural injection. Severe pain on injection suggests the possibility of an intraneural injection, and the needle should be immediately repositioned.

Saphenous Nerve Block

Indications
Saphenous nerve blockade is useful as a diagnostic, prognostic, or therapeutic procedure in painful disorders involving the knee, ankle, and foot.
After informed consent is obtained, the patient is placed in the supine or lateral position. The saphenous nerve is located at the medial surface of the medial condyle of the femur at about the same level as the apex of the patella. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. A 1½-in. (4-cm), 25-gauge needle is inserted perpendicular to the skin just below the medial surface of the tibial condyle. After negative aspiration, 5 to 10 mL of local anesthetic is injected subcutaneously to block the saphenous nerve (Fig. 67-18).

Comments
The saphenous nerve is the terminal branch of the femoral nerve. It provides cutaneous innervation to the skin overlying the medial, anteromedial, and posteromedial aspects of the leg from just above the knee to the level of the medial malleolus and, in some patients, to the medial aspect at the base of the great toe. There is no motor component.

Complications
The saphenous vein may accompany the saphenous nerve, and the patient should be made aware of the possibility of a hematoma from venous puncture. Other complications from the saphenous nerve block are rare, especially when care is taken to avoid an intraneural injection. Severe pain on injection suggests the possibility of an intraneural injection, and the needle should be repositioned immediately.

NERVE BLOCKS AT THE ANKLE

Nerve blockade at the ankle is primarily used to treat pain disorders in specific nerve distributions. There are five terminal branches of the tibial, common peroneal, and femoral nerves that supply the ankle and foot: posterior tibial, sural, superficial peroneal, deep peroneal, and saphenous nerves. These nerves are relatively easy to block at the level of the ankle.

In general, five nerve blocks form a ring of infiltration around the ankle at the level of the malleolus. It is important to
remember that large volumes of local anesthetic, especially those containing epinephrine, may cause vascular occlusion. Otherwise, neural blockade at the ankle is safe and highly successful.

**Tibial Nerve Block**

**Indications**

Tibial nerve blockade is used to treat pain disorders in the tibial nerve distribution of the foot.

**Techniques**

After informed consent is obtained, the patient is placed in the prone position with the foot supported by a pillow. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. A skin wheal is raised along the medial aspect of the Achilles tendon at the level of the superior border of the medial malleolus. A 1-in. (2.5-cm), 25-gauge needle is advanced through the wheal toward the posterior aspect of the tibia, behind the posterior tibial artery. If a paresthesia is elicited after negative aspiration, 3 to 5 mL of local anesthetic is injected after negative aspiration. If a paresthesia is not elicited, the needle is advanced until the tibial periosteum is contacted. The needle is withdrawn 0.5 cm, and after negative aspiration, 5 to 7 mL of local anesthetic is injected to block the posterior tibial nerve. A nerve stimulator may be used to identify the posterior tibial nerve by eliciting contraction of muscles in the sole of the foot (Fig. 67-19).

**Comments**

The posterior tibial nerve is located along the medial aspect of the Achilles tendon, lying just behind the posterior tibial artery. The nerve gives off a medial calcaneal branch to the medial aspect of the heel, then divides behind the medial malleolus into the medial and lateral plantar nerves. The medial plantar nerve supplies the medial two thirds of the sole of the foot as well as the plantar portion of the medial three and one-half toes. The lateral plantar nerve supplies the lateral one third of the sole and the plantar portion of the lateral one and one-half toes.

**Complications**

Intraneural injection may result in nerve damage. Severe pain on injection suggests the possibility of an intraneural injection, and the needle should be repositioned immediately. Hematoma and intravascular injection are possible due to the close proximity of the posterior tibial vessels. If an arterial puncture occurs, prolonged direct pressure is usually adequate to prevent the development of a hematoma.

**Sural Nerve Block**

**Indications**

Sural nerve blockade is used to diagnose and treat pain disorders in the sural nerve distribution.

**Techniques**

After informed consent is obtained, the patient is placed in a prone position with the foot supported by a pillow. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. A skin wheal is raised lateral to the Achilles tendon at the level of the lateral malleolus. A 1-in. (2.5-cm), 25-gauge needle is inserted to a depth of 1 cm, directed toward the lateral border of the fibula. If a paresthesia is elicited, 2 to 3 mL of local anesthetic is injected after negative aspiration. If a paresthesia cannot be elicited, after negative aspiration, 3 to 5 mL of local anesthetic is injected subcutaneously in a fan distribution from the lateral border of the Achilles tendon to the lateral border of the fibula to block the sural nerve (see Fig. 67-19).
The sural nerve is a cutaneous nerve that contains fibers from both the tibial and common peroneal nerves. It lies subcutaneous somewhat distally to the middle of the leg and travels with the short saphenous vein behind and below the lateral malleolus. It supplies the posterolateral surface of the leg, the lateral side of the foot, and the lateral aspect of the fifth toe.

Complications
Intraneural injection may result in nerve damage. Severe pain on injection suggests the possibility of an intraneural injection, and the needle should be immediately repositioned. Hematoma and intravascular injection are possible, owing to the close proximity of the sural vessels. If an arterial puncture occurs, prolonged direct pressure is usually adequate to prevent the development of a hematoma.

Superficial Peroneal Nerve Block
Indications
Superficial peroneal nerve blockade is used to diagnose and treat pain disorders of the superficial peroneal nerve distribution in the foot.

Techniques
After informed consent is obtained, the patient is placed in a supine position with the foot elevated on a pillow. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. A 1-in. (2.5-cm), 25-gauge needle is inserted between the extensor hallucis longus tendon and the anterior tibial tendon, just superior to the level of the malleoli. The extensor hallucis longus tendon can easily be identified by having the patient extend the great toe. If the artery can be palpated, the needle is placed just laterally to the artery. The needle is advanced toward the tibia, and after negative aspiration, 3 to 5 mL of local anesthetic is injected deep to the fascia to block the deep peroneal nerve (see Fig. 67-19).

The superficial peroneal nerve exits the deep fascia of the leg at the anterior aspect of the distal two thirds of the leg. From that point, the superficial peroneal nerve runs subcutaneously to supply the dorsum of the foot and toes, with the exception of the contigous surfaces of the great and second toes.

Complications
Complications are rare with the superficial peroneal nerve block.

Deep Peroneal Nerve Block
Indications
Deep peroneal nerve blockade is used to diagnose and treat pain disorders in the deep peroneal nerve distribution of the foot.

Techniques
After informed consent is obtained, the patient is placed in a supine position with the foot elevated on a pillow. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. A 1-in. (2.5-cm), 25-gauge needle is inserted between the extensor hallucis longus tendon and the anterior tibial tendon, just superior to the level of the malleoli. The extensor hallucis longus tendon can easily be identified by having the patient extend the great toe. If the artery can be palpated, the needle is placed just laterally to the artery. The needle is advanced toward the tibia, and after negative aspiration, 3 to 5 mL of local anesthetic is injected deep to the fascia to block the deep peroneal nerve (see Fig. 67-19).

Comments
The superficial peroneal nerve travels down the anterior portion of the interosseus membrane of the leg and extends midway between the malleoli onto the dorsum of the foot. At this point, the nerve lies laterally to the extensor hallucis longus tendon and...
A 1½- to 4-in. (4- to 10-cm) insulated needle is advanced raised over the main muscle bulk of the muscles to be injected. The area large enough to allow palpation of landmarks, and sterile technique is used to allow optimal access to the muscles involved. After informed consent is obtained, the patient is positioned prone position with the foot elevated on a pillow. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. A 1-in. (2.5-cm), 25-gauge needle is inserted immediately above and anteriorly to the medial malleolus and advanced to the anterior border of the tibia. After negative aspiration, 3 to 5 mL of local anesthetic is injected over the course of the needle to block the saphenous nerve (see Fig. 67-19).

**Complications**
Hematoma and intravascular injection are possible due to the close proximity of the anterior tibial vessels. If an arterial puncture occurs, prolonged direct pressure is usually adequate to prevent the development of a hematoma.

**Saphenous Nerve Block**
**Indications**
Saphenous nerve blockade is used to diagnose and treat pain disorders of the saphenous nerve distribution in the foot.

**Techniques**
After informed consent is obtained, the patient is placed in a prone position with the foot elevated on a pillow. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. A 1-in. (2.5-cm), 25-gauge needle is inserted immediately above and anteriorly to the medial malleolus and advanced to the anterior border of the tibia. After negative aspiration, 3 to 5 mL of local anesthetic is injected over the course of the needle to block the saphenous nerve (see Fig. 67-19).

**Comments**
The saphenous nerve is the terminal branch of the femoral nerve. It becomes cutaneous at the lateral aspect of the knee joint and follows the great saphenous vein to the medial malleolus. It supplies cutaneous innervation to the medial aspect of the lower leg anterior to the medial malleolus and the medial aspect of the foot, and may extend as far forward as the metatarsophalangeal joint.

**Complications**
Hematoma and intravascular injection are possible due to the close proximity of the great saphenous vessels. If an arterial puncture occurs, prolonged direct pressure is usually adequate to prevent the development of a hematoma.

**Intramuscular Nerve (Motor Point) Block**
**Indications**
Intramuscular nerve blockade is used for diagnostic, prognostic, and therapeutic treatment of non–velocity-dependent muscle tone, flexor spasm, and dystonia.

**Techniques**
After informed consent is obtained, the patient is positioned comfortably to allow optimal access to the muscles involved. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. A skin wheal is raised over the main muscle bulk of the muscles to be injected. A 1½- to 4-in. (4- to 10-cm) insulated needle is advanced through the wheal, with a nerve stimulator used to localize the motor nerve branches or motor points. The current is reduced until the minimum current is required to elicit muscle contraction. When the needle tip is within 1 mm of the motor nerve, and after negative aspiration, 1 to 2 mL of 4% to 6% phenol is injected for neurolysis (see Fig. 67-1).

**Comments**
Intramuscular nerve or motor point blockade is reported to have duration of effect from 1 to 36 months (median, 11.5 months). No dose-response or dose-duration of effects relationship has been demonstrated for motor point blocks (51,52). The needle is positioned to produce the maximal twitch at the lowest stimulus. The needle is usually adjacent to the nerve when 0.5 to 0.1 mA produces motor stimulation with an insulated needle, and 1 mA with an uninsulated needle. The motor points of each muscle cluster at the midpoint of the muscle fibers.

**Complications**
Significant complications are rare with intramuscular nerve injections, and transitory side effects include pain of mild intensity, tenderness and swelling at injection sites, and dysesthesia. Inadvertent neurolysis of a mixed nerve results in painful paresthesia in about 11% of patients.

**COMMON MUSCLE INJECTION TECHNIQUES**

**General**
Trigger points may occur in any muscle or muscle group of the body. They are commonly found in muscle groups that are routinely overstressed or those that do not undergo full contraction and relaxation cycles. Many trigger points are characterized by pain originating from small circumscribed areas of local hyperirritability involving myofascial structures, resulting in local and referred pain. Pain is aggravated by stretching, cooling, and compression of the affected area, which often gives rise to a characteristic pattern of referred pain (11,12).

Trigger points are best localized by deep palpation of the affected muscle, which reproduces the patient’s pain complaint both locally and in the referred zone. Trigger points are usually a sharply circumscribed spot of exquisite tenderness when they are present; passive or active stretching of the affected muscle routinely increases the pain. When compared with equivalent palpation pressure in normal muscle, the trigger point region displays isolated bands, increased tenderness, and referred pain. The muscle in the immediate vicinity of the trigger point is often described as ropey, tense, or a palpable band. The trigger point is injected after palpation of the affected muscle, and the point of maximal tenderness reproducing the pain complaint is identified. When the point of injection has been determined, it is best marked with the tip of a retracted ballpoint pen or needle hub by pressing the skin to reproduce temporary indentation to mark the point of entry. The patient is prepared in a standard aseptic fashion over an area large

the anterior tibial artery. It supplies motor innervation to the short extensors of the toes and cutaneous innervation to adjacent areas of the first and second toes.

**Complications**
Hematoma and intravascular injection are possible due to the close proximity of the anterior tibial vessels. If an arterial puncture occurs, prolonged direct pressure is usually adequate to prevent the development of a hematoma.
enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. The skin and subcutaneous tissue at the injection site are usually not anesthetized.

A 1½- to 2-in. (4- to 5-cm), 22- to 25-gauge needle is advanced into the muscle at the point of maximum tenderness. Before injecting the medication, an attempt should always be made to aspirate to avoid accidental or intravascular injection. Verification that the needle is at the trigger point may be established by the jump sign or reproduction of the pain complaint. Medication should be injected in a fanwise manner in the area of the trigger point.

**Indications for Trigger Point Injections**

Trigger point injections may be used to determine the source of pain and to provide maximum pain relief from myofascial pain and to facilitate physical therapy for the stretching of trigger points.

**Contraindications for Trigger Point Injections**

Absolute contraindications to trigger point injection include localized infection, a skin condition that prevents adequate skin preparation, the existence of a tumor at the injection site, history of allergy to local anesthetics, gross coagulation defects, septicemia, or an uncooperative patient.

**Complications**

The complications associated with trigger point injections include infection, increased pain, local anesthetic overdose, or intravascular injection that can result in CNS toxicity and, in some cases, pulmonary and cardiac arrest. Intraneural injection may result in nerve damage. Severe pain on injection suggests the possibility of an intraneural injection, and the needle should be immediately repositioned. Other complications depend on location of the trigger point injection and are discussed separately.

**Techniques**

Before the injection, the affected muscle is palpated and the trigger points located and marked. The skin is scrubbed with antiseptic and allowed to dry for 2 minutes. The wearing of sterile gloves is required so that muscle in the sterile field may be palpated throughout the procedure. Before the injection, the trigger point is repalpated and stabilized between the fingers for injection (Fig. 67-20). Routinely, a 1½-in. (4-cm), 21- to 25-gauge needle transverses the skin’s subcutaneous tissue and is advanced smoothly into the area of the trigger point.

Aspiration should be done to ensure there is no intravascular penetration. If this does occur, the needle should be repositioned and aspirated to ensure that blood vessels have been avoided, and then the medication is injected. A fanwise manner of injection often results in the longest pain relief, owing to increased distribution of local anesthetic (Fig. 67-21). The needle is then withdrawn, with pressure applied to minimize bleeding.

**SPECIFIC TRIGGER POINT INJECTIONS**

**Trapezius**

**Indications**

Trigger point injection of the trapezius muscle is used to treat myofascial pain.

**Techniques**

After informed consent is obtained, the patient is placed in the sitting or prone position. The trapezius muscle is palpated. The injection sites are identified as points of maximal tenderness to deep palpation reproducing the patient’s pain complaint. This may or may not result in referred pain. The patient is prepared in a standard aseptic fashion over an area large
enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. A 1½-in. (4-cm), 21- to 25-gauge needle is inserted at the point of maximal tenderness and advanced to the area of the trigger point. After negative aspiration, the trigger point area is injected with 4 mL of local anesthetic (Fig. 67-22).

Comments
The referred pain pattern for the upper trapezius is often along the posterior lateral aspect of the neck, as well as periarticular and temporal regions. The referred pain pattern for the mid trapezius often involves the shoulder and paraspinal region. The referred pain pattern for the lower trapezius usually involves the paraspinal region. The patient should be fully familiar with the stretching program for the trapezius muscle and be instructed in a home program. Failure to include a home stretching program usually results in short-term relief.

Complications
Significant complications are uncommon with trapezius trigger point injections.

Levator Scapulae
Indications
Levator scapulae trigger point injection is a useful diagnostic and therapeutic procedure for myofascial pain.

Techniques
After informed consent is obtained, the patient is placed in a sitting or prone position. The levator scapulae muscle is palpated along the attachment at the C1 to C4 vertebrae and the superior angle of the scapula. The injection sites are identified as points of maximal tenderness to deep palpation reproducing the patient’s pain complaint. This may or may not result in referred pain. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. A 1½-in. (4-cm), 21- to 25-gauge needle is inserted at the point of maximal tenderness and advanced to the area of the trigger point. After negative aspiration, the trigger point area is injected with 4 mL of local anesthetic (Fig. 67-23).

Comment
The entire body of the levator scapulae muscle should be palpated from origin to insertion and all trigger points injected.
Total injection should not exceed maximum safe dosage. The referred pain pattern for the levator scapular muscle often includes a posterior lateral neck and occipital and temporal regions. The patient should be fully familiar with the stretching program for the levator scapulae muscle and be instructed in a home program. Failure to include a home stretching program usually results in short-term relief.

Complications
Nerve root blockade may result from improper needle placement or injection of large quantities of local anesthetic in the vertebral region. Intraneural injection may result in nerve damage. Severe pain on injection suggests the possibility of an intraneural injection, and the needle should be repositioned immediately.

Supraspinatus
Indications
Supraspinatus injection is a useful diagnostic and therapeutic procedure for myofascial pain.

Techniques
After informed consent is obtained, the patient is placed in a sitting or prone position. The supraspinatus muscle is palpated for trigger points from the supraspinous fossa to the humerus. The injection sites are identified as points of maximal tenderness to deep palpation reproducing the patient’s pain complaint. This may or may not result in referred pain. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. A 1½-in. (4-cm), 21- to 25-gauge needle is inserted at the point of maximal tenderness and advanced to the area of the trigger point. After negative aspiration, the trigger point area is injected with 4 mL of local anesthetic (Fig. 67-24).

Comment
The referred pain pattern for the supraspinatus muscle often involves the posterior lateral aspect of the shoulder and upper extremity. The patient should be fully familiar with
the stretching program for the supraspinatus muscle and be instructed in a home program. Failure to include a home stretching program usually results in short-term relief.

Complications
Significant complications are uncommon with supraspinatus trigger point injections.

Infraspinatus
Indications
Infraspinatus injection is a useful diagnostic and therapeutic procedure for myofascial pain.

Techniques
After informed consent is obtained, the patient is placed in the sitting or prone position. The infraspinatus muscle is palpated from the infraspinous fossa of the scapula to the humerus. Trigger points are most often located below the spine of the scapulae. The injection sites are identified as points of maximal tenderness to deep palpation, reproducing the patient’s pain complaint. This may or may not result in referred pain. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. A 1½-in. (4-cm), 21- to 25-gauge needle is inserted at the point of maximal tenderness and advanced to the area of the trigger point. After negative aspiration, the trigger point area is injected with 4 mL of local anesthetic (Fig. 67-25).

Comment
The referred pain pattern for the infraspinatus often involves the deltoid muscle, as well as the area over the lateral shoulder and proximal upper extremity. Pain also may be referred in the infrascapular region. The patient should be fully familiar with the stretching program for the infraspinatus muscle and be instructed in a home program. Failure to include a home stretching program usually results in short-term relief.

Complications
Significant complications are uncommon with infraspinatus trigger point injections.

Pectoralis
Indications
Pectoralis muscle injection is a useful diagnostic and therapeutic procedure for myofascial pain.

Techniques
After informed consent is obtained, the patient is placed in the supine position. The pectoralis muscles are palpated. The injection sites are identified as points of maximal tenderness to deep palpation, reproducing the patient’s pain complaints. This may or may not result in referred pain. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. A 1½-in. (4-cm), 21- to 25-gauge needle is inserted at the point of maximal tenderness and advanced to the area of the trigger point. After negative aspiration, the trigger point area is injected with 4 mL of local anesthetic (Fig. 67-26).
Comment
The referred pain pattern for the pectoralis muscles usually involves the anterior chest wall and breast regions. The patient should be fully familiar with the stretching program for the pectoralis muscle and be instructed in the home program. Failure to include a home stretching program usually results in short-term relief.

Complications
Significant complications are uncommon with pectoralis trigger point injections; however, the anatomy of the region, including the close proximity of the thoracic cavity, must be carefully considered. The risk for pneumothorax is reduced by approaching the trigger point with the needle tangential to the thoracic wall.

Deltoid
Indications
Deltoid muscle injection is a useful diagnostic and therapeutic procedure for myofascial pain.

Techniques
After informed consent is obtained, the patient is placed in the sitting position. The anterior, middle, and posterior components of the deltoid muscle are palpated. The injection sites are identified as points of maximal tenderness to deep palpation, reproducing the patient’s pain complaint. This may or may not result in referred pain. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. A 1½-in. (4-cm), 21- to 25-gauge needle is inserted at the point of maximal tenderness and advanced to the area of the trigger point. After negative aspiration, the trigger point area is injected with 4 mL of local anesthetic (Fig. 67–27).

Comment
The referred pain pattern for the deltoid muscle usually involves the shoulder and proximal upper extremity. The patient should be fully familiar with the stretching program for the deltoid muscle and be instructed in a home program. Failure to include a home stretching program usually results in short-term relief.

Complications
Significant complications are uncommon with deltoid trigger point injections.

Quadratus Lumborum
Indications
Quadratus lumborum injection is a useful diagnostic and therapeutic procedure for myofascial pain.

Techniques
After informed consent is obtained, the patient is placed in a prone position. The quadratus lumborum muscle is palpated from the 12th rib to the iliac crest and from vertebral attachments L1 to L4 to its lateral border. The injection sites are identified as points of maximal tenderness to deep palpation, reproducing the patient’s pain complaint. This may or may not result in referred pain. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. A 1½-in. (4-cm), 21- to 25-gauge needle is inserted at the point of maximal tenderness and advanced to the area of the trigger point. After negative aspiration, the trigger point area is injected with 4 mL of local anesthetic (Fig. 67–28).

Comment
The referred pain pattern for the quadratus lumborum muscle usually involves the iliac crest, hip, and buttock. The patient should be fully familiar with the stretching program for the quadratus lumborum muscle and be instructed in a home program. Failure to include a home stretching program usually results in short-term relief.

Complications
Significant complications are uncommon with quadratus lumborum trigger point injections.

Paraspinal Musculature
Indications
The paraspinal muscle injection is a useful diagnostic and therapeutic procedure for myofascial pain.
Techniques
After informed consent is obtained, the patient is placed in the prone position. The appropriate thoracic and lumbar regions are palpated. The injection sites are identified as points of maximal tenderness to deep palpation, reproducing the patient’s pain complaint. This may or may not result in referred pain. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. A 1½-in. (4-cm), 21- to 25-gauge needle is inserted at the point of maximal tenderness and advanced to the area of the trigger point. After negative aspiration, the trigger point area is injected with 4 mL of local anesthetic (Fig. 67-29).

Comment
The referred pain pattern for the thoracic paraspinal muscles often involves the scapular and chest wall region, as well as the lower thoracic paraspinal muscles and abdomen region. The referred pain pattern for the lumbar paraspinal muscles often involves the buttock, iliac crest, and sacroiliac joint region. These muscles involve the erector spinae, semispinalis cervicis, longissimus capitis, longissimus cervicis, longissimus iliocostalis thoracis, iliocostalis lumborum, and semispinalis multifidus.

Complications
Significant complications are uncommon with paraspinal trigger point injections.

Gluteal
Indications
Gluteal muscle injection is a useful diagnostic and therapeutic procedure for myofascial pain.

Techniques
After informed consent is obtained, the patient is placed in the lateral position with the unaffected side down, or in the prone position. The gluteus maximus, minimus, and medius muscles are palpated. The injection sites are identified as points of maximal tenderness to deep palpation, reproducing the patient’s pain complaint. This may or may not result in referred pain. The patient is prepared...
Part VI | Management Methods

in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. A 1½-in. (4-cm), 21- to 25-gauge needle is inserted at the point of maximal tenderness and advanced to the area of the trigger point. After negative aspiration, the trigger point area is injected with 4 mL of local anesthetic (Fig. 67-30).

Comment
The referred pain pattern for the gluteus maximus usually involves the sacroiliac joint, hip, and buttock. The referred pain pattern for the gluteus medius often involves the iliac crest, sacroiliac joint, and buttock. The referred pain pattern for the gluteus minimus muscle usually involves the buttock and lateral aspect of the lower extremity. The patient should be fully familiar with the stretching program for the gluteal muscle and be instructed in a home program. Failure to include a home stretching program usually results in short-term relief.

Complications
Significant complications are uncommon with gluteal trigger point injections; however, the anatomy of the region, including the sciatic nerve, must be carefully considered with these injections. Intraneural injection may result in nerve damage. Severe pain on injection suggests the possibility of an intraneural injection, and the needle should be immediately repositioned. Temporary lower extremity weakness is possible from regional spread of the local anesthetic.

Piriformis
Indications
Piriformis muscle injection is a useful diagnostic and therapeutic procedure for myofascial pain.

Techniques
After informed consent is obtained, the patient is placed in the lateral Sims’ position. The piriformis muscle is palpated from the sacrum toward the hip. The injection sites are identified as points of maximal tenderness to deep palpation, reproducing the patient’s pain complaint. This may or may not result in referred pain. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. A 1½-in. (4-cm), 21- to 25-gauge needle is inserted at the point of maximal tenderness and advanced to the area of the trigger point. After negative aspiration, the trigger point area is injected with 4 mL of local anesthetic (Fig. 67-31).
CHAPTER 67 | INJECTION PROCEDURES

Comments
The referred pain pattern for the piriformis muscle often involves the buttocks, iliosacral region, and posterior hip. The patient should be familiar with the stretching program for the piriformis muscle and be instructed in a home program. Failure to include a home stretching program usually results in short-term relief.

Complications
Attention to the anatomy of the sciatic nerve in this region will prevent intraneural injection; otherwise, significant complications are uncommon with trigger point injections. Severe pain on injection suggests the possibility of an intraneural injection, and the needle should be repositioned immediately. Temporary lower extremity weakness is possible from regional spread of the local anesthetic.

Hip Adductor
Indications
Hip adductor muscle injection is a useful diagnostic and therapeutic procedure for myofascial pain.

Techniques
After informed consent is obtained, the patient is placed in the supine position and the affected limb flexed, adducted, and externally rotated. The adductor longus, adductor brevis, and adductor magnus are palpated along the medial aspect of the humerus and thigh. The injection sites are identified as points of maximal tenderness to deep palpation, reproducing the patient’s pain complaint. This may or may not result in referred pain. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. A 1½-in. (4-cm), 21- to 25-gauge needle is inserted at the point of maximal tenderness and advanced to the area of the trigger point. After negative aspiration, the trigger point area is injected with 4 mL of local anesthetic (Fig. 67-32).

Comment
The referred pain pattern for the adductor muscles of the hip often involves the proximal hip, medial thigh, anterior thigh, and knee. The patient should be fully familiar with the stretching program for the adductor muscle and be instructed in a home program. Failure to include a home stretching program usually results in short-term relief.

Complications
Significant complications are uncommon with hip adductor trigger point injections.

Gastrocnemius
Indications
Gastrocnemius injection is a useful diagnostic and therapeutic procedure for myofascial pain.

Techniques
After informed consent is obtained, the patient is placed in a supine position. The gastrocnemius and soleus muscles are palpated from the knee to the ankle. The injection sites are identified as points of maximal tenderness to deep palpation, reproducing the patient’s pain complaint. This may or may not result in referred pain. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. A 1½-in. (4-cm), 21- to 25-gauge needle is inserted at the point of maximal tenderness and advanced to the area of the trigger point. After negative aspiration, the trigger point area is injected with 4 mL of local anesthetic (Fig. 67-33).

Comment
The patient should be fully familiar with the stretching program for the gastrocnemius muscle and be instructed in a home program. Failure to include a home stretching program usually results in short-term relief. The referred pain pattern for the gastrocnemius and soleus muscles often involves the posterior knee, calf, heel, and plantar aspect of the foot.

Complications
Significant complications are uncommon with gastrocnemius trigger point injections.
COMMON JOINT INJECTION TECHNIQUES

General
The joint is usually injected from the extensor surface at a point where the synovium is closest to the skin. This site minimizes the interference from major arteries, veins, and nerves. When the point of injection has been determined, it is best marked with the tip of a retracted ballpoint pen or a needle hub by pressing the skin to produce a temporary indentation to mark the point of entry. The skin is then prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. The skin and subcutaneous tissue at the injection site may be anesthetized by injecting 1% lidocaine, with no epinephrine, using a 25- to 30-gauge needle. Alternatively, 5% lidocaine-prilocaine cream may be applied to the skin surface for 15 to 30 minutes prior to skin preparation or a vapocoolant spray may be applied to the skin surface after skin preparation to provide adequate anesthesia (53,54).

A 1½- to 2-in. (4- to 5-cm), 22- to 25-gauge needle is then pushed gently into the joint. Before injecting the medication, an attempt should always be made to aspirate to avoid accidental intravascular injection. After ensuring that the needle is in the joint space, the medication should be injected in a slow, steady fashion.

Ultrasound guidance has been reported in multiple studies to increase accuracy for needle placement for arthrocentesis and therapeutic steroid injection in joints and tendon sheaths (55–58). One study demonstrated that needle guidance with ultrasound increased the ability to obtain synovial fluid from joints to 97% of patients compared with 32% when using conventional techniques without ultrasound (58).

Indications for Intraarticular Injection
Intraarticular injections may be used to determine the source of pain as articular or extra-articular and to provide maximal control of inflammation in joints when nonsteroidal anti-inflammatory drug (NSAID) therapy has failed or is contraindicated. Intraarticular injections are indicated to decrease morbidity in self-limited, sterile, inflammatory conditions. Intraarticular injections provide rapid relief of inflammatory pain and facilitate physical therapy of an inflamed joint. Poorly controlled inflammation in more than three joints requires reconsideration of systemic corticosteroids.

The intraarticular injection of hyaluronate for osteoarthritis is a relatively new technique. The viscosupplement acts like synovial fluid to maintain lubrication of the joint. This may be used in early osteoarthritis but does not appear to be an efficacious treatment in advanced osteoarthritis (59).

Contraindications for Intraarticular Injection
Contraindications must be considered before the injection of any joint. Contraindications to intraarticular injection include overlying soft-tissue sepsis, bacteremia, anatomic inaccessibility, an uncooperative patient, articular instability, septic arthritis, avascular necrosis, osteonecrosis, and neurotrophic joints.

Steroid injection into Charcot’s joint is contraindicated because local steroids will not provide significant long-term relief of the symptoms. Avascular necrosis in Charcot’s joints has been correlated to corticosteroid injections. Another specific contraindication is injection of an unstable joint unless the instability is appropriately corrected by bracing or surgery. Traumatic arthritis, secondary to fracture through the joint, is another contraindication for steroid injection because the beneficial effects of the steroid injections are not long lasting. Severe osteoporosis in areas around the joint is also a contraindication for injecting steroids.

Injection of joints with surgical implants is relatively contraindicated because these joints are more prone to infection than intact joints and are usually inflamed secondary to infection rather than synovitis. Injection of corticosteroids into a nondiarthrodial joint is rarely of value because there is no synovial sac in which to decrease inflammation.

Corticosteroids
The amount and type of corticosteroid vary extensively in the musculoskeletal application. The type of corticosteroid chosen is often based on purpose, expected duration, availability,
and preference of the practitioner. Suggested guidelines for maximum amounts with intraarticular injections include (60):

- Large joints (e.g., hip)—80 mg triamcinolone acetonide or equivalent
- Medium joints (e.g., knee)—40 mg triamcinolone acetonide or equivalent
- Small joints (e.g., wrist)—20 mg triamcinolone acetonide or equivalent
- Very small joints (e.g., metacarpal phalangeal, zygapophyseal facet)—5 to 10 mg triamcinolone acetonide or equivalent

Suggested guidelines for maximum amounts with bursa injections include:

- Large bursa (e.g., subacromial)—20 mg triamcinolone acetonide or equivalent
- Medium bursa (e.g., olecranon)—10 mg triamcinolone acetonide or equivalent
- Small bursa (e.g., anserine, retrocalcaneal)—1 to 5 mg triamcinolone acetonide or equivalent

Complications

The few complications associated with corticosteroid injections include infection, postinjection inflammation, and tissue atrophy. The occurrence of joint infection is extremely rare with the use of appropriate sterile techniques. Hollander described an incidence of 0.005% in more than 400,000 consecutive injections (10). There is a paucity of research regarding the efficacy and duration of action of various corticosteroids in joint and soft-tissue injections. The use of less soluble corticosteroids (acetate suspensions, e.g., methylprednisolone diacetate, triamcinolone acetonide) results in the steroid remaining in the joint and an assumed longer action. This decreased solubility results in increased possibility of postinjection flare. The use of more soluble corticosteroids (phosphate solutions, e.g., prednisolone sodium phosphate) results in more rapid absorption and an assumed shorter action. This increased solubility results in increased possibility of soft-tissue atrophy (61).

Postinjection inflammation is often secondary to corticosteroid crystal–induced synovitis and may mimic iatrogenic infection. This normally lasts 4 to 12 hours and is treated with NSAIDs and local application of ice. If this persists beyond 24 hours, the patient should be reevaluated to rule out infection. Postinjection flares may last up to 48 hours. The incidence of postinjection inflammation has been estimated at 1% to 6%. Repeated intraligamentous injections may result in calcification and rupture of the ligaments. Penetration of the articular cartilage will result in damage. Traumatic injection is prevented by never injecting against resistance.

Multiple studies on the weight-bearing joints of humans and primates treated with frequent corticosteroid injections have produced conflicting results regarding the adverse effects of corticosteroids on articular cartilage (17,61). However, due to the continued controversy regarding negative effects of corticosteroid injections on cartilage, it is recommended that intraarticular injections should be performed with intervals of at least 3 months to minimize complications to the cartilage or supporting ligaments.

Tissue atrophy in the area of injection occurs when corticosteroid is placed outside the joint space or leaks from the joint space. If a portion of the injected intraarticular corticosteroid is absorbed into the systemic circulation, the result may be an elevation of blood sugar, hormonal suppression, and brief, generalized improvement in all inflamed joints.

Corticosteroid injections are not recommended immediately after an acute injury or immediately before an athletic event. A patient should have a period of joint immobilization, rest, and protection from further injury after injection (62).

Techniques

Before the injection, the appropriate landmarks are located and marked. The skin is scrubbed with antiseptic and allowed to dry for 2 to 3 minutes. The wearing of sterile gloves is required so that the bony landmarks in the sterile field may be palpated throughout the procedure. Standard sterile technique is required to minimize risk for a septic joint. It is preferable to use single-dose vials of the steroid preparation and local anesthetic agent because this further reduces the risk for infection. A 25- to 27-gauge needle is used with 1% lidocaine, and no epinephrine, to raise a small skin wheal for skin anesthesia. Routinely, a 1½-in. (4-cm), 21- to 25-gauge needle traverses the skin, joint capsule, and synovial lining, sliding smoothly into the joint cavity. One must avoid the periosteum of the bone as well as the articular cartilage during this procedure. Aspiration is done to ensure there is no intravascular penetration. The return of synovial fluid ensures the position of the needle in the joint space; however, often there is minimal to no aspirated fluid. If there is an effusion, the fluid is removed in a slow, steady fashion until all possible joint fluid is aspirated. If the fluid is yellow and clear, the likelihood of infection is minimal, and the corticosteroid may be injected. If the fluid appears turbid, it should be sent for synovial analysis, including culture and sensitivity testing for microorganisms. If infection is suspected, the steroid injection into the joint should be postponed until the culture and sensitivity reports are completed. In addition to examining the color of the fluid, the viscosity of the fluid may be determined by putting a couple of drops of the fluid between the gloved thumb and index finger and stretching it. Normal synovial fluid has a good viscosity and is usually able to stretch for 2 to 2.5 cm. However, if there is an inflammatory process occurring in the joint, the viscosity of the fluid will be significantly decreased with a hazy or cloudy presentation (Table 67-6). The synovial fluid analysis completed in the laboratory may include rheumatoid factor, albumin, complements, protein electrophoresis, glucose level, and cell count with differential. A high white-cell count may indicate an inflammatory process. The fluid for culture and sensitivity should be sent immediately to the laboratory because infections caused by fastidious gonococcal organisms do not survive long in the test tube. Corticosteroids should not be injected into a joint until infection, including that caused by mycobacteria or fungi, has been excluded.
Aspiration from the joint may be impeded by synovial tissue over the end of the needle, intraarticular debris, and excessive joint fluid viscosity. It also may be difficult to aspirate if the tip of the needle is against or imbedded in the articular cartilage or if the needle is not in the joint cavity. It is important to ensure that the entire needle is withdrawn intact, because there have been reports of the separation of the needle from the hub and of needle fracture in articular injections.

The ease with which medication can be injected into the joint provides an indication as to the appropriate placement of the needle. No resistance should be encountered during the injection. If this occurs, the needle should be repositioned and aspirated to ensure that blood vessels have been avoided, and then the medication should be slowly injected. After the medication has been injected, the needle should be cleared with a new syringe containing a small amount of lidocaine or saline. The needle is then withdrawn with pressure applied to minimize bleeding. Joint injections are used primarily to deliver corticosteroids and anesthetic agents to treat inflamed synovium, bursa, and tendon.

Long-acting steroid preparations may induce a crystal synovitis 24 hours after the injection that resolves spontaneously. The patient should be cautioned about possible short-term aggravation of symptoms in the affected joint. It is recommended to infiltrate the subcutaneous tissues and the joint capsule with 2 to 4 mL of 1% lidocaine without epinephrine. When long-acting anesthetic agents are injected into the joint, it is important to advise the patient to limit usage of that joint during the first 24 hours after injection to prevent injury.

### Table 67.6 Characteristics of Synovial Fluid

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Normal</th>
<th>Noninflammatory (E.g., Osteoarthritis, Traumatic Arthritis, Osteochondritis Dissecans, Aseptic Necrosis)</th>
<th>Group I Rheumatoid Arthritis (E.g., Seropositive and Seronegative Spondyloarthritides)</th>
<th>Group II Septic Arthritis (E.g., Bacterial Infection Tuberculosis)</th>
<th>Group III Crystal Synovitis (Gout and Pseudogout)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clarity</td>
<td>Transparent</td>
<td>Transparent</td>
<td>Transparent to opaque</td>
<td>Opaque, cloudy</td>
<td>Clear with flakes of fibrin</td>
</tr>
<tr>
<td>Color</td>
<td>Pale yellow</td>
<td>Yellow or straw</td>
<td>Yellow</td>
<td>Brown/green/yellow/gray</td>
<td>Yellow</td>
</tr>
<tr>
<td>Viscosity</td>
<td>High</td>
<td>High</td>
<td>Low</td>
<td>Very low (may be high with coagulase-positive Staphylococcus)</td>
<td>Low</td>
</tr>
<tr>
<td>WBC/mm³ Predominant cell</td>
<td>&lt;150 Mononuclear (&lt;25% PMN)</td>
<td>&lt;3,000 Mononuclear (&lt;25% PMN)</td>
<td>3–50,000 Neutrophil (&gt;70% PMN)</td>
<td>50–300,000 Neutrophil (70%–100% PMN)</td>
<td>3–50,000 Neutrophil (&gt;70% PMN)</td>
</tr>
<tr>
<td>Crystal Culture</td>
<td>No Negative</td>
<td>No Negative</td>
<td>No Negative</td>
<td>No Often positive</td>
<td>Yes</td>
</tr>
<tr>
<td>Culture</td>
<td>No Negative</td>
<td>No Negative</td>
<td>No Negative</td>
<td>No Often positive</td>
<td>Yes</td>
</tr>
</tbody>
</table>

### Injection of Specific Skeletal Structures

No concensus exists regarding amounts, types, and mixtures of medications used. Developed familiarity, personal preference, cost, availability, and other factors often influence the use of these agents. Some practitioners promote the injection of a combination of short- and long-acting preparations. The suggested dosages in each of the specific examples in this section are common to many colleagues and other authors.

#### Cervical Zygapophyseal Joint

**Indications**

The cervical zygapophyseal joints have been shown to be a potential source of pain from the cranium to the midthoracic spine (63). Cervical facet injections can provide diagnostic as well as therapeutic benefits for patients with a wide variety of head and neck pains.

**Techniques**

After informed consent is obtained, the patient is placed in the prone position with the neck flexed and head turned to the opposite side to open the facet joint. A cushion is placed under the chest to allow neck flexion. The skin entry site lies about two vertebral segments below the target joint. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. A 3-in. (8-cm), 22-gauge spinal needle is advanced superiorly to the inferior margin of the joint, it is important to advise the patient to limit usage of that joint during the first 24 hours after injection to prevent injury.
to ensure the needle is directed over the articular pillars and not allowed to stray medially toward the interlaminar space or excessively laterally. The needle is advanced until contact is made with the articular pillar, either above or below the targeted joint, and then redirected into the joint capsule. Lateral and anteroposterior fluoroscopic views are necessary to ensure that the needle is advanced to the joint midpoint. Injection of contrast media can be used to confirm proper placement in the joint. After negative aspiration, the joint is injected with a 1-mL or less mixture of 6 mg mixture of betamethasone sodium phosphate and betamethasone acetate and local anesthetic (Fig. 67-34). Excessive injectate volume may spread to the epidural space via communication with the joint.

**Comments**
Total injected volume should not exceed 1 mL, because the joint volume is usually 1 mL or less. Anterior needle placement should be avoided because the neural foramen, epidural space, and vertebral artery are in close proximity to the anterior surface of the cervical joint.

**Complications**
Serious complications from cervical zygapophyseal joint injections are uncommon when meticulous care is given to ensure proper needle placement before injection. Local postinjection pain and light-headedness may occur. Ataxia and dizziness may occur if local anesthetics are used at the more proximal segments, secondary to loss of postural tonic-neck reflexes and proprioceptive input to the cervical muscles. The type of local anesthetic used will determine the duration of these effects. Vertebral artery injection of even small amounts of local anesthetic can cause a seizure; stroke is also possible if particulate steroids are used. Epidural or spinal blockade can occur if the needle placement is medial, resulting in regional blockade, respiratory compromise, and hypotension.

---

**Costochondral Junction**

**Indications**
Costochondral junction injection can be very useful as a diagnostic or therapeutic procedure in patients with costochondritis and Tietze’s syndrome.

**Techniques**
After informed consent is obtained, the patient is positioned in the supine position. The involved costochondral joints are palpated for local tenderness and replication of pain complaint. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. A 1-in. (2.5-cm), 25-gauge needle is inserted at the point of maximum tenderness to the level of costochondral cartilage and withdrawn 1 mm. After negative aspiration, a 2-mL mixture of 40 mg of methylprednisolone acetate or equivalent and local anesthetic is injected into each involved costochondral junction.

**Comments**
It is not necessary to advance the needle into the costochondral joint. Infiltration of the superficial tissue over the interosseous groove of the joint at the point of maximal tenderness is usually adequate. Tenderness with costochondritis is often present over more than a single costochondral joint. Tenderness and swelling of a single costochondral joint is found with Tietze’s syndrome. The presence of chest wall pain does not exclude underlying heart or lung disease. Similar injection techniques are used for the costoclavicular junction (Fig. 67-35).

**Complications**
Serious complications are uncommon with appropriate needle placement. Pneumothorax is possible with inadvertent penetration of the thorax. Most pneumothoraces can be easily treated with administration of supplemental oxygen and close observation and, when necessary, needle aspiration of air. Only those pneumothoraces that result in significant dyspnea or those under tension require chest tube thoracotomy and vacuum drainage.

---

**Glenohumeral Joint**

**Indications**
Intraarticular injection of the glenohumeral joint can be used to treat rheumatoid arthritis, inflammatory arthropathy, or adhesive capsulitis.

**Techniques**
After informed consent is obtained, the patient is placed in the sitting position, with the shoulder internally rotated. The
The glenohumeral joint is palpated by placing the fingers between the coracoid process and humeral head. The joint space can be felt as a groove just lateral to the coracoid process. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. A 1½-in. (4-cm), 21- to 23-gauge needle is inserted 1 fingerbreadth inferiorly and laterally to the tip of the coracoid process. The needle is directed in the anteroposterior plane just lateral to the coracoid process and is advanced into the groove. The needle is very gently manipulated through the joint capsule into the synovial cavity. Aspiration is attempted until the needle has entered the synovial space. If there is an effusion of the joint, aspiration should be completed. After negative aspiration, or if the aspirated fluid is noninflammatory (clear and viscous), a 2- to 3-mL mixture of 20 mg triamcinolone acetonide (or equivalent) and local anesthetic should be administered (Fig. 67-36).

**Comments**
Care is required not to direct the needle medially into the neurovascular structures in the axilla.

**Complications**
Hematoma and intravascular injection are possible, owing to the close proximity of the axillary vessels. If an arterial puncture occurs, prolonged direct pressure is usually adequate to prevent the development of a hematoma. Corticosteroids should not be injected if there is any suspicion that the joint is infected. If the fluid appears infected, it should be sent for culture and sensitivity and the patient treated appropriately for the infection.

### Acromioclavicular Joint

**Indications**
Intraarticular injection of the acromioclavicular joint is used to treat an inflamed or painful joint, as well as pain secondary to acromioclavicular joint separation.

**Techniques**
After informed consent is obtained, the patient is placed in the sitting position. The acromioclavicular joint is palpated by placing the fingers at the tip of the distal clavicle and medial to the tip of the acromion. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. A 1-in. (2-cm), 25-gauge needle is inserted at the joint and advanced to the proximal margin of the joint surface. After negative aspiration, the periarticular area is injected with 2 mL of 1% lidocaine for diagnostic purposes. If the local anesthetic provides significant pain relief, the periarticular area should be injected with a 2-mL mixture of 10 mg of triamcinolone acetonide (or equivalent) and local anesthetic (Fig. 67-37A).
CHAPTER 67 | INJECTION PROCEDURES

Comments
It is not necessary to advance the needle into the acromioclavicular joint. Infiltration of the superficial tissue over the interosseous groove of the joint at the point of maximal tenderness is usually adequate.

Complications
Serious complications are uncommon with injection of the acromioclavicular joint.

Rotator Cuff Tendon and Subacromial Bursa
Indications
Corticosteroid injection procedures are used to diagnose and treat rotator cuff tendinitis or subacromial bursitis. These conditions are often due to nonspecific irritation of the subacromial bursa, lesions of the rotator cuff, calcific tendonitis, or rheumatoid arthritis.

Techniques
After informed consent is obtained, the patient is placed in a sitting position with the arm in the lap. The lateral aspect of the shoulder is palpated for the point of maximal tenderness, usually 1 to 2 cm inferiorly and 1 to 2 cm anteriorly to the angle of the acromion. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. A 1½-in. (4-cm), 21-gauge needle is inserted below the acromion at the point of maximal tenderness. The needle is gently manipulated under the acromion. Aspiration is attempted until the needle has entered the synovial space. The subacromial bursa is about 1 to 2 cm below the skin between the tip of the acromion process and the head of the humerus. If there is an effusion of the bursa, the aspiration should be completed. After negative aspiration, or if the aspirated fluid is noninflammatory (clear and viscous), a 5-mL mixture of 20 mg of triamcinolone acetonide (or equivalent) and local anesthetic is injected. Half of the mixture should be injected under the acromion in the bursa. The needle should be slightly withdrawn and redirected toward the anterior part of the rotator cuff, and the remainder of the mixture infiltrated (see Fig. 67-37B).

Comments
Shoulder x-rays may show the locations of calcific deposits. If noted, a 1½-in. (4-cm), 16- to 18-gauge needle is directed to this area and aspiration attempted. Three mL of the mixture is injected at this location. The needle is withdrawn slightly and redirected toward the anterior part of the rotator cuff, and the remainder of the mixture infiltrated. This type of injection is usually uncomfortable, and premedication with codeine or oxycodone should be considered.

Complications
Corticosteroids should not be injected if there is any suspicion that the bursa is infected. If the fluid appears infected, it should be sent for culture and sensitivity and the patient treated appropriately for the infection.

Bicipital Tendon
Indications
Peritendinous injection of the bicipital tendon is a useful diagnostic and therapeutic procedure for bicipital tenosynovitis.

Techniques
After informed consent is obtained, the patient is placed in the seated position with the arm externally rotated and lateral to the medial edge of the humeral head. The bicipital groove is located and the bicipital tendon palpated to determine the area of marked tenderness. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. A 1½-in. (4-cm), 21- to 23-gauge needle is inserted along the border of the bicipital tendon. A 6-mL mixture of 20 mg of triamcinolone acetonide (or equivalent) and anesthetic agent is injected 2 mL at a time at the point of maximal tenderness and 1 in. above and below this point along the border of the bicipital tendon sheath (Fig. 67-38).

Comments
There should be no significant resistance encountered when injecting the tenosynovium. Resistance suggests that the tip of the needle is within the body of the tendon. Steroid injection into the tendon should be avoided.
Injecting directly into the tendon rather than into the periten- dinous region may result in damage to the bicipital tendon.

Lateral Epicondyle of the Elbow

Indications

Lateral epicondyle injection is a useful diagnostic and therapeu tic procedure for lateral epicondylitis of the elbow (tennis elbow). The condition is usually secondary to occupational or sports-related trauma or recurrent trauma.

Techniques

After informed consent is obtained, the patient is placed in the sitting position, with the arm resting on the examination table, palm down, and the elbow flexed to 45 degrees. The elbow is palpated at the junction of the forearm extensor group at its attachment to the bone near the lateral epicondyle to determine the point of maximal tenderness. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. A 1½-in. (4-cm), 23-gauge needle is inserted at the point of maximal tenderness. After negative aspiration, a 5-mL mixture of 10 mg of triamcinolone acetonide acetate (or equivalent) and anesthetic agent is injected at the point of maximal tenderness (Fig. 67-39).

Comments

The point of maximal tenderness is usually just medial and distal to the lateral epicondyle over the common tendon of the forearm extensor group at its attachment to the bone.

Complications

Serious complications are uncommon with injection of the lateral epicondyle of the elbow.

Medial Epicondyle of the Elbow

Indications

Medial epicondyle injection is a useful diagnostic and therapeutic procedure for medial epicondylitis (golfer’s elbow or tortilla elbow).

Techniques

After informed consent is obtained, the patient is placed in the sitting position, with the arm resting on the examination table, palm up, and the elbow flexed to 45 degrees. The elbow is palpated at the junction of the forearm extensor group at its attachment to the bone at the lateral epicondyle to determine the point of maximal tenderness. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure.
A 1½-in. (4-cm), 23-gauge needle is inserted at the point of maximal tenderness. After negative aspiration, a 5-mL mixture of 10 mg of triamcinolone acetonide acetate (or equivalent) and anesthetic agent is injected at the point of maximal tenderness.

**Comments**
The point of maximal tenderness is usually just lateral and distal to the medial epicondyle over the common tendon of the forearm flexor group at its attachment to the bone.

**Complications**
Avoid injecting the ulnar nerve in the groove just behind the medial epicondyle.

**Olecranon Bursa of the Elbow**

**Indications**
Olecranon bursa injection is a useful diagnostic and therapeutic procedure for olecranon bursitis. This condition is usually secondary to trauma or rheumatoid arthritis.

**Techniques**
After informed consent is obtained, the patient is placed in the sitting position with the hand on the lap. The olecranon process of the ulna is palpated for swollen bursa. The point of maximum swelling is marked. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. A 1½-in. (4-cm), 21-gauge needle is inserted at the most prominent part of the olecranon bursa. Aspiration is attempted until the needle has entered the synovial space. If there is an effusion of the bursa, the aspiration is completed. After negative aspiration or if the aspirated fluid is noninflammatory (clear and viscous), the bursa is injected with a 3-mL mixture of 10 mg of triamcinolone acetonide (or equivalent) and local anesthetic (Fig. 67-40).

**Comments**
This procedure may require an 18-gauge needle to aspirate the bursa with highly viscous fluid.

**Complications**
Corticosteroids should not be injected if there is any suspicion that the bursa is infected. If the fluid appears infected, it should be sent for culture and sensitivity and the patient treated appropriately for the infection.

**Radiohumeral Joint (True Elbow Joint)**

**Indications**
Radiohumeral joint injection is used to diagnose and treat the painful and swollen elbow due to rheumatoid arthritis or non-specific inflammatory arthritides.

**Techniques**
After informed consent is obtained, the patient is placed in the sitting position with the elbow flexed to 90 degrees. The lateral epicondyle and posterior olecranon are palpated. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. A 1½-in. (4-cm), 21- to 23-gauge needle is inserted into the groove just above and lateral to the olecranon process, just below the lateral humeral epicondyle, and posterior to the head of the radius. The needle is gently manipulated into the joint. Aspiration is attempted until the needle has entered the synovial space. If there is an effusion of the joint, the aspiration is completed. After negative aspiration or if the aspirated fluid is noninflammatory (clear and viscous), the joint is injected with a 5-mL mixture of 10 mg of triamcinolone acetonide (or equivalent) and local anesthetic (Fig. 67-41).

**Comments**
The connective tissue surrounding the elbow joint should be evaluated as a possible source of pain before injection of the radiohumeral joint.

**Complications**
Corticosteroids should not be injected if there is any suspicion that the bursa is infected. If the fluid appears infected, it should be sent for culture and sensitivity and the patient treated appropriately for the infection.

**Carpal Tunnel**

**Indications**
Injection of the carpal tunnel is used to treat inflammation of the tissue of the tunnel resulting in median nerve entrapment.
Techniques
After informed consent is obtained, the patient is placed in the sitting position with the arm resting on the examination table. The wrist is positioned with the hand dorsiflexed over a towel. The injection site is on the volar wrist surface just proximal to the distal wrist crease between the palmaris longus and flexor carpi radialis tendons. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. A 1½-in. (4-cm), 23- to 25-gauge needle is directed distally at an angle of 60 degrees to the skin and gently manipulated through the flexor retinaculum ligament into the carpal tunnel. The tunnel is about 1 to 2 cm from the skin in this position. After negative aspiration, the carpal tunnel is injected with a 1-mL mixture of 10 mg of triamcinolone acetonide acetate (or equivalent) and local anesthetic (Fig. 67-42).

Comments
Anesthesia in the distribution of the median nerve verifies injection into the carpal tunnel. These paresthesias may last for 1 to 2 weeks.

Complications
The median nerve should not be injected. The patient will normally report a sharp, electrical sensation when the needle tip is against the median nerve, and excruciating pain if the needle
tip pierces the median nerve. If either of the above occurs, withdraw slightly and continue the procedure. The volume injected into the carpal tunnel should be kept to a minimum to reduce postinjection discomfort.

**Wrist Joint**

**Indications**

Wrist joint injection is a useful diagnostic and therapeutic procedure for inflammation due to rheumatoid arthritis and other inflammatory arthritides.

**Techniques**

After informed consent is obtained, the patient is placed in the sitting position with the arm resting on the examination table. The hand is placed palm down with the wrist positioned over a rolled towel. The wrist joint is approached from the dorsal aspect. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. A 1½-in. (4-cm), 23-gauge needle is inserted between the distal radius and ulna on the ulnar side of the extensor pollicis longus tendon. The needle is gently manipulated into the joint cavity to a depth of about 1 to 2 cm. Aspiration is attempted until the needle has entered the synovial space. If there is an effusion of the joint, aspiration is completed. After negative aspiration or if the aspirated fluid is noninflammatory (clear and viscous), the joint is injected with a 2- to 3-mL mixture of 10 mg of triamcinolone acetonide (or equivalent) and local anesthetic (Fig. 67-43).

**Comments**

Many of the synovial joints of the wrists are interconnected. No significant resistance should be encountered. If resistance is encountered, the needle may not be in the joint cavity. Scapholunate dislocation, carpal instability, avascular necrosis, or other etiology of chronic conditions should be considered before injection. Elastic bandage or splint immobilization for 24 hours after injection may decrease discomfort.

**Complications**

Intraneural injection may result in nerve damage. Hematoma and intravascular injection are possible due to the close proximity of the vessels. If an arterial puncture occurs, prolonged direct pressure is usually adequate to prevent the development of a hematoma. Corticosteroids should not be injected if there is any suspicion that the joint is infected. If the fluid appears infected, then it should be sent for culture and sensitivity and the patient treated appropriately for the infection.

**Abductor Tendon of the Thumb**

**Indications**

Abductor tendon of the thumb injection is a useful therapeutic procedure for tenosynovitis of extensor pollicis brevis and abductor hallucis longus (de Quervain's syndrome) usually associated with repetitive trauma disorder. This procedure involves injection of common tendon sheath of the long abductor and short extensor tendons of the thumb.

**Techniques**

After informed consent is obtained, the patient is placed in the sitting position with the arm resting on the examination table. The forearm is placed on the ulnar side midway between supination and pronation. The wrist is held in ulnar deviation over a rolled towel, stretching the tendons over the radial styloid. Tendons are palpated for the point of maximal tenderness. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. A 1½-in. (3-cm), 23-gauge needle is inserted in a proximal direction, parallel to the tendon at a tangential angle, aiming for the point of maximal tenderness. The needle is held in ulnar deviation over a rolled towel, stretching the tendons over the radial styloid. Tendons are palpated for the point of maximal tenderness. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. A 1½-in. (3-cm), 23-gauge needle is inserted in a proximal direction, parallel to the tendon at a tangential angle, aiming for the point of maximal tenderness. Once the needle is in the tenosynovium, 0.25 to 0.5 mL of local anesthetic can be injected without significant resistance. This results in a small sausage-shaped swelling along the length of the tendons. At this point, the syringe containing only local anesthetic is disconnected, and another syringe is used to inject the peritendinous area with a 2-mL mixture of 5 mg of triamcinolone acetonide (or equivalent) and local anesthetic (Fig. 67-44).
Comments
There should be no significant resistance encountered in the tenosynovium. There will be resistance encountered by the plunger of the syringe if the needle is in the tendon. Steroid injection into the tendon should be avoided.

Complications
Injecting directly into the tendon rather than into peritendinous region can result in damage to the abductor tendon of the thumb.

First Metacarpal Joint Indications
First metacarpal joint injection is used to treat pain and inflammation secondary to osteoarthritis.

Techniques
After informed consent is obtained, the patient is placed in a sitting position with the arm resting on the examination table. The forearm is placed on the ulnar side midway between supination and pronation, with the thumb adducted and held in flexion with the palm. The first metacarpal along the dorsal aspect to the groove at its proximal end is palpated. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. A 1-in. (2.5-cm) needle is inserted at the point of maximal tenderness. The needle is advanced into the joint space. Aspiration is attempted until the needle has entered the synovial space. After negative aspiration or if there is an effusion of the joint, the aspiration is completed. If negative aspiration or if the aspirated fluid is noninflammatory (clear and viscous), the joint is injected with a 1- to 3-mL mixture of triamcinolone acetonide (or equivalent) and local anesthetic (Fig. 67-45).

Comments
Piercing the radial artery and extensor pollicis tendon should be avoided.

Complications
Radial artery injury, extensor pollicis tendon injury, and increased pain for 1 to 3 days are uncommon, but may result from this injection. Hematoma and intravascular injection are possible due to the close proximity of the axillary vessels. If an arterial puncture occurs, prolonged direct pressure is usually adequate to prevent the development of a hematoma. Corticosteroids should not be injected if there is any suspicion that the joint is infected. If the fluid appears infected, it should be sent for culture and sensitivity and the patient treated appropriately for the infection.

Interphalangeal Joint Indications
Interphalangeal joint injection is used as a therapeutic procedure to treat inflammation of the metacarpal phalangeal and interphalangeal joints due to rheumatoid arthritis and other inflammatory arthritides.

Comments
Piercing the radial artery and extensor pollicis tendon should be avoided.
Techniques
After informed consent is obtained, the patient is placed in the sitting position with the arm resting on the examination table. The hand is placed with the joint extended for approach from the lateral or medial aspect, with slight traction applied to the finger. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. A 1-in. (2.5-cm), 25- to 27-gauge needle is inserted at the borders of the joint and advanced gently to the joint capsule. Pericapsular injection without attempting to enter the joint is appropriate. The pericapsular area is injected with a 1-mL mixture of 10 mg of triamcinolone acetonide (or equivalent) and local anesthetic (Fig. 67-46).

Comments
No effort should be made to aspirate fluid unless infection is suspected.

Complications
Serious complications are uncommon with appropriate needle placement.

Lumbar Zygaphophyseal Joint
Indications
The lumbar zygaphophyseal (facet) joints have been shown to be a potential source of pain involving the low back and buttocks. Facet joint injections can provide diagnostic as well as therapeutic benefits for patients with low back pain.

Techniques
After informed consent is obtained, the patient is placed in the prone position with a pillow under the pelvis to flatten the lumbar curve. The lumbar spine is palpated for the point of maximum tenderness. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. A 3-in. (8-cm), 22-gauge needle is inserted at the point of maximum tenderness. The needle is advanced to the joint, with care taken to ensure that the needle is directed over the articular pillars and not allowed to stray medially toward the interlaminar space or excessively laterally. The needle is advanced until contact is made with the articular pillar, either above or below the targeted joint, and then redirected into the joint capsule. Lateral and anteroposterior fluoroscopy views are necessary to ensure that the needle is advanced to the joint midpoint. Injection of contrast medium may be used to confirm proper placement in the joint. After negative aspiration, the joint is injected with a 1-mL or less mixture of 10 mg of triamcinolone acetonide (or equivalent) and local anesthetic (Fig. 67-47).

Comments
Total injected volume should not exceed 1 mL, because joint volume is usually 1 mL or less. Anterior needle placement should be avoided because the dural sleeve, spinal cord, and epidural space are in close proximity to the anterior surface of the cervical joint.

Complications
Serious complications from lumbar zygaphophyseal joint injections are uncommon when meticulous care is given to ensure proper needle placement before injection. Local postinjection pain may occur. Intravascular injection of local anesthetic may cause a seizure. Epidural or spinal blockade can occur if the needle placement is medial, resulting in regional blockade, respiratory compromise, and hypotension.

FIGURE 67-46. Interphalangeal joint injection. Approach for first metacarpal joint aspiration and injection. (Reprinted with permission from Steinbrocker O, Neustadt DH. Aspiration and Injection Therapy in Arthritis and Musculoskeletal Disorders. Hagerstown, MD: Harper & Row; 1972:64, with permission.)

FIGURE 67-47. Lumbar facet joint injection. Fluoroscopic approach for lumbar facet joint injection (arrows).
Sacroiliac Joint

Indications
Sacroiliac joint injection is used to treat inflammation of the sacroiliac joints secondary to trauma, rheumatoid arthritis, degenerative joint disease, or stress secondary to mechanical changes in posture or gait.

Techniques
After informed consent is obtained, the patient is placed in the prone position. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. A 6-in. (16-cm), 22-gauge needle is inserted. The needle is advanced under fluoroscopy to the joint. After negative aspiration, joint penetration is confirmed with 1 mL of diatrizoate meglumine injection USP 60% (Renografin-60). After needle location is confirmed, a 1.5 to 2.0 mL (64) mixture of 40 mg of methylprednisolone acetate (or equivalent) and local anesthetic is injected (Fig. 67-48).

Comments
The sacroiliac joint is difficult to aspirate, owing to the depth and bony structures. Fluoroscopic guidance with the use of contrast media is recommended. It is rare to aspirate fluid from this joint.

Complications
Serious complications are uncommon with appropriate needle placement.

Coccygeal Junction

Indication
Infiltration of the coccyx region can be useful as a therapeutic procedure in coccydynia after exclusion of infection or other significant pathology.

Techniques
After informed consent is obtained, the patient is positioned in the lateral Sims’ position with the left side down for right-handed clinicians. With the upper leg flexed, the buttocks are separated, allowing easy access to the sacrococcygeal junction. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. The area of tenderness is localized by palpating from the tip of the coccyx to the sacrococcygeal junction. The palpat ing hand is kept in position, and a 1½-in. (4-cm), 21-gauge needle is inserted at the point of maximal tenderness perpendicular to the skin. After negative aspiration, a 3-mL mixture of 20 mg of methylprednisolone acetate or equivalent and local anesthetic is injected into the tender area using a fan pattern (Fig. 67-49).

Comments
It is not necessary to advance the needle into the sacrococcygeal junction. Infiltration of the superficial tissue at the point of maximal tenderness is usually adequate.
Complications
Perianal numbness may be noted for 24 hours after injection. Serious complications are uncommon with appropriate needle placement.

Hip Joint
Indications
Hip joint intraarticular injection is used to treat inflammation of the hip secondary to rheumatoid arthritis or osteoarthritis.

Techniques
After informed consent is obtained, the patient is placed in the supine position with the leg straight and externally rotated. A point is marked at 2 cm below the anterosuperior spine of the ilium and 3 cm laterally to the palpated femoral pulse at the level of the superior edge of the greater trochanter. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. A 3.5-in. (9-cm), 21-gauge needle is inserted at the mark in the posterior medial direction at an angle 60 degrees to the skin. The needle is advanced through the tough capsular ligaments to the bone and slightly withdrawn. Under fluoroscopy, contrast medium is injected to confirm appropriate needle placement. Aspiration is attempted until the needle has entered the synovial space. If there is an effusion of the joint, the aspiration is completed. After negative aspiration or if the aspirated fluid is noninflammatory (clear and viscous), the joint is injected with a 2- to 4-mL mixture of 20 mg triamcinolone acetonide (or equivalent) and local anesthetic (Fig. 67-50).

Comments
The hip joint is difficult to aspirate or inject due to the depth and limited landmarks. Fluoroscopic guidance with the use of contrast media is recommended. It is rare to aspirate fluid from this joint.

Complications
Avascular necrosis of the hip has been reported as a result of repeated intraarticular injection of corticosteroids. Hematoma and intravascular injection are possible, owing to the close proximity of the femoral vessels. If an arterial puncture occurs, prolonged direct pressure is usually adequate to prevent the development of a hematoma. Corticosteroids should not be injected if there is any suspicion that the joint is infected. If the fluid appears infected, it should be sent for culture and sensitivity and the patient treated appropriately for the infection.

Trochanteric Bursa
Indications
Trochanteric bursa injection is used to diagnose and treat bursitis of the hip. This often presents as pain in the lateral thigh during ambulation. Pain may be elicited by placing the hip in external rotation and abduction.

Techniques
After obtaining informed consent, the patient is positioned lying on the side and facing the clinician, with the painful hip exposed. The hips and knees are flexed and the affected
hip adducted. The protuberance of greater trochanter on the lateral aspect of the thigh is palpated for the point of maximal tenderness. This is usually two fingerbreadths below the tip of the trochanter. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. A 3-in. (8-cm), 21-gauge needle is inserted perpendicular to the skin at the point of maximal tenderness. The needle is advanced with aspiration attempted until the needle has entered the synovial space. If there is an effusion of the joint, the aspiration is completed. If the aspirated fluid is non-inflammatory (clear and viscous), the bursa should be injected with a 3-mL mixture of 20 mg of triamcinolone acetonide (or equivalent) and local anesthetic. If unable to enter the synovial space, the needle is advanced to the bone and then withdrawn 2 mm. After negative aspiration, a 3 mL mixture of 20 mg of triamcinolone acetonide (or equivalent) and local anesthetic is injected (Fig. 67-51).

Comments
The clinician should consider other causes of pain if the problem persists after bursa injection and appropriate rehabilitation.

Complications
Corticosteroids should not be injected if there is any suspicion that the bursa is infected. If the fluid appears infected, it should be sent for culture and sensitivity and the patient treated appropriately for the infection.

Abductor Tendon of the Hip
Indications
Injection of the abductor tendon is a useful diagnostic and therapeutic procedure for tendonitis at the insertion of the gluteal musculature into the greater trochanter.

Techniques
After informed consent is obtained, the patient is positioned lying on the side facing the clinician with the painful hip exposed. The hips and knees are flexed and the affected hip adducted. The hip is palpated above the tip of the trochanter to determine the point of maximal tenderness. A 3.5-in. (9-cm), 21-gauge needle is inserted at the point of maximal tenderness and directed toward the tip of the greater trochanter, approximating the insertion of the gluteal fasciae. The needle is advanced vertically to a depth that would reach the hip abductor tendon. This depth would vary from one patient to another, which is estimated by palpation of the hip abductor tendon. In obese patients, a lumbar puncture needle would have to be used to reach the tendon area. After negative aspiration, the area of maximal tenderness is injected with a 10-mL mixture of 10 mg of triamcinolone acetonide (or equivalent) and local anesthetic.

Comments
The injection is not placed into the tendon but rather into the peritendinous region. Wide infiltration with a corticosteroid and local anesthetic mixture is recommended. Tenderness in the vicinity of the hip joint are often associated with osteoarthritis of the hip. Injection of these sites may provide significant pain relief.

Complications
Injecting directly into the tendon rather than into the peritendinous region may result in damage to the abductor tendon of the hip.

Knee Joint
Indications
Intraarticular corticosteroid injection of the knee joint is used to treat noninfective inflammatory joint disease secondary to rheumatoid arthritis, seronegative spondyloarthritides, or the chondrocalcinosis inflammatory phase of osteoarthritis.

Techniques
After informed consent is obtained, the patient is placed in the sitting position with the knee flexed to 90 degrees. The patellar tendon is palpated and the middle of the patellar tendon is marked. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. A 1½-in. (4-cm), 21-gauge needle is inserted horizontally and advanced to the intercondylar notch. Aspiration is attempted until the needle has entered the synovial space. If there is an effusion of the joint, the aspiration is completed. After negative aspiration or if the aspirated fluid is noninflammatory (clear and viscous), the joint is injected with a 2-mL mixture of 10 mg of triamcinolone acetonide (or equivalent) and local anesthetic (Fig. 67-52).

Comments
An alternate approach may be used to access the suprapatellar pouch, which is continuous with the synovial space of the
knee. The patient is placed in the supine position with the leg fully extended. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. Throughout the procedure, the patella should be grasped between the examiner’s thumb and forefinger and should be able to be moved from one side to another to ensure that the quadriceps muscle is relaxed. A 1½-in. (4-cm), 21-gauge needle is inserted horizontally into the suprapatellar pouch at a point lateral and posterior to the patella at the level of the cephalad edge. A small amount of pressure is placed on the patella, pushing it to the side of needle insertion. This improves the ability to direct the needle during advancement (Fig. 67-53). The patient should be advised to minimize walking activity for 24 hours after injection, to minimize dispersion of the corticosteroid from the joint. If fluid is exceptionally viscous, a 1½-in., 18-gauge needle may be required to aspirate the joint.

Complications
Corticosteroids should not be injected if there is any suspicion that the joint is infected. If the fluid appears infected, then it should be sent for culture and sensitivity and the patient treated appropriately for the infection. It is contraindicated to inject the joint in a person with hemophilia, unless the risk for intraarticular bleed has been minimized. Corticosteroid injection into the knee joint may impair epiphyseal growth in children, resulting in a significant leg-length discrepancy.

Complications
Corticosteroids should not be injected if there is any suspicion that the joint is infected. If the fluid appears infected, then it should be sent for culture and sensitivity and the patient treated appropriately for the infection. It is contraindicated to inject the joint in a person with hemophilia, unless the risk for intraarticular bleed has been minimized. Corticosteroid injection into the knee joint may impair epiphyseal growth in children, resulting in a significant leg-length discrepancy.

Anserine Bursa of the Knee
Indications
Anserine bursa injection is a useful diagnostic and therapeutic procedure in bursitis resulting from osteoarthritis or direct trauma. Pain is noted inferior to the anterior medial surface of the knee when climbing stairs. Pain is reproduced with the knee in flexion-extension while internally rotating the leg.

Techniques
After informed consent is obtained, the patient is placed in the supine position with the knee in extension. The knee is palpated for the point of maximal tenderness over the medial tibial flare. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. A 1½-in. (4-cm), 21-gauge needle is inserted perpendicular to the skin at the point of maximal tenderness. The needle is advanced to the periosteum and withdrawn slightly. After negative aspiration, a 4-mL mixture of 2 mg of triamcinolone acetonide (or equivalent) and local anesthetic is injected (Fig. 67-54).

Comments
The anserine bursa is one of the most common bursae to become inflamed in the lower extremity. Kneepads are recommended for athletes with anserine bursitis secondary to trauma.

Complications
Serious complications are uncommon with injection of the anserine bursa.

Tibiotalar Joint
Indications
Tibiotalar joint injection is a useful therapeutic procedure with inflammation secondary to osteoarthritis, rheumatoid arthritis,
or chronic pain from instability. Pain most often occurs with ankle extension and flexion with weight bearing.

Techniques
After informed consent is obtained, the patient is placed in the supine position with the leg extended and the ankle extended over the end of the examination table. The area just anterior to the medial malleolus at the articulation of the tibia and the talus should be palpated and marked. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. A 1½-in. (4-cm), 21-gauge needle is inserted at the mark perpendicular to the skin. The needle is advanced slightly laterally, penetrating the capsule of the joint. The needle is directed into the tibiotalar joint to a depth of about 2 to 3 cm. Aspiration is attempted until the needle has entered the synovial space. If there is an effusion of the joint, the aspiration is completed. After negative aspiration or if the aspirated fluid is noninflammatory (clear and viscous), the joint is injected with a 2-mL mixture of 10 mg of triamcinolone acetonide (or equivalent) and local anesthetic (Fig. 67-55).

Comments
Injection of this joint is usually secondary to osteoarthritis resulting from trauma or from repetitive overuse injury such as from ballet dancing. If the swelling and tendonitis are around the lateral aspect of the joint, entry is accomplished just below the lateral malleolus. Gout is not an indication for injecting this joint.

Complications
Corticosteroids should not be injected if there is any suspicion that the joint is infected. If the fluid appears infected, then it should be sent for culture and sensitivity and the patient treated appropriately for the infection.

Subtalar (Talocalcaneal) Joint
Indications
Subtalar joint injection is used to treat inflammation secondary to rheumatoid arthritis and other inflammatory arthropathies.
Techniques
After informed consent is obtained, the patient is placed in the prone position with the feet extending over the end of the examination table and the foot flexed to about 90 degrees. The location of the subtalar joint, about 1 to 2 cm distal to the tip of the lateral malleolus and posterior to the sinus tarsus, should be palpated and marked. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. A 1½-in. (4-cm), 21-gauge needle is inserted perpendicular to the skin at the mark and advanced medially into the subtalar joint. Aspiration is attempted until the needle has entered the synovial space. If there is an effusion of the joint, the aspiration is completed. If the aspirated fluid is noninflammatory (clear and viscous), the joint should be injected with a 2-mL mixture of 10 mg of triamcinolone acetonide (or equivalent) and local anesthetic (Fig. 67-56).

Comments
Injection of this joint is usually secondary to osteoarthritis resulting from trauma or from repetitive overuse injury such as from ballet dancing. Gout is not an indication for injecting this joint.

Complications
Corticosteroids should not be injected if there is any suspicion that the joint is infected. If the fluid appears infected, then it should be sent for culture and sensitivity and the patient treated appropriately for the infection.

Retrocalcaneal Bursa
Indications
Retrocalcaneal bursa injection is a useful therapeutic procedure for bursitis secondary to repetitive overuse disorder or rheumatoid arthritis.

Techniques
After informed consent is obtained, the patient is situated in the side lying position. The lateral malleolus and the Achilles tendon are palpated. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. A 1½-in. (4-cm), 23- to 25-gauge needle is inserted between the lateral malleolus and the Achilles tendon perpendicular to the skin. The needle is advanced slowly to about half the thickness of the width of the Achilles tendon. After negative aspiration, a 2-mL mixture of 2 mg of triamcinolone acetonide (or equivalent) and local anesthetic is injected (Fig. 67-57).

Comments
This disorder may be seen in runners as they increase mileage early in the season or from an improperly fitting running shoe.

Complications
Corticosteroids should not be injected if there is any suspicion that the bursa is infected. If the fluid appears infected, then it should be sent for culture and sensitivity and the patient treated appropriately for the infection.
Plantar Heel Fascia

Indications
Plantar heel fascia injection is used to treat inflammation at the insertion of the long plantar ligament at the anterior aspect of the calcaneus, secondary to chronic overuse disorder or spondyloarthritides.

Techniques
After informed consent is obtained, the patient is placed in the prone position with the feet extending over the end of the examination table. The plantar aspect of the heel is palpated in the area of the attachment of the plantar fascia to the calcaneus to determine the point of maximal tenderness. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. A 1½-in. (4-cm), 23- to 25-gauge needle is inserted at the point of maximum tenderness on the plantar surface of the heel, perpendicular to the skin. The needle is gently advanced until the tip touches the underlying bone and then is withdrawn 2 mm. After negative aspiration, a 2-mL mixture of 20 to 10 mg of triamcinolone acetonide (or equivalent) and local anesthetic is injected. If proper palpation of the point of maximal tenderness is difficult, one half of the mixture of local anesthetic and corticosteroid should be injected into the region of maximal tenderness and the remainder injected in a fanwise manner around the plantar fascia attachment (Fig. 67-58A).

Comments
This is a significantly painful procedure with or without cutaneous anesthesia. After injection, the patient is discouraged from excessive walking until the local anesthetic wears off and is encouraged to wear a heel cushion inside the shoe.

Complications
Serious complications are uncommon with appropriate needle placement.

Metatarsophalangeal Joint

Indications
Metatarsophalangeal joint injection is a useful procedure in the treatment of joint inflammation secondary to rheumatoid arthritis.

Techniques
After obtaining informed consent, the patient is positioned for optimal access to the dorsal surface of the foot. The metatarsophalangeal joints are palpated for swelling and point tenderness. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. Light traction is applied to the toe of the joint to be injected. A ½- to 1-in. (1.5- to 2.5-cm), 25-gauge needle is inserted perpendicular to the skin, directly into the joint space. Aspiration is attempted until the needle has entered the synovial space. If there is an effusion of the joint, the aspiration is completed. After negative aspiration or the aspirated fluid is noninflammatory (clear and viscous), the joint is injected with a 0.5-mL mixture of 5 mg of triamcinolone acetate (or equivalent) and local anesthetic (see Fig. 67-58B).

Comments
These joints are often limited to 0.5 mL of solution. The first metatarsophalangeal joint may be approached from the medial side with the needle advanced tangentially under the extensor tendon.

Complications
Corticosteroids should not be injected if there is any suspicion that the joint is infected. If the fluid appears infected, then
it should be sent for culture and sensitivity and the patient treated appropriately for the infection.

**Metatarsal Joint**

**Indications**
This injection procedure is used to diagnose and treat Morton’s metatarsalgia and Morton’s neuroma.

**Techniques**
After informed consent is obtained, the patient is positioned for optimal access to the dorsal aspect of the foot. The metatarsal joint interspaces are palpated for swelling and tenderness. The patient is prepared in a standard aseptic fashion over an area large enough to allow palpation of landmarks, and sterile technique is used throughout the procedure. A 1-in. (2-cm), 25-gauge needle is inserted at the point of maximal tenderness, perpendicular to the skin, and advanced about 1 cm. After negative aspiration, a 2-mL mixture of 5 mg of triamcinolone acetonide (or equivalent) and local anesthetic is injected (Fig. 67-59).

**Comments**
Morton’s metatarsalgia often involves the first and second interdigital spaces. Morton’s neuroma is neuritis of the plantar digital nerves located between the third and fourth metatarsal heads and occasionally in the nerve between the second and third metatarsal heads. This procedure is used to treat postoperative scar pain often present after surgical removal of a Morton’s neuroma.

**Complications**
Serious complications are uncommon with appropriate needle placement.

**CONCLUSION**
The injection procedures outlined in this chapter are appropriately used in conjunction with other aspects of rehabilitation to reduce pain and increase function. Sterile technique and knowledge of anatomic relationships are required.

Aspiration is performed before injecting and is repeated as necessary throughout administration of medication to prevent intravascular injection. For optimal results and safety with injection techniques, the practitioner needs the necessary skill, training, and education to perform the procedures. Haste and failure to observe the necessary precautions by the practitioner increase the risks for any procedure. Knowledge of the anatomy pertaining to each injection procedure is a requirement for safe and successful outcomes in even the simplest injection.

**REFERENCES**


**SUGGESTED READINGS**

Spinal Injection Procedures

Spinal pain, especially low back pain (5% incidence and 60% to 80% lifetime prevalence in the United States), is very common (1). Low back pain is the leading cause of disability in people younger than 45 years. Although spinal pain often improves and resolves, a significant proportion of patients have ongoing symptoms and pain recurrence. Low back pain is a costly disorder with an annual cost approaching $50 billion (1). Therefore, a comprehensive rehabilitation approach that improves outcomes for patients with spinal pain can have a significant positive medical and economic impact. Adequate pain control can minimize disability, maximize function, improve quality of life, and potentially improve long-term outcomes by preventing the development of chronic pain syndromes. Spinal injection procedures have become an integral part of comprehensive rehabilitative management for individuals with spinal pain. Judicious performance of these interventional procedures on carefully selected patients can provide optimal pain control, reduce disability, and improve functional outcome. This chapter is intended to discuss common spinal interventional procedures in an evidence-based manner and provide some instruction on performance of these procedures.

**Epidural Steroid Injection**

Epidural steroid injections (ESIs) for the management of lumbar radicular pain provide the advantage of delivering potent anti-inflammatory agents in a localized fashion to the area of affected nerve roots, thereby decreasing the systemic side effects often seen with orally administered steroids. Lumbar ESIs have been endorsed by the North American Spine Society (2) and the Agency for Health Care Policy and Research (3) as an integral part of nonsurgical management of radicular pain from lumbar spine disorders. Cervical and thoracic ESIs are also being used to treat cervical and thoracic radicular pain respectively, but have not, as of the time of this writing, received similar endorsement.

**Pathophysiology of Radicular Spinal Pain**

Radicular pain is the result of nerve root irritation or inflammation (4–13). It is often described as a sharp, lancinating, and cramping discomfort that may feel like it is “shooting” from the spine along a dermatomal distribution. Other clinical manifestations of nerve root inflammation may include dermatomal hypesthesia or hyperalgesia, weakness of muscle groups innervated by the involved nerve roots, diminished deep tendon reflexes, and positive neural tension signs such as the straight-leg-raising test or Spurling’s maneuver.

Animal research in rats has revealed severe local inflammation in the epidural space and nerve root after injection of autologous nuclear material into the epidural space (6). High levels of PLA2, an enzyme that regulates the initial inflammatory cascade, have been demonstrated in herniated disc material from surgical samples in humans. Leukotriene B4, thromboxane B2, and inflammatory products also have been discovered within herniated human discs after surgery (4). Animal models have demonstrated that injection of PLA2 into the epidural space induces local demyelination of nerve roots and subsequent ectopic nerve discharges, which is considered to be the primary pathophysiologic mechanism of radicular pain (5). Interleukin 1 and tumor necrosis factor-α (TNF-α) have been detected in the herniated nucleus pulposus and found to play an important role in the pathogenesis of radicular pain from a herniated nucleus pulposus (11,12,14,14a).

Radicular low back pain caused by spinal stenosis probably occurs through the impedance of normal nerve root vascular flow and subsequent development of nerve root malnutrition, nerve root edema, and nerve root dysfunction (13). Chronic nerve root compression can induce axonal ischemia, impede venous return, promote extravasation of plasma proteins, and cause local inflammation (7). Venous congestion and arterial compromise can induce radiculopathy.

In a rat study, the compression of dorsal root ganglion (DRG) resulted in the reduction in K+, a hyperpolarizing shift in TTX-S Na+ current activation, and an enhanced TTX-R Na+ current. These phenomena may all contribute to the enhanced neuronal excitability and thus to the pain and hyperalgesia associated with the compression of DRG. In a separate study, an inflammatory soup (IS) consisting of bradykinin, serotonin, prostaglandin E2, and histamine (each 10−6 M) was applied topically to the DRG that had undergone chronic compression. IS remarkably increased the discharge rates of somata of a chronically compressed dorsal root ganglion (CCD) in rat neurons and evoked discharges in more silent-CCD than control neurons. Inflammatory mediators, by increasing the excitability of DRG somata, may contribute to chronic compression-induced neuronal hyperexcitability and to hyperalgesia and tactile allodynia (15). That these findings of lumbar radicular pain may be associated with increased excitability of involved DRG neurons was confirmed by a similar animal study (16).
In summary, clinical practice and animal research suggest that radicular spinal pain is the result of nerve root inflammation in the epidural space that is provoked by leakage of disc material, compression of the nerve root, and/or irritation of DRG from spinal stenosis. In addition, the pain may also relate to the enhanced neuronal excitability of the DRGs associated with chronic compression that occurs in neuroforaminal stenosis due to spondylosis or nucleus pulposus herniation.

Because radicular pain appears to originate from inflammation within the epidural space and nerve root, analgesic effects of epidural corticosteroids most likely are related to their anti-inflammatory effect. Underlying corticosteroid anti-inflammatory mechanisms include inhibition of phospholipase A₂ and inflammation (14, 14a), inhibition of neural transmission in nociceptive C fibers (5), reduction of capillary permeability, and nonselective inhibition of TNF-α (12) and IL-1 (17).

### Rationale for Epidural Injection of Corticosteroids and Local Anesthetics

Given the fact that the efficacy of corticosteroids apparently depends on their anti-inflammatory mechanism, it stands to reason that epidural injection should have higher efficacy in reducing nociceptive-induced radicular pain if applied earlier in the inflammatory process. In contrast, at least some of the nerve root fibers undergo fibrosis and necrosis in chronic radicular pain states. This perhaps explains in part why corticosteroid injections are less beneficial in the chronic setting. In fact, one study demonstrated little efficacy of cervical ESIs in atraumatic (i.e., stretched or compressed nerve roots) neuropathic radicular pain (68). In addition to timing of administration, the actual route of administration is important in treating radicular pain. For example, one study found that corticosteroid administration through the intravenous approach only offered short-term (<3 days) benefit (18). However, an open trial without control patients showed that intravenous methylprednisolone resulted in improvement in all 11 patients at the median follow-up time of 3.8 months (19). At the time of this writing, there have been only a couple case reports and no peer-reviewed, randomized controlled studies on the benefit of oral steroids in the treatment of radiculopathy. Some case reports provide initial data (20, 21).

Local anesthetics are often administered along with corticosteroids during epidural injections. By blocking sodium channels, local anesthetics impair peripheral neurotransmission of pain impulses, normalize the hyperalgesic state of the nervous system, and prevent and/or reduce the neuronal plasticity in the central nervous system by reducing the peripheral nociceptive input. Perhaps this, in part, explains the well-recognized phenomenon in clinical practice of pain relief after injection of local anesthetic often outlasting the physiological action of the anesthetics. In chronic (>6 months in duration) radiculopathy, neuropathic pain likely plays a greater role than pain due to inflammation. Since local anesthetics act directly on axons rather than acting as anti-inflammatory agents like corticosteroids, one would anticipate that they have efficacy in chronic radiculopathy. In fact, one study demonstrated the important role of local anesthetics in reducing chronic cervical radicular pain and discussed their possible mechanism via the neuronal plasticity that has been proposed to play a role in chronic radiculopathy (22).

### Indications for ESIs

The primary indication for ESIs is radicular pain associated with a herniated nucleus pulposus or spinal stenosis. A variety of other indications have been reported with variable results (23, 25–29). These include radicular pain associated with lumbar spine compression fracture, facet or nerve root cysts, postlaminectomy back pain, cervical strain syndromes with associated myofascial pain, and posttherapeutic neuralgia (23, 25–29).

### Contraindications for ESIs

Contraindications for epidural corticosteroid injections include systemic infection, local infection at the site of planned injection, bleeding disorder or full anticoagulation, history of significant allergic reactions to the components of the solution for injection, severe central canal stenosis at the level of planned injection, and lumbar ESI in pregnant women (23, 25–29). Caution should be used when performing injections in patients with poorly controlled diabetes and in individuals who have a history of severe or uncontrolled hypertension or congestive heart failure (CHF), because of the potential for steroid-induced fluid retention.

### Techniques

Cervical, thoracic, and lumbar epidural injections may be performed through either interlaminar or transforaminal approaches, and lumbosacral injections may also be performed through the caudal route (23, 25–29).

### INTERLAMINAR EPIDURAL INJECTION

#### Technique

The patient is placed in a prone position, ideally with a pillow or abdominal roll under the abdomen to help open up the lumbar interlaminar space by reversing the lumbar lordosis. The skin is then prepped and draped in a sterile manner. The targeted interlaminar space is identified using an anteroposterior (AP) fluoroscopic view, the vertebral body endplates at the targeted level are “squared off” by adjusting the relative cephalad-caudal orientation of the fluoroscope, and the fluoroscope position is further adjusted so that the proposed needle entry site into the epidural space is centered with respect to the fluoroscopic view in order to reduce parallax error. After the local skin and underlying tissues are anesthetized with 1% lidocaine, a 17- or 20-gauge epidural needle (e.g., Tuohy or Crawford) of appropriate length, depending upon body habitus, is inserted at the injection site. The epidural needle then penetrates the skin, subcutaneous tissue,
paraspinal muscles (paramedian approach) or the interspinous ligament (midline approach), and ligamentum flavum, where increased resistance is usually felt. At this point, the needle stylet is removed and the epidural needle is connected, ideally via extension tubing, to a Luer-Lok low friction glass or plastic syringe filled with about 2 mL of preservative-free saline. (Although the syringe can alternatively be filled with air, this can theoretically lead to an air embolus with inadvertent intrathecal injection and is believed to cause a higher incidence of postepidural headaches.) As the operator’s one hand advances the needle slowly into the ligamentum flavum, the other hand exerts steady gentle pressure on the plunger of the syringe. Depending upon the experience of the injectionist and the patient’s body habitus, the entire procedure can either be done using an AP view, or additional lateral views can also be obtained to help judge the depth of penetration. Once the needle penetrates the ligamentum flavum, loss of resistance should be detected by the hand holding the Luer-Lok syringe because saline will be suddenly injected owing to the negative pressure within the epidural space. Aspiration is then performed to ensure no CSF or blood return. (If blood is present, the needle position should be readjusted until no blood return is found. If CSF return is present, the needle is either withdrawn and the procedure attempted at an adjacent level or a caudal or transforaminal approach considered for the epidural.) A small amount of contrast (usually in the range of up to several milliliters) is then injected to visualize an epidurogram pattern that can be described as a Christmas tree, a bunch of grapes, or a vacuolated pattern (Fig. 68-1). Two other contrast patterns are possible if there has been false loss of resistance (in which the needle has not yet penetrated into the epidural space) or accidental needle penetration through the subarachnoid membrane. In these cases, contrast pattern recognition is essential. For example, in situations of false loss of resistance, the injected contrast typically appears as a local accumulation of contrast, whereas a typical myelogram revealing a relatively tubular (column-shaped) contrast pattern is generated when there has been subarachnoid membrane penetration. In the latter situation, the needle should be withdrawn, and the injection can be reattempted at an adjacent interlaminar space or by switching to a caudal or transforaminal approach. Once the needle is confirmed in the epidural space and no vascular pattern is observed upon contrast injection, a mixture of 4 to 10 mL of solution containing 80 to 125 mg of preservative-free methylprednisolone or 12 mg of preservative-free betamethasone sodium phosphate (Celestone Soluspan) and preservative-free 1% lidocaine with or without saline is injected into the epidural space through the epidural needle.

Several procedural modifications are recommended for cervical or thoracic interlaminar epidural injections due to the presence of the underlying spinal cord. For example, the cervical or thoracic interlaminar epidural injections should not be performed at the level of herniated nucleus pulposus or spinal stenosis, to avoid further potential spinal canal compromise and spinal cord compression. Furthermore, consideration should be given to directing the needle so that it contacts the inferior aspect of the lamina, to provide a clearly felt sense of depth prior to engaging the ligamentum flavum. The needle is then withdrawn slightly and directed into the ligamentum flavum. Further needle advancement should be performed using a lateral view and in addition to using the loss-of-resistance technique, the needle tip should not be advanced further than the laminar line to avoid the potential penetration of the dura mater or spinal cord injury. Epidural dye pattern recognition should be performed after a minimal amount of contrast has been injected since a total volume of less than 4 mL is recommended in these body regions (Figs. 68-2 and 68-3).
TRANSFORAMINAL EPIDURAL INJECTION

Comment
In one systematic review comparing transforaminal, interlaminar, and caudal ESIs, the authors concluded that there was moderate evidence for long-term (>6 weeks) relief of lumbar radicular pain using the transforaminal and caudal approaches, but limited evidence using the interlaminar approach (30). The authors also concluded that there was moderate evidence for relief of cervical nerve root pain using both the transforaminal and interlaminar approaches (30). Another study concluded that the transforaminal approach was more effective than the interlaminar or caudal approaches in treating lumbar pain (31). A review showed transforaminal lumbar ESIs under fluoroscopic guidance to be more cost effective than blind interlaminar and caudal ESIs (32).

As of this writing, no peer-reviewed, randomized controlled studies have compared transforaminal and interlaminar ESIs in the thoracic region.

Technique
For lumbar transforaminal epidural injection (TEI), the patient is placed in a prone position with a pillow or abdominal roll under the abdomen to at least reduce and ideally reverse the lumbar lordosis in order to open up the foramen. Using an ipsilateral oblique fluoroscopic view, the x-ray tube (source) of the C-arm fluoroscope is generally angulated in either a caudal direction (for L5-S1 and L4-5 TEI) or cephalic direction (for L3-4 and above TEI) to square the inferior end-plate of the vertebral body, and to place the superior articular process of the subjacent segment pointing at 6 o’clock of the pedicle of the above level that appears as a Scottie dog eye. Local skin is then prepped and draped in a sterile manner. A local skin wheel is raised with 1% lidocaine at the needle entry site and the subcutaneous tissue in the needle trajectory path is infiltrated with 1% lidocaine. A 22- or 25-gauge spinal needle of appropriate length is inserted and directed down and parallel to the fluoroscopic beam toward the “safe triangle.” The safe triangle is formed by the lower border of the pedicle, the lateral margin of the vertebral body, and the traversing nerve root. To avoid deep needle placement and potential injury to the vasculature or nerve root or DRG in the neuroforamen, the novice injectionist should advance the needle until the needle tip touches the lower edge of the Scottie dog eye, the junction of the transverse process and the superior articular process. The needle is then slightly withdrawn for 2 to 3 mm and redirected inferiorly just under the lower edge of the transverse process for about 0.5 mm. Further advancement of the needle should be under AP and cross table (lateral) views. The final needle tip position should be at the posterior half of the neuroforamen just under the pedicle in the lateral view to minimize the potential injury to the vasculature, nerve root, or DRG. In the AP view, the needle tip should not be medial to the medial edge of the pedicle to avoid penetrating the dura mater. For S1 transforaminal injections, the eye of the Scottie dog can also be used as an injection landmark. Using a slightly caudad and ipsilateral fluoroscopic view, the S1 Scottie dog image is outlined. The needle should be directed to the outer upper quadrant of the neuroforamen. In the lateral view, the needle tip should not pass the anterior margin of the sacral canal that appears as a radiological lucent strip. A neurogram pattern should then be visualized under an AP view (Fig. 68-4).

For the L5-S1 foramen, the C-arm source often needs to be tilted in a caudal direction to accommodate any remaining lumbar lordosis. An ipsilateral oblique projection is then used to visualize the Scottie dog and the target is identified as the
region immediately under the pedicle, slightly lateral to the 6 o’clock position (Fig. 68-5). This position leads to needle placement in the neuroforamen, ventral to the nerve root. Lateral imaging is used to demonstrate the needle depth, which should be located at the superior portion of the intervertebral foramen, just under the pedicle (Fig. 68-6). An AP view is then obtained to ensure that the needle tip is located at the “safe triangle,” slightly lateral to the 6 o’clock position of the pedicle. The safe triangle is formed by the lower border of the pedicle, the lateral margin of the vertebral body, and the traversing nerve root. A needle position located within the safe triangle and lateral to the 6 o’clock position is deemed safe because it will not penetrate the nerve, blood vessels, or dura mater. Nevertheless, because of the precarious location of the nerve root and the DRG, caution should be exercised by advancing the needle slowly upon entering the neuroforamen, to avoid needle penetration of these neurologic structures. If the patient complains of radicular pain or paresthesias, the needle should be withdrawn and redirected superiorly. Once the needle is deemed at the proper position, approximately 1.0 mL of the contrast is injected under live fluoroscopic view. The needle should be redirected if there is vascular uptake of the contrast. The contrast and corticosteroid are injected in a similar fashion as for lumbar TEI. A critical part of the injection is the correct identification of a clear rectangular-shaped clear space window under the fluoroscope. The upper and lower borders of the rectangle are the lower edge of the lamina of the same vertebral segment and the upper edge of the inferior vertebral endplate of the same segment. The lateral and medial borders of the rectangle are the medial edge of the rib head and the pars interarticularis of the same segment, respectively. After proper skin preparation and local anesthesia, the spinal needle is inserted and directed toward the lower edge of the Scottie dog eye using the same technique as in the lumbar TEI. Caution should be exercised not to direct the needle outside the clear rectangle window. If the needle strays too far laterally outside the rectangular window, it can penetrate the pleura, resulting in a pneumothorax. Needle placement too medially outside the rectangular window can result in spinal cord injury. The final needle position should be in the posterior half of the neuroforamen in the lateral view and the 6 o’clock position of the pedicle in an AP view.

A cervical TEI is performed, with the patient in a supine position. The fluoroscope should be directed in a similar region immediately under the pedicle, slightly lateral to the 6 o’clock position (Fig. 68-5). This position leads to needle placement in the neuroforamen, ventral to the nerve root. Lateral imaging is used to demonstrate the needle depth, which should be located at the superior portion of the intervertebral foramen, just under the pedicle (Fig. 68-6). An AP view is then obtained to ensure that the needle tip is located at the “safe triangle,” slightly lateral to the 6 o’clock position of the pedicle. The safe triangle is formed by the lower border of the pedicle, the lateral margin of the vertebral body, and the traversing nerve root. A needle position located within the safe triangle and lateral to the 6 o’clock position is deemed safe because it will not penetrate the nerve, blood vessels, or dura mater. Nevertheless, because of the precarious location of the nerve root and the DRG, caution should be exercised by advancing the needle slowly upon entering the neuroforamen, to avoid needle penetration of these neurologic structures. If the patient complains of radicular pain or paresthesias, the needle should be withdrawn and redirected superiorly. Once the needle is deemed at the proper position, approximately 1.0 mL of the contrast is injected under live fluoroscopic view. The needle should be redirected if there is vascular uptake of the contrast. The contrast and corticosteroid are injected in a similar fashion as for lumbar TEI. A critical part of the injection is the correct identification of a clear rectangular-shaped clear space window under the fluoroscope. The upper and lower borders of the rectangle are the lower edge of the lamina of the same vertebral segment and the upper edge of the inferior vertebral endplate of the same segment. The lateral and medial borders of the rectangle are the medial edge of the rib head and the pars interarticularis of the same segment, respectively. After proper skin preparation and local anesthesia, the spinal needle is inserted and directed toward the lower edge of the Scottie dog eye using the same technique as in the lumbar TEI. Caution should be exercised not to direct the needle outside the clear rectangle window. If the needle strays too far laterally outside the rectangular window, it can penetrate the pleura, resulting in a pneumothorax. Needle placement too medially outside the rectangular window can result in spinal cord injury. The final needle position should be in the posterior half of the neuroforamen in the lateral view and the 6 o’clock position of the pedicle in an AP view. The contrast and corticosteroid are injected in a similar fashion as when performing a lumbar TEI.

A cervical TEI is performed, with the patient in a supine position. The fluoroscope should be directed in a similar fashion as for lumbar TEI. A critical part of the injection is the correct identification of a clear rectangular-shaped clear space window under the fluoroscope. The upper and lower borders of the rectangle are the lower edge of the lamina of the same vertebral segment and the upper edge of the inferior vertebral endplate of the same segment. The lateral and medial borders of the rectangle are the medial edge of the rib head and the pars interarticularis of the same segment, respectively. After proper skin preparation and local anesthesia, the spinal needle is inserted and directed toward the lower edge of the Scottie dog eye using the same technique as in the lumbar TEI. Caution should be exercised not to direct the needle outside the clear rectangle window. If the needle strays too far laterally outside the rectangular window, it can penetrate the pleura, resulting in a pneumothorax. Needle placement too medially outside the rectangular window can result in spinal cord injury. The final needle position should be in the posterior half of the neuroforamen in the lateral view and the 6 o’clock position of the pedicle in an AP view. The contrast and corticosteroid are injected in a similar fashion as when performing a lumbar TEI.
posterior wall of the neuroforamen. At this point, the needle tip is withdrawn and directed slightly anteriorly to “walk off” the superior articular process and slip into the neuroforamen. The C-arm is turned to the AP view to assess the needle depth. The needle should be advanced in millimeter-by-millimeter increments in the AP view to ensure that the needle is not advanced past the center of the lateral mass (Fig. 68-8). Overzealous advancement of the needle into the inner half of the lateral mass can potentially lead to penetration of the dura into the subarachnoid space or into the spinal cord. The desired final needle location is the posterior wall of the targeted neuroforamen in the oblique view and the lateral half of the lateral mass in the AP view. After negative aspiration of the cerebrospinal fluid (CSF) or blood, 0.5 to 1 mL of contrast is injected under real-time imaging to exclude a vascular pattern. The needle should be repositioned if there is either blood flashback in the needle hub or a vascular pattern upon contrast injection. If the patient complains of paresthesias or radicular pain, the needle also needs to be repositioned. With satisfactory needle position, the injected nonionic water soluble contrast often outlines the exiting spinal nerve and fills the neuroforamen with epidural spreading or an epidurogram. After the satisfactory position, 1.0 mL of a test dose of 1% lidocaine is injected, and the patient is monitored for 2 minutes for any changes in vital signs or consciousness or neurological deficits in the extremities that would indicate an intravascular injection. For patients without abnormal signs, 40 mg of methylprednisolone, 6 mg of betamethasone sodium phosphate, or 10.25 mg of nonparticulate dexamethasone in a total volume of less than 2 mL per neuroforamen may then be injected (25,26,35). To prevent inadvertent arterial embolism into the spinal cord and brain stem, a nonparticulate soluble corticosteroid such as dexamethasone is recommended for cervical transforaminal ESI.

**CAUDAL LUMBAR EPIDURAL INJECTIONS**

Caudal lumbar epidural injections are performed by inserting a needle through the sacral hiatus into the epidural space at the sacral canal (Figs. 68-9 and 68-10). The patient is placed in a prone position. The legs are slightly abducted and feet turned inward to separate the gluteal fold to facilitate palpation of the sacral cornu. The sacral skin is prepped and draped in a sterile manner. AP imaging can be used to visualize the location of the

---

**FIGURE 68-7.** Oblique view of transforaminal cervical epidural injections showing needles in the posterior walls of the C4/5 and C5/6 neuroforamina.

**FIGURE 68-8.** AP view of transforaminal cervical epidural injection demonstrating left C6 neurogram.

**FIGURE 68-9.** Caudal epidural injection. Needle tip in the sacral canal. Note the “smoke up the chimney” pattern of contrast in the epidural space of the sacral canal.
sacral hiatus. Alternatively, lateral imaging is used to view the bone defect, consistent with the opening of the sacral hiatus. The skin and the tissues overlying the sacral hiatus are anesthetized with 1% lidocaine. A 22- or 25-gauge spinal needle of appropriate length or a Tuohy epidural needle is inserted into the sacral hiatus. Loss of resistance can sometimes be felt upon needle penetration through the sacral ligament. Several milliliters of the contrast are injected in order to produce a sacral epidurogram to note the level that the contrast reaches. In the lateral view, a typical epidural contrast spread within the sacral canal resembles “smoke up a chimney” (Fig. 68-9), and in the AP view, it often looks like a “Christmas Tree” (Fig. 68-11).

If a vascular pattern is observed, the needle should be withdrawn and redirected. Upon proper positioning, a mixture of 10 to 20 mL of solution containing 80 to 125 mg of preservative-free methylprednisolone or other equivalent doses of corticosteroid, preservative-free normal saline, and preservative-free 1% lidocaine is slowly injected into the epidural space through the spinal or epidural needle.

**TIMING, FREQUENCY, DOSE, AND VOLUME OF EPIDURAL INJECTIONS**

**Comment**

Optimal timing of ESIs is unknown, although there is evidence of better benefit if ESIs are performed within 3 months of radicular pain onset (36,37). The general consensus is that most patients with radicular symptoms should undergo a few weeks of treatment including oral medications, physical therapy or manual medicine, and relative rest from activities that exacerbate their pain, before undergoing ESIs (3,23,26). If a patient does not have success with such a program, or if the therapy cannot progress because the patient’s pain is too severe, an ESI is indicated for pain control. In contrast, ESIs can be considered earlier in patients with severe radicular pain not responding to even opioid medication or with pain that is significantly interfering with a patient’s sleep and/or function (26,28). Early ESIs also carry the theoretical benefit of controlling inflammation at an early stage (5,7,38) and possibly preventing permanent neural damage such as nerve fibrosis from prolonged inflammation (8). A study demonstrated that an ESI has higher efficacy (>75% pain relief) for patients with radicular pain within 3 months duration, whereas less benefit was found in patients with sciatica longer than 7 months (39). Another study compared epidural injections with bupivacaine alone versus injections of bupivacaine with methylprednisolone in patients with lumbar radicular pain longer than 6 months in duration. At 3-month follow-up, both treatments reduced pain but there was no additional benefit with corticosteroids (40).

The time interval between epidural injections should vary depending upon the steroid preparation used. Because injected methylprednisolone is reported to remain in situ for about 2 weeks (41), the clinician should probably consider waiting for about 2 weeks before fully assessing a patient’s response or administering a repeat injection.

Studies have suggested that the total maximum methylprednisolone dose should be about 3 mg/kg of body weight because excessive salt and water retention can occur at doses above this due to the mineralocorticoid properties of corticosteroids. In general, it is felt that up to three to four ESIs within a year may be performed if clinically indicated (23). Some clinicians schedule and proceed with a series of three ESIs regardless of the clinical response to the first preceding injection(s). Although the efficacy of this approach is unclear, as there are no medical outcome studies to support or refute such a regimen, it may be best to reassess the response to a

---

**FIGURE 68-10.** Caudal epidural injection. AP view demonstrating contrast in a “Christmas tree” pattern within the epidural space.

**FIGURE 68-11.** Right L5 TEI. Injecting 0.5 mL of contrast outlined the right L5 spinal nerve. However, contrast also spread through the epidural space and outlined the right S1 nerve root as well.
given injection at the time of an intervening office visit before proceeding with another injection. Using this approach, the clinician can determine if another injection is still needed and can more readily alter their planned injection technique, rather than trying to make this assessment at the time of the scheduled injection itself.

Recommended injection volumes and the corticosteroid doses are dictated mainly by the approach used as shown in Tables 68-1 and 68-2 (23,25,26,29). The epidural steroid can be injected in a preservative-free diluent such as lidocaine (1% to 2%) or normal saline (25,28,29). In the cervical spine, it is recommended that local anesthetics and steroids be injected separately to prevent a potential embolus of poorly dissolved steroid particles within a local anesthetic diluent. A recent study demonstrated that a nonparticulate dexamethasone has similar efficacy compared with particulate triamcinolone, and carries the lowest potential risk of embolization with inadvertent intravascular injection when used in ESIs (41).

### TABLE 68.1
<table>
<thead>
<tr>
<th>Volume of Injectate for Epidural Injections</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Lumbar</td>
</tr>
<tr>
<td>Cervical and thoracic</td>
</tr>
</tbody>
</table>

ESI, epidural steroid injection.
Data from reference 14, 16, 17, 20.

### TABLE 68.2
<table>
<thead>
<tr>
<th>Dosage of Corticosteroids for Epidural Injections</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Lumbar</td>
</tr>
<tr>
<td>Methylprednisolone</td>
</tr>
<tr>
<td>Triamcinolone</td>
</tr>
<tr>
<td>Diacetate (Aristocort)</td>
</tr>
<tr>
<td>Triamcinolone</td>
</tr>
<tr>
<td>Acetonide (Kenalog)</td>
</tr>
<tr>
<td>Betamethasone</td>
</tr>
<tr>
<td>Dexamethasone</td>
</tr>
<tr>
<td>Methylprednisolone</td>
</tr>
<tr>
<td>Triamcinolone</td>
</tr>
<tr>
<td>Diacetate (Aristocort)</td>
</tr>
<tr>
<td>Triamcinolone</td>
</tr>
<tr>
<td>Acetonide (Kenalog)</td>
</tr>
<tr>
<td>Betamethasone</td>
</tr>
<tr>
<td>Dexamethasone</td>
</tr>
</tbody>
</table>

ESI, epidural steroid injection.
Data from references 16, 17, and 45.

### FLUOROSCOPIC GUIDANCE AND CONTRAST

Fluoroscopic guidance and contrast enhancement are essential for accuracy when performing epidural injections (43). Published data show that even in experienced hands, epidural injections without fluoroscopic and contrast-enhanced guidance (i.e., “blind injections”) often result in inaccurate placement (Table 68-3) (43). These misplacements include the needle being inadvertently positioned into the subarachnoid, intravascular (Table 68-4), or subcutaneous regions (caudal approach) or fascial plane superficial to the ligamentum flavum for interlaminar ESI. Misplacement into the subarachnoid or intravascular regions has major potential safety implications, particularly for those injections that include local anesthetics as part of the injectate. Use of detection of flash back of blood in the needle hub to gauge the intravascular placement of needle is not a reliable substitute for looking for a vascular pattern after contrast injection (44). Although injection accuracy should also logically affect efficacy, there is very limited data on the efficacy of fluoroscopic-guided ESIs compared with blind ones. One such study demonstrated that fluoroscopic-guided transforaminal ESIs provided better pain relief than blind interlaminar ESIs (45). ESIs using fluoroscopic guidance have also been shown to reduce procedure-related complications compared to non–image-guided injections (46–48).

As a result of these factors, it is recommended that ESIs be performed under fluoroscopic guidance and with radiographic contrast, documenting appropriate needle placement in order to improve their accuracy, and by extension their safety and efficacy (49).
Efficacy of Epidural Injections

Recent studies have demonstrated good efficacy of lumbar ESIs when proper needle placement is confirmed by using fluoroscopic guidance and radiographic contrast (50,51). A meta-analysis of 12 published randomized controlled trials concluded that ESIs are effective (52). In a systematic review of randomized trials on lumbar epidural injections, Abdi et al. concluded that there was moderate evidence that caudal and TEIs are effective in providing long-term (>6 weeks) pain relief and limited evidence for the effectiveness of lumbar interlaminar ESIs (30). Other studies have suggested that 60% to 75% of patients receive some relief after ESIs (53,54). Benefits include relief of radicular pain and low back pain (generally relieving leg pain more than back pain), improved quality of life, reduction of analgesic consumption, improved maintenance of work status, and a decreased need for hospitalization and surgery in many patients (27,50–56). One study showed no difference in analgesic use in patients with sciatica who had received three ESIs (58). Another study reported that patients were more likely to start taking opioids and more likely to receive surgery after receiving multiple (>3) injections than patients receiving fewer injections (59). However, the population of patients receiving multiple steroid injections was more likely to have had more advanced disease such as spinal stenosis.

A prospective cohort study was conducted on cervical TEIs for both neck pain and radicular pain from herniated discs or spondylosis. Twenty-one such patients awaiting surgery received cervical TESIs 2 times, at 2-week interval with 12 months follow-up. All patients had reduction in neck and radicular pain, and five of these patients cancelled the surgery (59). In contrast, a prospective randomized study involving 20 patients with cervical radicular pain confirmed by selective nerve root block (SNRB) and with magnetic resonance imaging (MRI) evidence of corresponding segmental pathology demonstrated that there was no difference in radicular pain reduction between steroid/local anesthetic and saline/local anesthetic groups at 3-week follow-up (60). A limitation of this study, however, was that it only involved small numbers of patients and that it is unknown whether saline/local anesthetic is a true control.

There are more studies in support of ESIs for low back pain (7,8,42,53–55) than there are negative studies (56). Problems with some of these supportive studies, however, include the fact that most of these studies did not use fluoroscopy and radiographic contrast to document accurate placement of the injected substance into the epidural space. Furthermore, many of these injections were not performed at the presumed level of pathology, even though this has been demonstrated to be critical to the success of ESIs (61). These methodologic problems are likely contributing factors to the mixed assessment that ESIs have received. A review of six prospective randomized clinical trials of fluoroscopic-guided transforaminal ESIs, selective nerve blocks, or periradicular nerve injections concluded that there is moderate (level III) evidence that TESIs are safe and effective in reducing radicular pain. However, more prospective, randomized, placebo-controlled studies using sham procedures are needed to provide more conclusive evidence for the efficacy of TESIs in treating lumbar radicular symptoms (62). A recent review article concluded that with proper patient selection, ESIs are a reasonable alternative to surgery for short-term pain relief, reduced medication use, and increased patient activities while awaiting natural recovery (63).

Aside from technical considerations, response to ESIs has been shown to be related to several other factors such as the type and quantity of steroid preparation used, volume of injectate, underlying pathophysiology, and the duration of symptoms (23,26,28). In general, radicular pain or radiculopathy induced by herniated nucleus pulposus appears to respond better to corticosteroid injection than that induced by spinal

---

**TABLE 68.3** Incorrect Needle Placement Associated with “Blind” ESI

<table>
<thead>
<tr>
<th></th>
<th>Lumbar Interlaminar</th>
<th>Caudal</th>
<th>Lumbar Transforaminal</th>
<th>Cervical Transforaminal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Misplacement</td>
<td>17%–25%</td>
<td>12%–38%</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>


**TABLE 68.4** Incidence of Intravascular Uptake (IVU) Associated with ESI

<table>
<thead>
<tr>
<th></th>
<th>Lumbar Interlaminar</th>
<th>Caudal</th>
<th>Lumbar Transforaminal</th>
<th>Cervical Transforaminal</th>
</tr>
</thead>
<tbody>
<tr>
<td>IVU</td>
<td>1.9%</td>
<td>10.9%</td>
<td>11.2%</td>
<td>19.4%</td>
</tr>
</tbody>
</table>

steno sis. There is essentially no literature that correlates the type of disc herniation with the response of ESIs. It is the authors’ collective experience and observations that patients with large lumbar disc herniations obliterating the neuroforamen or extraforaminal herniations often have less benefit from ESIs. One study demonstrated that radiculopathy induced by the combination of spinal stenosis and disc herniation has less favorable outcome with ESI. In lumbar spinal stenosis, the efficacy of ESI correlated with the degrees and the levels of stenosis categorized by MRI (64). Patients with single-level lumbar spinal stenosis generally respond better than those with multilevel lumbar spinal stenosis. ESIs provide better efficacy in reducing pain and opioid consumption for patients with mild to moderate rather than severe stenosis. But a prospective cohort study with 12-month follow-up in patients with severe degenerative lumbar spinal stenosis found that fluoroscopic-guided and contrast-enhanced caudal ESIs reduced bilateral radicular pain and improved standing and walking tolerance (65). In contrast to radiculopathy due to herniated discs and/or spinal stenosis, radiculopathy caused by epidural scar tissues or trauma such as nerve root stretch injury often responds poorly to ESI.

A recent prospective, randomized study on lumbar TESIs demonstrated positive efficacy in treating radicular low back pain. The success rate for TESI is 84%, compared to 48% with trigger point injection, at 1.4 years of follow-up (66). Another prospective, randomized controlled clinical trial compared perineural (transforaminal) epidural injection with conventional posterior (interlaminar) epidural injection with steroid, and perivertebral injection with local anesthetic as a control group (27). The result demonstrated that perineural injection was the most effective approach. Both perineural and conventional epidural injection with steroid were better than that with saline alone (27).

Uncontrolled studies have generally reported favorable outcome of cervical epidural injections for cervical radiculopathy with structural abnormalities such as cervical disc herniation (66,67) and spondylosis (68). However, the prospective, randomized, blinded and controlled clinical trials on the outcome of cervical and thoracic epidural injections have not been reported yet in the peer-reviewed literature.

At the time of this writing, there have been no prospective randomized trials on thoracic ESIs that have been published in the peer-reviewed literature.

Patients should be educated that ESI alone may not be the only solution to give them long-term benefits. ESI is just one of many nonoperative treatments used to treat low back pain or radicular symptoms. Other treatments may include short-term bed rest; medications (e.g., analgesics, muscle relaxants); a properly designed program of physical therapy; and management of any psychological, financial, marital, and work-related problems. A comprehensive treatment approach is likely to produce better outcomes for patients with low back pain than any single modality used in isolation (23,26,51). Recently published research on the outcome of ESIs has supported this notion of multifaceted treatment (50,69).

Recent Advances and Investigations on the Management of Radicular Pain
A study was performed on lumbar TEsIs for radiculopathy using autologous conditioned serum (ACS) containing enriched IL-1 antagonist. The ACS group showed statistical superiority over both triamcinolone groups (5 and 10 mg) with regard to the VAS score for pain from week 12 to the final evaluation at week 22, statistical superiority at week 22 compared to the triamcinolone 5 mg group, and no significant difference compared to the 10 mg triamcinolone group. This is an exciting finding; as autologous blood is not associated with the same side effect concerns associated with corticosteroids, and theoretically can be used more frequently than corticosteroids. Additional studies are needed to confirm this finding (70).

An animal study examined the potential benefits from anti-TNF-α [α] therapy in reducing neurotoxic effects induced by the nucleus pulposus on neuronal tissues (71). Two open-label human clinical trials, one using intravenous infliximab (a monoclonal antibody against TNF-α) and the other using (etanercept) a soluble TNF-α receptor antagonist, in patients with sciatica from disc herniation demonstrated significant efficacy in pain reduction (72,73). Although these basic science and human studies initially implied potential clinical use of anti-TNF-α medication as a treatment for patients with radiculopathy due to disc herniation, there were disappointing long-term findings related to the evaluation of the efficacy of an anti-TNF-α treatment versus a placebo injection in disc herniation-induced sciatica in a randomized controlled setting. Specifically, 3-month results showed no difference in the patient-reported symptoms or in the more objective outcomes (SLR, days on sick leave, discectomies) between intravenous infliximab 5 mg/kg and placebo (74). The 1-year results also confirmed the earlier findings (75). Clearly, further studies using multiple intravenous infusion of the anti-TNF-α agents or epidural injection of the similar substances are necessary to clarify any efficacy of anti-TNF-α treatment in radicular pain.

Safety and Complications of Epidural Injections
A retrospective cohort study reviewing the immediate complications of 2,217 patients who received selective lumbar nerve root blocks under fluoroscopy, reported a 5.5% minor complication rate (76). When performed by a skilled, experienced clinician within an appropriate setting and on carefully selected patients, the chance of a significant complication from an ESI is remote (23,25,26,77–83). Like any procedure that punctures the skin, bleeding and soft tissue infection are potential but rare risks. More common risks of epidural injection are acute back pain, postural puncture headache (0.5% to 1% for lumbar interlaminar and 0.6% for caudal epidural injections), nausea, vomiting, dizziness, vasovagal reactions, and epidural hematoma (0.001%) (23,25,26,29). Nerve root injury, arachnoiditis, and meningitis also have been reported but are very rare. Lumbar transforaminal and caudal epidural injections were associated with 9.6% and 15.6% of minor complications, respectively (82,83). Anterior spinal cord syndrome has been reported after lumbar transforaminal ESIs, presumably due to
SELECTIVE NERVE ROOT BLOCK

Diagnostic Nerve Root Block

Because of the overlap pattern of dermatomal innervation and the anatomic variants of spinal nerves, clinical history and physical examination alone are often not sufficient to accurately diagnose the segmental level of a spinal nerve lesion. In addition, current imaging studies and electrodiagnostic tests have limited sensitivity and specificity in reaching a conclusive diagnosis of radicular pain at a specific spinal level. Therefore, a diagnostic SNRB can be an important test with respect to providing a physiologic diagnosis of the level of radicular pain. By selectively depositing a limited volume of local anesthetic directly around the spinal nerve rather than in the epidural space, pain relief after the SNRB identifies the spinal nerve blocked as the involved level. A diagnostic SNRB is indicated when imaging and/or electrodiagnostic testing studies are not corroborative with the clinical findings, or these tests demonstrate multilevel pathology and the exact pain generators are unclear. Studies reveal that a diagnostic SNRB is 87% to 100% accurate when intraoperative findings are used as the gold standard (77–79). Surgery performed at the positive SNRB level had higher success rate than surgery done at a level with a negative SNRB in the lumbar spine (80). In the cervical spine, SNRB also helped guide with a high level of success the evaluation of radicular pain in the multilevel degenerative cervical spine and subsequent surgery (91).

A diagnostic SNRB is performed with the needle tip directed to the posterior lateral portion of the neuroforamen, as for a transforaminal ESI. However, for a selective spinal nerve block, the needle tip should remain outside the neuroforamen pointing at the 5 o’clock position of the pedicle above for a left-sided SNRB or at 7 o’clock for a right-sided SNRB. The needle tip should also be localized immediately lateral to the superior articular process. Special care should, therefore, be exercised not to impale or transfix the exiting spinal nerve and it is contraindicated to inject steroid or local anesthetics directly into a spinal nerve due to their neurotoxic effects. The patients should be examined before the injection to document the maneuvers and activities that produce radicular pain. The same provocative maneuvers or activities should be repeated for comparison after the diagnostic injection. Selective spinal nerve block as a diagnostic procedure is considered positive when the patient’s radicular symptoms are reproduced upon gentle needle contact with the nerve sheath, followed by relief of the radicular pain after diagnostic blockade with local anesthetic.

Despite the apparent advantages of diagnostic selective spinal nerve block, the test has several limitations. Because of the overlap of the dermatomal distribution, blockade of one segment will not necessary produce clear-cut sensory changes. Second, because the anesthetic blockade is placed at the spinal nerve, a successful block can impede pain transmission not only from the spinal nerve but also from sites distal to the spinal nerve. Furthermore, the specificity of the blockade also depends on the spread of the injectate and the exact location of the needle tip. One study demonstrated that 1 mL of injected contrast in an L4 SNRB spread onto the L5 nerve root in 46.1%, and 1 mL of injected contrast in an L5 SNRB spread onto an S1 nerve root in 57.7% of subjects (92). If the needle tip is placed too close to the neuroforamen, even 0.5 mL of injectate can spread to the adjacent nerve root level through the neuroforamen, thus compromising the specificity of the segmental test (Fig. 68-11). However, the efficacy of blockade with anesthetic less than 0.5 mL is questionable. A study that utilized multi-slice computed tomography (CT) revealed that only 0.6 mL of injectate with contrast during cervical TEIs could be accepted as being selective enough for diagnostic investigations (93). It appears that a volume somewhere between 0.5 and 1 mL of local anesthetic should be used in performing an SNRB. More randomized, controlled studies are needed to determine the optimal volume of injectate for a diagnostic selective spinal nerve block and to ascertain the true value of a diagnostic SNRB in aiding with the selection of appropriate patients for spinal decompression surgery.

CONCLUSIONS

ESIs are effective treatment for radicular pain, at least in the short-term period, and particularly in the lumbosacral region. The use of fluoroscopic guidance and contrast enhancement can improve injection accuracy and thus presumably improve
safety and effectiveness. While ESIs are often reserved for patients who have failed to improve with other nonsurgical treatments, performing them earlier may further enhance their effectiveness. Further research is needed regarding the frequency of repeat injections.

**ZYGAPOPHYSEAL JOINT INJECTION AND RADIOFREQUENCY MEDIAL BRANCH NEURETOMY**

**Introduction**

The prevalence of lumbar zygapophyseal joint (Z-joint) pain has been reported to be approximately 6% in a primary care setting (94). However, in a tertiary spine center, lumbar Z-joint pain has been reported to range from 15% in younger individuals (96) to 40% in older populations (98) with chronic low back pain. In individuals with chronic neck pain after a whiplash injury, Z-joint pain occurred in approximately 50% of patients (99,100). More recent literature examining 500 patients has reported a prevalence of 55% cervical facet joint pain, 42% thoracic facet joint pain, and 31% lumbar Z-joint pain in patients with chronic spine pain identified with a double block model as described later (95). Lumbar facet joint pain has been reported in 16% of patients with chronic postsurgical lumbar spine pain (97). Studies of a larger population using a dual anesthetic block paradigm may be helpful in further identifying the prevalence in the general population for both acute and chronic low back pain.

**Anatomy**

Z-joints are pairs of small synovial joints in the posterior aspect of the spine, formed when the inferior articular process of one vertebra articulates with the superior articular process of the subjacent vertebra. Each lumbar Z-joint has a 1 to 2 mL capacity (101). Cervical and thoracic Z-joints can hold volumes of less than 1 mL (103) and 0.5 to 0.6 mL (104,105), respectively.

The Z-joint at C2-3 is innervated by the third occipital nerve from the superficial medial branch of the C3 dorsal ramus. Below the C2-3 Z-joint, each cervical Z-joint is innervated by the medial branches from the level above and below. The medial branches wrap around the articular pillar transverse processes. The C7 medial branch is located higher, as it is pushed up by the base of the transverse process. The joints between C0-1 (atlando-occipital) and C1-2 (atlandoaxial) joints are technically not Z-joints, due to their anterior location. They are innervated by the anterior rami of C1 and C2, respectively. Significant variability in the location of the medial branches has been reported, particularly with regard to the anatomy in the cervical spine (102).

The medial branches innervating the thoracic Z-joints have a different course in relation to the transverse processes (105). The thoracic medial branches instead wrap around the junction between the transverse process and the superior articular process. As in the lumbar spine, they often exit in the middle portion of the intertransverse space, and they typically cross the superolateral corners of the transverse processes and then pass medially and inferiorly across the posterior surfaces of the transverse processes. Furthermore, at mid-thoracic levels (T5-8), the inflection occurs at a point superior to the superolateral corner of the transverse process. Therefore, the superolateral corners of the transverse processes are generally the more accurate target points for diagnostic blockade or radiofrequency denervation of the thoracic medial branches.

In the lumbar spine, each Z-joint is innervated by two medial branches of dorsal rami; one from the same level and the other from the level above (106). For example, the Z-joint at L4-5 is innervated by the medial branches of the L4 and L3 dorsal rami. The medial branches travel along the junction of the transverse process and superior articular process. After exiting the mamilloaccessory notch covered by the mamilloaccessory ligament, the medial branches send branches to innervate the same level of the Z-joint and the Z-joint below. However, the L5 dorsal ramus crosses the groove between the sacral ala and the superior articular process of the sacrum. The L5 medial branch to the L5-S1 Z-joint is very short because it does not branch out until it comes just under the L5-S1 Z-joint.

It is important to remember that the lumbar medial branch rests on the subjacent level rather than the same level of the transverse process. For example, the L3 medial branch rests on the L4 transverse process rather than on the L3 transverse process (Figs. 68-12 and 68-13). Therefore, to denervate the L4-5 Z-joint, it is necessary to target the L3 and L4 medial branches, which rest on the junction of L4 and L5 transverse processes and superior articular processes, respectively. One exception is the blockade or denervation of the L5 medial branch. Due to the short branch below the Z-joint, denervation of the L5 medial branch can only be performed by destruction of the L5 dorsal ramus proper at the groove between the sacral ala and the superior articular process.

**Pathophysiology of Z-joint Pain**

The Z-joint is a well-innervated structure. The Z-joint capsule contains both nociceptive and mechanosensitive receptors (101,106-108). Immunocytochemical studies have demonstrated that the Z-joint capsule or synovial folds contain substance P, calcitonin gene-related product, vasoactive intestinal polypeptide (VIP) and neuropeptide, and tyrosine hydroxylase (101,107,108). Ostensibly, the Z-joint capsule is a potential pain generator if it is injured. In addition, like any synovial joints, Z-joints can develop synovitis under certain circumstances.

Typical Z-joint pathology often derives from pathologic mechanical stress or inflammation. In the lumbar spine, the sagittal orientation of upper lumbar Z-joints and the relative lateral oblique orientation of lower lumbar Z-joints make them vulnerable to injury from extension and torsion forces.

In the cervical spine, both human and animal studies of whiplash injuries have demonstrated evidence of multiple pathologic findings, including Z-joint capsule tearing, intra-articular hemorrhage, articular cartilage, muscle injury, and
subchondral bone fracture (110). These pathologic changes can serve as a basis of pathologic nociception and neck pain. However, none of these pathologies can be detected with conventional imaging studies (110).

Patients with lumbar Z-joint pain often have more pain in the lateral aspect of the low back unilaterally or bilaterally but not centrally (111). Pain can often be made worse with oblique extension of the lumbar spine (111,112). Pain is often worse after overnight rest or inactivity. Local tenderness and “muscle spasm” over the involved Z-joint are frequently noted. Although Z-joint pain can present with referred pain, groin pain or thigh pain, neural tension signs are negative and there are generally no neurological deficits, though recently a study where a rat model was used to induce Z-joint inflammation has documented associated radiculopathy as a possible sequela (109). Patients with cervical Z-joint pain experience pain located in the axial cervical areas. However, several studies have failed to demonstrate any pathognomonic physical exam findings of Z-joint pain (111,112). Imaging studies are also unable to confirm or refute the diagnosis of Z-joint pain (110,113,114).

**Indications**

Given the lack of a clinical diagnostic gold standard, clinicians have used regional anesthesia to identify Z-joint pain. Abolishment of low back pain after anesthetic Z-joint injection or medial branch block confirms the Z-joint as the pain generator (Figs. 68-14 and 68-15). Z-joint injection is indicated in patients with acute back and neck pain of suspected Z-joint origin, with no evidence of neurologic deficits, and whose pain pattern resembles that evoked in normal volunteers upon stimulation of their Z-joints. However, since the majority of acute back and neck pain, including Z-joint pain, will resolve in several weeks, the injection is often reserved for individuals

**FIGURE 68-12.** AP view of left L5-S1 Z-joint (facet) arthrogram. Notice the contrast pooling in the superior and inferior articular recesses. This 33-year-old man injured his low back lifting a 70-lb piece of metal overhead. Unable to sustain the weight, he extended his low back obliquely to the right side to drop the metal. He felt a “pop” in his left low back and sustained persistent significant left low back pain. Several minutes after this left L5-S1 intra-articular Z-joint injection with 0.5 mL of 1% lidocaine, his low back pain was completely abolished.

**FIGURE 68-13.** AP lumbar MBB at the right L2, L3 and L4 vertebral levels, demonstrating needle tips and contrast injected at the base of SAPs and the transverse processes.

**FIGURE 68-14.** Oblique view of left L4-5 Z-joint injection showing the needle tip inside the Z-joint space.
with severe pain that has failed to respond to 4 to 6 weeks of conservative therapy including oral analgesics, directed physical therapy, and relative rest. Fortunately, injection can be performed earlier if pain is inhibiting therapy progress.

Contraindications
As is true of all spinal injection procedures, Z-joint injections are contraindicated in individuals with the following conditions: infection, bleeding diathesis, pregnancy (for use of fluoroscopy), allergy to the medications to be injected (contrast medium, local anesthetics, corticosteroid), and unstable medical conditions such as unstable angina or poorly controlled hypertension or diabetes mellitus. Injections are elective procedures, so an effort should be made to properly select patients so as to ensure safety and optimize outcome.

Techniques
For lumbar Z-joint injection, the patient is placed in a prone position with a pillow under the abdomen to distract the Z-joint. After sterile preparation and skin draping, intermittent fluoroscopic views are used to identify the level of the Z-joint. For the L5-S1 level, the fluoroscope is tilted in a caudad direction to accommodate the lumbar lordosis, and rotated ipsilaterally until the joint space first comes into view, which is the posterior opening of the Z-joint, that is, the needle entry point. Although further rotating the image intensifier will more clearly visualize the Z-joint space, the visualized joint space at this angle is actually the middle or anterior Z-joint opening, not the posterior opening which is the injectionist’s target. For upper lumbar Z-joint injections, it may require less oblique rotation to better visualize the Z-joint space. The needle entry site is marked with a metal instrument. The skin and the underlying tissues are infiltrated with 1% lidocaine. For diagnostic Z-joint injections, the underlying tissues should not be infiltrated with anesthetic in order to maximize injection specificity. A 22- or 25-gauge 3.5 in. spinal needle is then inserted at the anesthetized site and directed toward either the superior or inferior articular processes of the targeted facet joint using a “tunnel view” technique by which the entire needle shaft is parallel to the fluoroscopy beam in such a way that the needle hub appears as a dot. Once the needle contacts the bone, the tip of the needle is then “walked” off into the Z-joint space. Occasionally, due to the osteoarthritic changes, the needle cannot gain entry. The needle can then be directed into the inferior articular recess just off the lower margin of the articular processes. Once the needle is felt to be in the articular space or to have penetrated the Z-joint capsule, 0.2 to 0.3 mL of the water soluble and nonionic contrast is injected to outline the Z-joint and to confirm that the needle tip is not located inside the vascular or epidural space. For therapeutic benefit, 1.0 mL of a mixed solution containing 20 mg of methylprednisolone acetate and 1% lidocaine is slowly injected into the Z-joint.

Medial Branch Block
For lumbar medial branch blocks, skin preparation and C-arm fluoroscope positioning are essentially the same as for Z-joint injections. The difference is the target, which is the “Scottie dog’s eye” rather than “ear,” because the former represents the anatomical site where the medial branch is situated whereas the latter represents the articular processes forming the Z-joint. After local skin is infiltrated with 1% lidocaine, a 22- or 25-gauge 3.5 in. needle is inserted and directed until the needle touches the middle portion of the “Scottie dog’s eye.” The fluoroscope is then turned to the cross table or lateral view; the needle should be located at the site posterior to the spinal lamina. At the anteroposterior (AP) view, the needle tip should be at or slightly medial to the lateral margin of the superior articular process. At this point, the needle bevel should be turned to face medially, and 0.2 to 0.3 mL of the contrast is injected under real-time fluoroscopy to ensure that a vascular pattern or neuroforaminal spread upon the dye injection has occurred. To ensure block specificity, less than 0.5 mL of either 2% lidocaine or 0.5% bupivacaine is used to block each medial branch.

For block of the L5 dorsal ramus, the fluoroscope should be rotated ipsilaterally oblique about 10 to 15 degrees. The needle is then directed in “tunnel view” down to the junction at the superior articular process of the S1 vertebra and the ala of the sacrum. In the AP view, the needle tip should be at the lateral margin of the S1 superior articular process.

Efficacy
Diagnostic Z-Joint Injection or Medial Branch Block
The literature has demonstrated that Z-joint injections and medial branch blocks can be used for the diagnosis of Z-joint pain with comparable sensitivity and specificity (Fig. 68-16)
Potential causes of false-negative responses include venous uptake of the injected anesthetic, aberrant innervation of the target Z-joint, and other technical issues. Since both intravascular injection and epidural spread of injectate can reduce the injection specificity, fluoroscopy plays an important role in maximizing outcomes.

If injection is being performed solely for diagnostic, not therapeutic, purposes, then only local anesthetic should be injected, without corticosteroid. Regarding the placebo-controlled blocks, a positive diagnosis was recorded only if the patient’s pain was completely and reproducibly relieved by each of the local anesthetic but not via the normal saline.

As stated above, double block with local anesthetics of different durations on two separate occasions can reduce (although not eliminate) the false-positive response and placebo effect (122). What constitutes a positive diagnostic response remains the physician’s subjective judgment. For a true concordant response with comparative blocks (100), the patient should obtain complete pain relief with a duration that is consistent with each particular anesthetic solution, though recent published placebo-controlled studies have used a criterion of at least 80% pain relief. This concordant response criterion for identifying Z-joint pain yields a good specificity of 88%, but only marginal sensitivity of 54%, thus suggesting significant false-negatives (100). Expanding the comparative blocks diagnostic criteria to include all patients with reproducible relief, irrespective of duration, increases sensitivity to 100% but lowers specificity to 65% (100). Whether the criterion should be structured to more aggressively prevent false-positive responses versus to prevent false-negative responses may depend on what subsequent treatment will be based on the determination. For additional treatments that can irreversibly alter the patient’s anatomy (e.g., via surgery or radiofrequency ablation [RFA]), preventing false-positive responses becomes increasingly important (100).

Using the dual block technique in patients with chronic low back pain, lumbar Z-joint pain was demonstrated in 15% of younger patients and 40% of older patients (96,98). In chronic neck pain after whiplash injury from motor vehicle injury, Z-joint pain occurred in 54% of patients (99). Another study of patients with chronic neck pain for more than 6 months demonstrated that the prevalence of Z-joint pain was 36% (123). The C2-3 and C5-6 Z-joints were found to be the most common symptomatic joints (99,123). The C2-3 joint was found to be a pain generator in 50% of patients with chronic cervicogenic headache after a whiplash injury (99). It is interesting that using the intra-articular and/or medial branch block, studies have demonstrated that the coexistence of Z-joint pain and discogenic pain in lumbar region is only 4%, while in the cervical spine it is 40% (124,125).

Caution should be exercised using provocation of pain as a sole diagnostic criterion for patient undergoing a diagnostic lumbar Z-joint injection. One study has demonstrated
no significant correlation between pain provocation during Z-joint injection and the analgesic response (126).

**Therapeutic Effect**
The therapeutic benefit of Z-joint injection with corticosteroids remains controversial (127). A past study has suggested positive efficacy of C2-3 Z-joint corticosteroid injections for cervicogenic headache after a whiplash injury (128). Controlled studies have failed to demonstrate such efficacy although the studies had various design flaws (Table 68-5), including the use of saline as a placebo, patient selection bias without using a double block technique, and injection in isolation of other treatments. Selection of patients without using the double block paradigm may potentially include those patients with false-positive results. Finally, injection is not recommended to be used in isolation of other treatments. Rather, when tolerable, patients should be involved in a directed therapy program if they experience significant pain reduction (therapeutic window) after the injection.

A recently published systematic review of the literature from 2004 to 2006 has demonstrated strong evidence for diagnostic injections in the cervical and lumbar spine and moderate evidence for diagnostic injections in the thoracic spine (115). Another published review from 2007 has noted “limited” evidence for intra-articular cervical facet joint injections and “moderate” evidence for intra-articular lumbar Z-joint injections for pain relief (129). It is interesting to note, however, that this study did conclude that there actually was moderate evidence for short-and long-term pain relief from medial branch blocks.

**Radiofrequency Neurotomy**
Radiofrequency neurotomy interrupts the nociceptive afferent from the Z-joint by thermally coagulating the two medial branches that innervate a given Z-joint. The exposed terminal portion of radiofrequency probe delivers heat at 80°C. For each Z-joint (except the C2-3 joint, which is innervated by the third occipital nerve), two medial branches need to be ablated.

**Indications**
Radiofrequency neurotomy can provide relative long-term benefit symptoms from persistent or recurrent Z-joint pain despite conservative care (that have had transient benefit from Z-joint injections), and for patients with substantial (e.g., at least 80%) pain relief after dual blocks of the medial branch with two local anesthetics of different duration on two occasions.

**Technique**

**Lumbar Medial Branch Neurotomy**
Since the “groove” at the junction of the superior articular process and the transverse process can be clearly viewed in a certain fluoroscopic position, this “groove view” has been proposed for the starting position of a lumbar medial branch neurotomy. Specifically, the C-arm intensifier is obliqued ipsilaterally approximately 10 to 15 degrees and tilted cephalad 20 degrees. A 22-gauge 100 mm radiofrequency probe with a 5-mm active tip is then introduced using the “tunnel view” until it contacts the bone and advanced along the “groove” where the medial branch resides. The radiofrequency probe creates an effective circumferential lesion around the probe but does so poorly distal to the tip. To avoid incomplete heating, the radiofrequency probe needs to be placed on and parallel to the path of the medial branches crossing the transverse processes. In AP imaging, the needle should be medial to the lateral silhouette of the superior articular process. A lateral view should be obtained to ensure that the

---

**TABLE 68.5 Summary of Controlled Studies of Z-Joint Injections with Corticosteroid**

<table>
<thead>
<tr>
<th>Patient Selection</th>
<th>Study Group</th>
<th>Results (Pain Relief)</th>
<th>Comments on Study Design</th>
</tr>
</thead>
<tbody>
<tr>
<td>One hundred and nine patients with unilateral or bilateral LBP &gt;3 mo (76)</td>
<td>Three groups with Z-joint injections: 1. Corticosteroid and local anesthetics 2. Saline 3. Pericapsular injection with corticosteroid and local anesthetics</td>
<td>64% at 1 h 36% at 3 mo No benefit attributed to corticosteroids</td>
<td>Large volume used (&gt;8 mL), potentially ruptured Z-joint capsule Did not use Z-joint injection for patient selection Failed to exclude placebo effect by using saline</td>
</tr>
<tr>
<td>Forty-one patients with neck pain from whiplash injury selected with two separate diagnostic MBB (86)</td>
<td>Z-joint block with betamethasone Z-joint injection with bupivacaine</td>
<td>No difference</td>
<td>Injection in isolation</td>
</tr>
</tbody>
</table>

LBP, lower back pain; MBB, median branch block.

needle is not positioned anterior to the posterior aspect of the neuroforamen. Motor stimulation using 2.0 Hz and less than 2.0 V should not induce any muscle twitching or movement in the lower extremity. A single lesion at 80 degrees celsius for 60 seconds is performed. A second and third lesion is performed after repositioning the needle 1.0 mL cephalad and caudad to ensure proper coagulation of the length of the medial branch.

**Cervical Medial Branch Neurotomy**

For cervical medial branch ablation, the patient is placed in a prone position and the C-arm is tilted in a cephalad direction to obtain a pillar view. The first radiofrequency lesioning probe is inserted parallel to the articular pillar, directed down the fluoroscopic beam and slightly medial until it touches the dorsal aspect of the pillar. The needle is then walked lateral until it slips off the bone. A second radiofrequency probe is inserted with approximately 30 degrees of ipsilateral obliquity and slightly caudad so that this needle can be positioned more anteriorly. The rest of the needle advancement is the same as for the first needle. This dual needle placement positions the two probes to allow subsequent adequate denervation along the entire length of the medial branch. Prior to lesioning, a lateral view is taken to ensure that the tips of the radiofrequency probes are not anterior to the anterior margin of the articular pillars. Motor stimulation is conducted at a frequency of 2.0 Hz and at less than 2.0 V intensity. If muscle twitching or upper extremity movement occurs, this indicates that the radiofrequency probe is too close to the anterior rami and needs to be repositioned. Next, 0.5 mL of 1% lidocaine is injected through the radiofrequency cannula. A lesion is then performed at this position at 80°C for 60 seconds. The probe is then repositioned 1 mm caudad and cephalad, respectively, and two additional lesions are made for a total of six lesions using the two radiofrequency probes. A similar technique is employed for the third occipital nerve.

**Efficacy**

Proper patient selection is the key for optimal outcome from radiofrequency neurotomy. One recent prospective study has demonstrated good efficacy from the procedures when patients with presumed Z-joint pain are selected using the double block paradigm with comparative local anesthetic (130). The patients were diagnosed with a lumbar Z-joint pain if they obtained at least 80% pain reduction after medial branch blocks with 0.5 mL of 2% lidocaine on one occasion and 0.5% bupivicaine on another. At 12 months following the radiofrequency medial branch neurotomy, 60% of patients achieved at least 90% pain reduction and 87% of patients had 60% pain relief. The success of denervation was seen in virtually all patients as demonstrated by post-radiofrequency needle EMG of the corresponding segmental multifidi (130). The initial population of 41 patients accepted into the study and clinically felt to have facet-mediated pain was reduced to only 15 candidates for actual RFA, which further emphasizes the importance of double-block screening for Z-joint pain diagnosis. A recent randomized study using single Z-joint blocks for selecting patients with subsequent radiofrequency neurotomy demonstrated modest success with an average VAS pain reduction of 2.0 points in 66.7% of the lesion group versus 35.7% of sham group patients at 8 weeks to 12 months follow-up (131). An obvious concern with this study is that the criteria for a positive block mandated a relatively low threshold of only 50% pain relief, and 40 of 92 patients subjected to the diagnostic block had a positive response in this regard. This would imply a prevalence of facet joint pain higher than what has been reported in most populations. Another study showed no treatment effect at 12 weeks, as assessed by the Roland-Morris scale (2.6% change) and Oswestry scale (1.9% change) and VAS (–7.6% change) (132). Again, the known 38% false-positive rate for a single-block makes the study results and conclusions less convincing (126,133).

Efficacy of radiofrequency neurotomy for cervical Z-joint pain (other than from the C2-3 joint that was excluded from this study), has also been demonstrated in a randomized, double-blinded, and placebo-controlled trial (134,135). The patients were selected by placebo-controlled medial branch blocks. The total duration of pain relief was defined as the period until the patient judged that pain had returned to 50% of the pre-procedural level. Twenty-four patients were randomized into radiofrequency neurotomy and sham radiofrequency neurotomy (radiofrequency probes placed but radiofrequency was not turned on) groups. Fifty percent pain relief lasted 263 days in the radiofrequency group and 8 days in a controlled group. A second study of 28 patients with long-term follow-up and repeat radiofrequency neurotomy demonstrated a median duration of pain relief of 422 days. In the 11 patients who underwent repeat radiofrequency neurotomy, the median duration of pain relief was 219 days. The authors concluded that radiofrequency neurotomy provides clinically significant pain relief, and can be repeated if pain recurs (136).

Recent review articles on randomized controlled trials of radiofrequency neurotomy for spinal pain concluded that RF neurotomy was efficacious for both chronic low back pain and neck pain after flexion-extension injuries (124), and that there is evidence of moderate strength for use of RFA in the cervical and lumbar spine (129).

**Complications of Z-Joint Procedures**

Potential complications of fluoroscopic-guided, contrast-enhanced lumbar Z-joint injections or medial branch blocks are rare. The most common post-procedural problem is transient pain at the injection site. However, there have been rare case reports of meningitis, inadvertent spinal anesthesia, and infection after Z-joint injections (137,138).

Recurrence of back and neck pain after radiofrequency neurotomy may be due to incomplete ablation or medial branch regeneration (139). Since the dorsal root ganglia are left intact, the ablated medial branches may regenerate.
radiofrequency neurotomy does not permanently denervate the Z-joints or cause inherent instability in the spine, the concern of Charcot facet joint development has little ground and no such cases have been published (140). Also, because the DRG remains intact, deafferentation pain should not occur. Side effects from radiofrequency neurotomy are rare when the procedure is performed correctly but include local pain and infection.

**CONCLUSIONS**

The Z-joints of the spine can be a significant source of pain. Diagnostic medial branch blockade is a valid technique for confirming the Z-joint as a pain generator. Proper selection of patients with the dual block paradigm may minimize false-positive responses and increase therapeutic efficacy for intra-articular injections. In carefully selected patients with recalcitrant symptoms, radiofrequency neurotomy of medial branches may offer long-term benefit.

**SACROILIAC JOINT INJECTION**

The sacroiliac (SI) joint can be a significant source of low back pain (141–143). Etiologies of SI pain include spondyloarthropathy, crystal arthropathy, septic arthritis, trauma, and pregnancy diasthesis (144). In addition, SI joint dysfunction (pain from a biomechanical disorder without a demonstrable lesion) has been proposed as a possible etiology of SI pain (145). Among patients with chronic low back pain, a study using a single block technique of the SI joint with a local anesthetic, estimated the prevalence of SI joint pain as between 13% and 30% (141). A study on 54 patients with unilateral low back pain suspected from the SI joint, using a dual local anesthetic block technique, demonstrated an 18.5% prevalence of SI joint-based pain (142).

**Diagnosis of the SI Joint Pain**

The value of clinical data from history and physical examination in the diagnosis of SI joint pain remains controversial (141–143,150,151). Although SI joint pain frequently manifests as pain in the sacral sulcus areas, SI joint pain can refer to the buttock, lower lumbar region, groin, and lower limb (152). However, none of these symptoms, signs or various provocative tests are pathognomonic for SI joint pain. Other sources of low back pain, such as lower lumbar Z-joint arthropathy or degenerative disc disease, can present similarly. By using fluoroscopically guided SI joint blocks to confirm cases of SI joint pain (Fig. 68-17), several authors have shown that clinical medical history and pain provocation tests are not reliable in the diagnosis of SI joint pain (141–143,150,151).

In a clinical trial involving 84 patients with possible low back pain from the SI joint, Dreyfuss et al. studied clinical history and the 12 physical examination tests deemed most reliable by a panel of experts for isolating SI joint pain (143). Fluoroscopically guided intra-articular SI joint injections of local anesthetic and corticosteroid were performed to confirm the diagnosis. The criterion for a positive result was the achievement of at least 90% pain relief postinjection. The study demonstrated that neither the history nor the physical examination data was of significant value in diagnosing SI joint pain (141). Maigne et al. investigated 54 patients with clinical features of chronic low back pain compatible with the origin in the SI joint with the following features: unilateral buttock pain,
tenderness over the SI joint, normal lumbar CT scan, failure of previous epidural, or facet injections (142). They applied seven “SI pain provocation tests” before and after the dual block of SI joint. To be considered diagnostic, patients had to report 100% pain relief after block with 1% lidocaine and at least 75% pain relief after block with 0.25% marcaine. There was an 18.5% rate of positive responders to this dual block. The result demonstrated that none of the SI joint pain provocation tests were able to isolate SI joint pain (142). Slipman et al. performed a diagnostic fluoroscopically guided SI joint injection in 50 consecutive patients with low back pain presumed to be from SI joint (162). A reduction of the VAS rating by at least 80% was considered a positive response to SI joint block. The authors concluded that the various SI joint provocative maneuvers were not useful in diagnosing SI joint pain (162).

A diagnostic imaging study done with bone scan was found to have low sensitivity and high specificity for diagnosing the SI joint syndrome (149). This study excluded patients with SI joint pain from inflammation, such as in a seronegative spondyloarthropathy. However, another study using single photon emission CT (SPECT) was performed in 54 patients with symptoms of low back pain of at least 3 months duration, the presence of higher erythrocyte sedimentation rate, and higher C-reactive protein levels who had not received anti-inflammatory drugs. The results demonstrated high sensitivity (97%) and specificity (90%) in diagnosing inflammatory disease within SI joints (153).

Indications

Although exact guidelines for administering an SI joint injection are unclear, one set of guidelines is as follows: a diagnostic SI joint injection is indicated in patients with pain over the sacral sulcus who have failed to respond to 4 to 6 weeks of directed physical therapy and oral nonsteroidal anti-inflammatory agents (154,155).

Technique

The patient is placed in a prone position. The skin over the sacral area is prepped and draped in a sterile manner. By rotating the C-arm fluoroscope slightly contralaterally, but occasionally ipsilaterally, intermittent fluoroscopy is used to identify the medial joint line when it just separates from the lateral joint line of the SI joint. Some adjustment of the C-arm in the caudal or cephalad plane may then be used to best isolate and visualize the lower portion of the SI joint. The targeted area is the small lucent area just below the joint line. The skin entry site is selected slightly lower than the targeted area, and is infiltrated with a small amount of 1% lidocaine. A 22- or 25-gauge 2.5 to 3.5-in. spinal needle is inserted and directed down to contact the ilium. The needle is then withdrawn 2 to 3 mL and redirected toward the inferior-medial aspect of the joint into the lucent area (156). Typically, the needle tip will bend if it enters the SI joint, and a tactile sense of a sliding into the joint will be appreciated (157). Applying a slight curve at the tip of the needle prior to use may help assist this process (156). A small amount of contrast is then injected to outline the SI joint. If the needle fails to plunge and no contrast flow is seen on fluoroscopy, one technique advocates that the needle be slowly extracted a millimeter at a time, while continuing to maintain pressure on the plunger until there is a loss of resistance (157). Once an SI joint arthrogram without a vascular uptake pattern is demonstrated, anesthetic with or without steroid is injected (depending upon if the injection intent is for therapeutic or diagnostic use, respectively). One milliliter of 2% lidocaine or 0.5% bupivacaine mixed with 40 mg/mL of triamcinolone acetonide, or other equivalent corticosteroid, is injected into the SI joint (158–161). A total of no more than 2.0 mL of volume is generally injected due to the limited volume of the SI joint (160).

Efficacy

Diagnostic Injection of SI Joints

Because the gold standard for proof of SI joint pain is unclear, the sensitivity and specificity of diagnostic SI joint injections has not been clearly established.

Maigne et al. (142) have suggested that the SI joint block has diagnostic value only for pain from intra-articular sources, not for SI joint pain from extra-articular sources such as the periosteum, interosseous ligaments, erector spinae muscles, or fascial elements, all of which contain nociceptors and hence are possible pain generators (147,148). Therefore, an SI joint block procedure that involves injection of an agent into the extra-articular components rather than the joint cavity may show better correspondence to the clinical features. Future studies should address whether the combination of periacetabular and extra-articular SI joint injection with corticosteroid can improve outcomes.

Therapeutic Injection of the SI Joint

The efficacy of SI joint corticosteroid injections has been reported in prospective and retrospective studies of patients with spondyloarthropathy (163,164). In a retrospective study, Slipman et al. reported a significant benefit from SI joint steroid injection in patients with SI joint syndrome (165). Thirty-one patients with chronic SI joint syndrome received an average of 2.1 fluoroscopic-guided SI joint corticosteroid injections. The average follow-up was 94.4 weeks. Of the 29 patients who completed the study, there was a significant improvement in the Oswestry disability score, VAS, and work status (165). Although these retrospective results are encouraging, there are currently no prospective studies on the efficacy of fluoroscopically guided therapeutic SI joint corticosteroid injections.

Radiofrequency Ablation of the SI Joint

Radiofrequency ablation (RF-A) has recently been proposed as a potential long lasting treatment for SI joint pain, and has been gaining more popularity along with other nonsurgical spinal procedures. RF-A involves de-innervation of the SI joint nerves believed to be responsible for generating pain (166,167). It is indicated as a treatment for those patients who have failed more conservative measures, yet only received transient benefit from diagnostic and/or therapeutic injections of the SI joint (168). The true effectiveness of RF-A of the SI joint is unclear, as of yet (169). In contrast to RF-A in
treating lumbar spine facet-mediated pain, which directly targets the medial branches of the dorsal rami, which innervate the facet joints (170), the SI joint has complex innervations (171). Therefore, no consistent procedural technique has been described in the literature. Multiple studies have, in fact, been done using various techniques for RF-A of the SI joint, which are summarized in Table 68-6. The table shows that there are variations among the techniques used in regard to structures, nerves, and patterns of ablation to the SI joint.

The table underscores the fact that there is no standard pattern of ablation and not enough available prospective data to determine which rami or branches should be ablated, or if a pattern technique (i.e., “leap frog” vs. “strip lesion”) is more efficacious. The studies do not show uniformity and additional studies to determine if RF-A is useful for treating chronic SI joint pain are warranted.

CONCLUSIONS

The SI joint can be a common source of low back pain. Diagnostic SI injections may be more reliable than physical exam maneuvers in confirming SI pain. SI joint corticosteroid injections may improve both pain and function. SI joint RF-A is being investigated as a potential tool for achieving lasting relief of SI joint pain.

DISCOGRAPHY (DIAGNOSTIC DISC INJECTION)

Although still controversial, discography (diagnostic intervertebral disc injection) is both an imaging study and a provocative physiologic study for determining whether an intervertebral disc is in fact a pain generator in a given patient (Table 68-7). Inserting a spinal needle into the center of the intervertebral disc and injecting contrast dye provides both physiologic information on whether a degenerative disc is painful and on anatomic features of the intervertebral disc. There is currently no other method to establish reliably whether a disc is a patient’s pain generator.

Clinical Presentation of Discogenic Pain

Patients with lumbar discogenic pain typically have low back pain but can also have pain referred to the buttock, hip, groin, thigh, or distal lower limb. (175). Discogenic pain is typically

<table>
<thead>
<tr>
<th>TABLE 68.6</th>
<th>Summary of RF-A of the SI Joint Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author(s)</strong></td>
<td><strong># of Patients</strong></td>
</tr>
<tr>
<td>Buijs et al. (174)</td>
<td>43</td>
</tr>
<tr>
<td>Burnham and Yasui (171)</td>
<td>9</td>
</tr>
<tr>
<td>Cohen and Abdi (168)</td>
<td>9</td>
</tr>
<tr>
<td>Ferrante et al. (166)</td>
<td>33</td>
</tr>
<tr>
<td>Gevargez et al. (173)</td>
<td>38</td>
</tr>
<tr>
<td>Vallejo et al. (172)</td>
<td>22</td>
</tr>
<tr>
<td>Yin et al. (159)</td>
<td>14</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TABLE 68.7</th>
<th>Review of Normal and Abnormal Discography Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pain</strong></td>
<td><strong>Volume Accepted</strong></td>
</tr>
<tr>
<td>Normal</td>
<td>No pain</td>
</tr>
<tr>
<td>Abnormal</td>
<td>Concordant</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

worse with lumbar flexion and unsupported sitting, as intradiscal pressures have been found to be higher in these positions (176). Physical examination should reveal a normal neurologic examination if the discogenic pathology does not affect the nerve roots.

**Radiographic Correlation**

MRI and CT of the lumbar spine can be useful initial assessment tools because they are noninvasive tests that allow for visualization of multiple discs. Although they have high sensitivity (Figs. 68-18 and 68-19) for detecting anatomic disc abnormalities, surgically proven internal disc disruption has been reported in cases of normal-appearing MRIs (177). MRI carries a high rate of false-positive findings, as shown in studies of asymptomatic patients (178). In addition, they cannot provide the physiologic information about whether an abnormal-appearing disc is actually a pain generator. Although high-intensity signal zones (HIZ) in the posterior annulus on MRI have been linked to discogenic pain (Fig. 68-19), the HIZ can also occur in asymptomatic patients (179–182). With no pathognomonic features of discogenic pain and a high false-positive rate of anatomic findings on noninvasive diagnostic tests, the diagnosis of symptomatic lumbar disc diseases requires a physiologic study for better clinical correlation of a patient’s pain with CT or MRI abnormalities. Provocation discography is a physiologic test for discogenic pain.

**Anatomy and Pathophysiology of Discogenic Pain**

The intervertebral disc is a well-innervated structure with A-delta and C-pain fibers (183) containing nociceptive substances such as substance P, calcitonin gene-related product, VIPs in the annulus fibrosis (184–186). Nerve growth factor has been found in both the annulus fibrosis and nucleus pulposus, which may increase pain sensitization (187). In healthy intervertebral discs, only the outer one third of the annulus fibrosis is innervated. Study of intraoperative samples from degenerative intervertebral discs of patients with chronic back pain demonstrated evidence of inward growth of nerve fibers along the radial fissures into the inner annulus (185,186). The presence of the neural structures and nociceptive fibers is believed to be the anatomic basis of chronic low back pain due to degenerative disc diseases. Discogenic pain may occur in internal disc disruption, which is a condition characterized by a degraded nucleus pulposus with radial fissures extending into the peripheral annulus fibrosis (Fig. 68-19) (188,189). The outer margin of the annulus is intact. The nucleus pulposus, reaching the innervated outer annulus through the annular fissures, invokes an intensive local inflammatory process. These inflammatory substances irritate and sensitize the nociceptive fibers in the outer annulus. The threshold for nociceptive mechanical stimulation is lowered in these chemically sensitized nociceptors. Chronic discogenic pain may result from mechanical stimulation of sensitized nociceptors with normal lumbar disc loading. In fact, intraoperative mechanical stimulation of the posterior annulus in the presumed painful segment induced low back pain in one study (190). Degenerative disc disease is believed to account for some 40% of patients with chronic low back pain of unclear origin (191).

**Indications**


- Patients with unremitting spinal pain of greater than 4 months and unresponsive to all appropriate methods of conservative therapy
Patients in whom other investigations have failed to explain the source of pain

Chronic back pain patients who are contemplating intradiscal or surgical procedures such as spinal fusion

technique

Discography is generally performed in a radiologic suite or an operating room. The patient is placed in either a prone or oblique side-lying position. A pillow is placed under the patient's abdomen to reverse lumbar lordosis. The lumbosacral area is prepared and draped in a sterile fashion. The patient's vital signs should be monitored and oxygen saturation recorded with a pulse oximeter. A peripheral intravenous line should be established for light conscious sedation.

An AP fluoroscopic view is then used to identify the appropriate disc level. The fluoroscope is tilted in either a cephalad or caudad direction to best visualize the target disc space on radiograph. The C-arm is then rotated ipsilaterally to place the superior articular process of the subjacent vertebral body in a position that bisects the vertebral body above. An appropriate size spinal needle, typically a 25-gauge, 3.5-in. spinal needle is used to infiltrate the skin and subcutaneous tissue down to the superior articular process with 1% lidocaine. Caution should be exercised not to inject local anesthetic overzealously in the superior articular process area in order to avoid potential spread to the epidural or nerve root areas, thus compromising the patient's ability to perceive pain during the subsequent provocative discography. To reduce the chance of discitis, a two-needle technique with 18- or 20-gauge, 3.5-in. introducers and 22- or 25-gauge, 6-in. inner needles, is recommended. The introducer needle is inserted and directed to the outer edge of the superior articular process in the AP view and just to the anterior border of the superior articular process in the lateral view. The inner needle is then inserted through the introducer and slowly advanced to the center of the nucleus pulposus using alternating AP and lateral views. If the patient complains of radicular pain or paresthesias in a nerve root distribution during needle advancement, the needle should be withdrawn and redirected. If a pressure-controlled system is used for the injection, the needle is connected to an injection system with a manometer that is filled with nonionic and water-soluble contrast. The injectionist then injects contrast slowly, monitoring the pressure reading and the patient's reports of pain simultaneously. The opening pressure is the pressure reading at the first appearance of dye in the disc on fluoroscopy. The endpoint of the injection occurs when the patient reports concordant pain (defined as reproduction of pain in the same location and intensity), or when 2 mL of total volume is injected into the disc or the pressure reading reaches 90 lb per square inch (psi). To help minimize the chance of discitis, 5 to 10 mg of cefazolin can be injected into each disc before needle removal. One level above and below the disc with concordant pain should be also studied so as to serve as a control (Figs. 68-20 to 68-22).

Information obtained from discography includes the volume of contrast injected, the patient's pain response (no pain, dissimilar or discordant pain, similar pain, and exact pain provocation or concordant pain), degree of resistance to injection, morphology of the nucleogram, and postdiscogram CT morphology of the disc (193).

Postinjection CT scanning provides an axial view of the injected discs (Fig. 68-23). Patterns of radial and concentric annular fissures are more clearly defined in this plane. Postdiscography CT scanning should be performed within 2 hours of the discogram to prevent diffusion of dye out of the nucleus. The Dallas Discogram Description, in addition to recording the pain and contrast volume injected, describes morphologic degrees of annular degeneration and disruption (194). Annular degeneration and disruption are graded by the percentage of the contrast injected that fills the annulus and annular fissures toward the outer annulus as revealed by the...
contrast (Table 68-8) (195). A study by Derby et al. found that there was a significant correlation between the extent of annular disruption on CT and the rate of symptomatic disc on discography, with significant differences between grade 3 and 5 versus grade 0 and 2 (196).

A positive discogram requires a concordant pain response, an abnormal nucleogram, and a normal control level. In contrast, multilevel painful discs without a normal control disc on injection are unexplainable and cannot be regarded as positive.

**Validity of Discography**

In a prospective and controlled study in 1990, by applying modern manometry and postdiscography CT technology, refined needle placement techniques, less irritating contrast dye, stressing concordant pain, or discordant pain in addition to disc morphology (195), Walsh concluded that using stringent criteria, the false-positive rate of lumbar discography in asymptomatic individuals is 0% and the true-positive rate in symptomatic patients is 89% (189). However, Walsh’s study is limited by its small sample size (i.e., only ten controls and seven patients).

In a retrospective study, Derby et al. reported that a patient with a chemically sensitive disc, defined as one having a concordant pain response provoked with intradisc pressure of less than 15 psi as measured by a manometer, has a better outcome with interbody fusion versus intertransverse fusion surgery (197). This better surgical outcome presumably resulted from the removal of the mechanical load or stimulation to the chemically sensitive disc through removal of the painful intervertebral discs and stabilization with interbody fusion (197). Nevertheless, these investigators pointed out that discography should never stand alone as a diagnostic tool and sole factor for a clinical decision. The presence of a chemically sensitive disc does not rule out other coexisting sources of pain, nor does it exclude patients with excessive pain magnification. Discography should be used in conjunction with other diagnostic tests, such as MRI and CT scans, as well as the patient’s history, physical examination findings, response to a comprehensive therapeutic program, and psychometric profile information. Therefore, careful candidate selection is the key to maximizing clinically valuable data from discography. A recent study by Carragee et al. in 2000, although controversial, revealed a high false-positive rate with lumbar discography in a mixed population of 26 asymptomatic patients (182). A significant positive pain response and pain-related behavior with discography were found in 10% of the pain-free group and in 83% of the somatization disorder group. Discs with annular disruption were more likely to be painful on injection. The study concluded that if strict criteria are applied, the rate of false-positive discography may be low in subjects with normal psychometric profiles. Schwarzer et al. in 1995 (183) reported that a discogram is likely to provide highly specific information on the painful lumbar disc when the following characteristics are present: (a) unremitting low back pain persisting after 6 months of conservative care;
Complications of Discography
Potential complications of discography include discitis (0.05% to 4%), nerve root injury, subarachnoid puncture, chemical meningitis, bleeding, and allergic reaction (192). Use of sterile technique, a two-needle technique, as well as intradiscal antibiotics, can reduce the risk for infection (198).

SUMMARY
In addition to providing imaging of disc morphology, discography is the only provocative physiologic test that can provide information on whether a degenerative disc is the source of pain. In appropriately selected patients, discography is a safe, reproducible, objective diagnostic tool when testing includes volume, pressure, fluoroscopic abnormalities, and pain provocation. Discography results help guide surgical management of chronic back pain. A negative discogram or a multilevel positive discogram without a painless control level can help to exclude a patient as a surgical candidate. Conversely, a single- or double-level positive discogram, along with other clinical data, can help to guide operative planning with respect to fusion or other surgical procedures. This potentially optimizes surgical outcomes. Whether diagnostic discography will actually change surgical outcome depends on many factors, such as a patient’s psychosocial factors, job situation, and the surgical technique employed. The true value of diagnostic discography will likely remain controversial until a prospective, randomized, double-blind, and controlled clinical trial on the outcomes of surgery based on discography is carried out.

COCCYX INJECTIONS
Coccyx pain (coccydynia) apparently occurs far less commonly than lumbosacral pain. Coccydynia can be a severe and persistent pain, causing significant suffering, frustration, and functional limitations (199). These patients often have a history of coccyx trauma (e.g., from a fall or childbirth) resulting in contusions, fractures, dislocations, or other injuries. However, other cases are idiopathic. Typical symptoms include focal tailbone pain, particularly worse with sitting and sometimes immediately worse upon going from sit-to-stand (199). Careful screening should seek to exclude malignancies of the spine (e.g., chordoma) and intrapelvic structures (e.g., rectum) (199,200). Useful diagnostic studies may include x-rays, MRI, CT scans, and colonoscopy (199). Dynamic radiographs comparing coccygeal alignment and angulation while sitting (weight bearing on the coccyx) versus standing may reveal dynamic instability (dislocations) not visualized via non—weight-bearing studies (201,202).

Injections for coccyx pain may include focal corticosteroids placed at the posterior coccyx or into a sacrococcygeal or coccygeal joint, ideally performed under fluoroscopic guidance to maximize injection accuracy and minimize the risk for inadvertent puncture of the rectum or other nearby structures. Diagnostic nerve blocks can include local anesthetic blockade of the somatic posterior coccygeal nerve fibers as well as the sympathetic nerve fibers (ganglion Impar) anterior to the coccyx (203–207). For both diagnostic and therapeutic injections, it is frequently helpful to simultaneously block both the anterior (sympathetic) and posterior (somatic) nerves. This combination can more completely shut off allafferent inputs from the coccyx, with resultant relief implying that the coccyx is the source of pain. Also, effective local anesthetic blockade can provide therapeutic benefit, sometimes including complete and permanent relief (208). The mechanism is perhaps via disrupting hyperactive and/or hypersensitive afferent reflex arcs or sympathetically maintained pain. The various technical approaches to the ganglion Impar include injecting a spinal nerve needle from inferior to the coccyx (above the anus), or passing the needle through the sacrococcygeal joint or through the intracoccygeal joints (203–207). Fluoroscopic guidance is crucial for safe and effective performance of sympathetic nerve blocks at the ganglion Impar, especially given the close proximity to the bacteria-laden rectum. Nerve ablation in the coccyx region may be beneficial for selected patients with coccydynia (207). Most patients with tailbone pain will obtain adequate relief via nonsurgical treatments such as injections, thus often avoiding the need for surgery and its potential complications (209).

CHAPTER CONCLUSIONS
In selected patients, spinal injections can offer diagnostic and therapeutic benefits. Spinal injection procedures have become an important part of clinical armamentarium for the musculoskeletal physiatrist. Efforts should be directed toward selecting appropriate patients and incorporating injection procedures into an overall rehabilitation program to relieve pain, maximize function, and improve quality of life.

REFERENCES


PART VI | MANAGEMENT METHODS


Palliative Care Symptom Management

INTRODUCTION

Palliative care involves a holistic approach to comprehensive symptom management. This represents a significant departure from the emphasis on disease-modifying treatment that dominates allopathic medicine. The specialty originally developed as a body of medical expertise geared toward the care of dying patients. Its roots are intertwined with those of the hospice movement. Unfortunately, the equation of palliative care with hospice medicine persists to the field’s detriment. This limited view fails to recognize that palliative care principles can be applied to alleviate symptoms and enhance quality of life (QOL) in many stages of illness, not solely the terminal.

Palliative care specialists have advocated a parallel model of care delivery with integrated use of disease-modifying and symptom-oriented therapies across disease continua. This model proposes that palliative care be provided from the time of diagnosis onward with emphasis appropriate to patients’ symptom burden irrespective of their prognoses. As suggested by the schematic of this model presented in Figure 69-1, palliative care plays an increasing role with disease progression, and symptom control ultimately becomes the sole agenda at the end of life. This approach is based on the belief that palliative care has a role in the management of all disease states associated with an intense and adverse symptom complex.

Origins

Before the advent of the modern medical era, interventions were largely palliative given the limitations of disease-modifying treatment. Comfort-oriented care was integral to medical management, frequently representing practitioners’ sole option. Physicians’ reliance on skillful symptom control informed medical education, research, and discourse. A steady shift in emphasis away from this symptom-oriented approach began with the integration of antibiotics into mainstream care as these were among the first dramatically effective disease-modifying agents in widespread use. Although simultaneous medical advances occurred, none paralleled the dramatically improved outcomes afforded by antimicrobials. Growing interest in disease mechanisms led to a marked expansion in pathophysiological research and consequent refinement of our understanding of responsible agents and mechanisms. Elucidation of disease processes opened the door for generations of “rational” medical therapies. A path was forged from the lab bench to the bedside, and the medical establishment became heavily invested in scientifically driven disease modification.

The success of scientifically driven disease modification is unequivocal. However, its achievements have not come without casualties. Expertise and interest in symptom control waned with medicine’s focus on disease-targeted interventions. In the early 1960s, clinicians began to recognize that dying patients or those whose diseases failed to adequately respond to therapy were being marginalized and suffering as a consequence. Clinician avoidance of dying patients has been well documented (1–3). In an effort to more sensitively and consistently care for the dying, the hospice movement was born. From humble origins, hospice care expanded and is now globally embraced. With the growth of care delivery models for patients with terminal illnesses came a demand for empiric research in symptom management, greater physician expertise, and empirically derived treatment algorithms. Palliative care represents the response to these demands.

Relevance to Physical Medicine and Rehabilitation

Enhanced disease-modifying therapies are allowing patients to survive longer in the advanced stages of many diseases. This trend has characterized the treatment of acquired immunodeficiency syndrome (AIDS) and many cancers. Although life may be extended, these patients are not spared functional decline (3). Progressive debility and dependence are increasingly leading these patient populations to seek physiatric attention. Multiple surveys have demonstrated that patients with advanced disease strongly desire functional autonomy and are deeply concerned over the prospect of becoming an increasing caretaker burden (4–6). It is critical that physiatrists rise to the challenge of providing expert and effective functional restoration within the confines imposed by their diseases. The fiscal, emotional, and social burden of unmitigated impairment is extremely high (7–9). Studies suggest that rehabilitative interventions can successfully restore function even in hospice-bound patients and those with widespread metastatic disease (10,11).

To date, little interface has occurred between rehabilitation medicine and palliative care. This is an unfortunate circumstance for both fields, particularly for the many patients with advanced disease who stand to benefit from the fields’ complementary skill sets of symptom control and functional restoration. Symptoms must be adequately controlled if patients are to participate fully
and benefit from rehabilitation. Without physiatric expertise, patients are unable to translate their enhanced comfort into self-directed activities, social integration, and improved QOL.

**Chapter Overview**

The majority of this chapter discusses the most prevalent adverse symptoms in patients with advanced disease. Palliative care is, in essence, expert, multimodal symptom control in refractory disease states. The pathophysiology, causes, and treatment of each symptom or clinical problem are addressed. The second, briefer, portion of this chapter outlines exercise as palliation and describes current, integrated palliative care delivery models. Both parts of the chapter have a common goal of imparting a body of knowledge geared toward the preservation of dignity and autonomy in symptomatic patients with progressive illnesses.

**SYMPTOM ASSESSMENT**

A uniform approach to symptom assessment will maximize the chances of appropriate diagnosis and treatment. Assessment of all symptoms should be conducted in a rigorous and structured fashion. Most symptoms can be caused by a range of etiologies, some of which may be amenable to definitive, disease-modifying treatment. For example, pain from discrete bone metastases is best managed with focal irradiation. Effective treatment requires identification of all contributing factors. Assessment should be conducted with the awareness that symptoms are dynamic processes that vary in intensity, quality, frequency, and level of associated distress. For some symptoms, reliable assessment instruments have been developed and validated. Table 69-1 lists commonly used instruments.

A critical dimension of palliative care, distinct from other fields of medicine, is the overriding emphasis on patients’ subjective experience. Subjective distress and discomfort displace disease processes as the primary therapeutic targets. Observer ratings of symptom severity, both those of caregivers and clinicians, correlate poorly with patient ratings and are generally an inadequate substitute for patient reporting. Therefore, the presence and the severity of symptoms must be accepted at each patient’s word.

An approach that has been widely endorsed for the comprehensive assessment of pain can be readily applied for the assessment of any symptom. Semantics may interfere with reliable symptom reporting. The terms nausea, pain, anxiety, and fatigue vary in significance and import to different patients. For example, some patients use the term nausea to describe abdominal discomfort, pain, distention, or early satiety. Soliciting multiple qualitative descriptors helps to clarify the precise characteristics of the symptom being evaluated. For many symptoms, the extent of physiologic and anatomic disturbance does not correlate directly with symptom intensity. Symptoms can augment one another when they occur in clusters. Interactions among symptoms should be explored to clarify the degree to which each symptom, or its treatment, induces or exacerbates other physical or psychological symptoms. Evaluation should attempt to distinguish whether...
### Table 69.1 Commonly Utilized, Validated Instruments for Global Symptom Assessment, as well as the Assessment of Specific Symptoms

<table>
<thead>
<tr>
<th>Domain of Assessment</th>
<th>Instrument</th>
<th>Reference</th>
</tr>
</thead>
</table>

### Evaluation of QOL

<table>
<thead>
<tr>
<th>Domain of Assessment</th>
<th>Instrument</th>
<th>Reference</th>
</tr>
</thead>
</table>

### Evaluation of pain

<table>
<thead>
<tr>
<th>Domain of Assessment</th>
<th>Instrument</th>
<th>Reference</th>
</tr>
</thead>
</table>

### Evaluation of impaired cognition

<table>
<thead>
<tr>
<th>Domain of Assessment</th>
<th>Instrument</th>
<th>Reference</th>
</tr>
</thead>
</table>

### Evaluation of depression

<table>
<thead>
<tr>
<th>Domain of Assessment</th>
<th>Instrument</th>
<th>Reference</th>
</tr>
</thead>
</table>

### Evaluation of anxiety

<table>
<thead>
<tr>
<th>Domain of Assessment</th>
<th>Instrument</th>
<th>Reference</th>
</tr>
</thead>
</table>

### Evaluation of delirium

<table>
<thead>
<tr>
<th>Domain of Assessment</th>
<th>Instrument</th>
<th>Reference</th>
</tr>
</thead>
</table>

(continued)
TABLE 69.1 | Commonly Utilized, Validated Instruments for Global Symptom Assessment, as well as the Assessment of Specific Symptoms (Continued)

<table>
<thead>
<tr>
<th>Domain of Assessment</th>
<th>Instrument</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evaluation of dyspnea</td>
<td>The Medical Research Council Scale</td>
<td>Medical Research Council Committee on Research into Chronic Bronchitis: Instruction for Use of the Questionnaire on Respiratory Symptoms. Devon, England: W.I. Holman; 1966.</td>
</tr>
</tbody>
</table>

Symptoms are (a) concurrent but unrelated in clinical etiology, (b) concurrent and related to the same pathological process, (c) concurrent with one symptom directly or indirectly consequent to a pathologic process initiated by the other symptom, or (d) concurrent with one symptom consequent to therapy directed against the other. These distinctions are challenging, but critical for successful comprehensive management.

Obtaining a detailed history with special emphasis on patients’ medication histories is fundamental to symptom management. Medications are an extremely prevalent source of adverse symptoms. Patients may be taking multiple medications for primary disease modification, symptom control, prophylaxis against complications, and medical comorbidities. Interactions between these medications must be considered. A medication can indirectly influence a symptom by reducing the protein binding or slowing the metabolism of medication(s) directly responsible for the symptom. Comorbid medical conditions should also be carefully reviewed. Diabetes, chronic renal insufficiency, autoimmune disorders, hepatic failure, peripheral vascular disease, among many other common diseases, can both influence adverse symptoms and be influenced by them. Too often, clinical attention becomes exclusively focused on the primary disease process, leading to neglect of other potentially important pathologic contributors.

Appropriate imaging and serologic evaluations will vary contingent on the symptom being evaluated and the clinical context in which it occurs. The choice of appropriate diagnostic tests should be guided by the stage of disease, the prognosis, the risk/benefit ratio of any proposed test or intervention, and the desires of the patient and the family. Evaluations that are associated with significant discomfort, risk, or expense, or that demand an extended time commitment must be weighed in light of the extent to which clinical management will be meaningfully altered.

The introduction of item response theory (IRT)-based instruments into the mainstream is a recent development in symptom assessment that is likely to have far-reaching influence on both clinical and research endeavors (25–29). Given the emphasis that the National Institutes of Health and preeminent psychometricians have placed on IRT-based assessment, readers will be well served by some familiarity with these concepts. Simply put, IRT represents an alternative to classical test theory, which maintains that symptom assessment tools must be administered as fixed-length instruments in their entirety without varying item order or presentation to remain valid. In contrast, IRT asserts that individual items, not multi-item instruments, are the foundation of accurate assessment and that reordering items or utilizing subsets of items does not compromise validity (30). IRT-based methods allow clinicians and researchers to select limited numbers of items from much larger item pools in order to match the sensitivity of the selected items with the symptom intensity of their population of interest (31). For example, items (e.g., “My pain is so bad that I can’t breathe: Always, Often, Sometimes, etc.”) that are sensitive to the intense pain experienced by cancer patients may be uninformative when administered to patients with relative mild pain related to tendonopathy. More comprehensive and intelligible descriptions of IRT, Rasch modeling, and computer adaptive testing are widely available. Interested readers are encouraged to explore the following resources: *Item Response Theory: Principles and Applications* by Hambleton and Swaminathan and *Health Status Assessment for the Twenty-First Century: Item Response Theory, Item Banking and Computer Adaptive Testing* by Revicki and Cella in *Quality of Life Research*, 2004.

PAIN

Pain becomes a disturbing reality for a majority of patients at some time over the course of progressive illness, particularly those with cancer and AIDS. The degrading impact of
uncontrolled pain on functional autonomy and QOL has been well documented (32). Adequate pain control is essential if QOL and functional autonomy are to be preserved.

Epidemiology

Pain epidemiology has been most extensively studied in cancer cohorts with reported prevalences ranging from 14% to 100% (33), with prevalences of 50% to 70% being noted in patients receiving active treatment (34). Prevalence rates increase to 60% to 90% near the end of life (35). Between 36% and more than 50% of ambulatory cancer patients report pain severe enough to significantly impede function (36). Inadequate pain management has been well documented, with as many as 40% of ambulatory cancer patients receiving insufficient analgesia according to World Health Organization (WHO) standards (36). In one report, 82% of patients with advanced cancer admitted to hospice were undermedicated and in significant pain (37). Patients who are female, of minority status, or former substance abusers are at highest risk for undertreatment (36). The prevalence of uncontrolled pain is alarming, given that the combined use of pharmacologic and interventional analgesic approaches is able to control 90% of severe cancer pain (38). However, it is worth noting that estimates of cancer pain undertreatment derive from research published in 1993 and may not accurately reflect current rates or predisposing characteristics.

Causes

Pain in the setting of advanced illness often arises simultaneously from multiple sources. This is particularly true in the context of disease progression, which may create significant diagnostic and therapeutic uncertainty. An organizational structure facilitating characterization of most disease-related pain syndromes has been proposed that involves characterizing the pain’s precipitant (e.g., disease- or treatment-related) and presumed pathophysiology (e.g., nociceptive somatic, nociceptive visceral, or neuropathic) (39). This rubric has been most consistently applied to patients with cancer-related pain. Table 69-2 lists common cancer pain syndromes and places them within this organizational framework. The distinction between nociceptive and neuropathic pain pathophysiology has both prognostic and therapeutic significance (40). Nociceptive pain responds extremely well to disease-modifying and opioid-based therapies. Neuropathic pain is pain that arises through deranged functioning of the nervous system (40). Neuropathic pain frequently proves more refractory to treatment than nociceptive pain but may respond to adjuvant analgesics.

Pain generators have been extensively described for cancer patients. Direct tumor effects on bone, nerve, soft tissue, or visceras account for 65% to 75% of pain (41–43). Anticancer therapies are responsible for 15% to 25% of cancer-related pain (41–43). Pain syndromes may develop following surgery (e.g., postthoracotomy pain), radiation therapy (e.g., osteoradionecrosis), or chemotherapy (e.g., peripheral neuropathy). Five to ten percent of patients report pain unrelated to their cancer or its therapy (41–43). Patients with AIDS, far-advanced arthritides, and severe vasculopathies frequently experience intense pain characteristic of these syndromes.

Treatment

A wide array of analgesic strategies is currently available. Pain management approaches commonly used in advanced cancer are listed in Table 69-3. Opioid-based pharmacotherapy is the preferred strategy for managing cancer pain. This approach has been generalized to the treatment of pain arising from other terminal conditions on the basis of theoretical rather than empiric grounds. The widespread success of this approach in controlling cancer pain has led to strong endorsement by numerous international and national organizations, as well as a federal task force (44–47). Countless chapters, books, and review articles published over the past decades have advocated the skillful integration of disease-modifying therapy with the use of opioid and nonopioid analgesics.

Disease-modifying therapy is the treatment of choice, irrespective of etiology, because it may permit definitive, lasting pain relief. For cancer patients, disease-modifying treatments are geared toward reduction or eradication of tumor. Therapies may include surgery, chemotherapy, radiation therapy, and the delivery of radiopharmaceuticals. The reduction of tumor effects on normal tissue may reduce or eliminate pain. Despite widespread practice of delivering chemotherapy to address cancer pain, the analgesic benefits of this
Unless contraindicated, patients should undergo serial nonopioid trials, particularly those with pain related to osseous metastases or inflammation. The maximal effective dose and degree of patient responsiveness varies considerably among NSAIDs (39). Therefore, sequential trials may be needed to identify the optimal agent and dose. NSAID-related gastropathy and antiplatelet effects in thrombocytopenic patients are a significant concern. Adverse consequences may be reduced through the use of cyclooxygenase-2-selective inhibitors or gastroprotective agents (55). Care must be taken in patients with advanced illness to monitor for hepato- and nephrotoxicity related to acetaminophen and NSAID use, respectively (40).

Adjuvant analgesics, also referred to as coanalgesics, comprise a large class of drugs with the capacity to offer pain relief in certain conditions. A majority of adjuvant analgesic trials have been conducted in patients with chronic neuropathic pain (e.g., postherpetic neuralgia). Their efficacy in this setting has served as justification for their use in pain arising from terminal disease states. The appropriateness of this strategy is attested by the inclusion of adjuvant analgesics as an integral part of all current guidelines (50,56–58). Adjuvant analgesics’ role as the preferred agents for control of neuropathic pain is supported by numerous clinical trials and literature reviews (59,60). The majority of adjuvant analgesics used to treat pain in advanced disease can be classed as antidepressants, anticonvulsants, or oral local anesthetics. However, the successful use of corticosteroids, antispasmodics, neuroleptics, and osteoclast inhibitors has been anecdotally described, and some agents have been found effective in controlled trials. The most frequently used agents are discussed below.
The tricyclic antidepressants have been most extensively studied among the adjuvants and have well-documented efficacy (61). Secondary amines (e.g., nortriptyline and desipramine) may offer less analgesia relative to tertiary amines (e.g., amitriptyline and imipramine), but have fewer associated anticholinergic side effects (61,62). Selective serotonin reuptake inhibitors have produced poor to mixed results in clinical trials. If cotreatment of depression is desired, the serotonin norepinephrine reuptake inhibitor duloxetine may be a superior choice, although its efficacy for cancer pain remains untested (63).

Trials have established the anticonvulsant carbamazepine as a first-line agent for intermittent lancinating pain. However, the associated risk of causing leukopenia limits its use in patients at risk for hematological abnormalities (64). Both gabapentin and pregabalin produce good results in the treatment of lancinating continuous neuropathic pain (65–68). Their benign side-effect profiles and the lack of drug-drug interactions make them reasonable first- or second-line choice for adjuvant analgesic therapy. Oxcarbazepine, levetocetam, and lamotrigine may be trialed on theoretical grounds as their utility in controlling pain associated with terminal illness lacks empirical support (40).

Topical agents such as 2% to 4% viscous lidocaine may alleviate pain from mucosal lesions. The degree of analgesia afforded by dermally applied topical preparations is unclear. In a recent blinded, controlled trial, lidocaine-impregnated patches offered no greater relief than placebo for postincisional neuropathic pain in cancer patients (69). A variety of topical salves combining various agents including amitriptyline, ketamine, gabapentin, etc. have been anecdotally endorsed, but empirical evidence is limited and their costs are often prohibitive (70,71).

In addition to NSAID therapy, several adjuvant analgesics are uniquely beneficial in the management of pain arising from bone metastases. These include steroids, bisphosphonates, and radiopharmaceuticals. Dexamethasone is the steroid of choice at doses ranging from 1 mg twice daily to 100 mg daily, with standard doses being 16 to 24 mg/d (40,72). Bisphosphonates inhibit osteoclast activity and effectively alleviate pain from bone lesions (73). Single doses of the radiopharmaceuticals, strontium-89 and samarium, may lastingly alleviate pain arising from diffuse bone metastases as demonstrated in breast and prostate cancer cohorts (74,75).

An extensive international literature describes the control of cancer pain through the use of high-dose, long-term opioid therapy. Recommendations that doses be increased to “effect or side effect” stem from the lack of a therapeutic ceiling for opioid analgesia, although large dose increments may be required at high dose ranges to achieve incremental benefit. Many patients require doses as high as or higher than the equivalent of 5 g of oral morphine per day (76,77). Clinical success and an ever-growing literature has given rise to increased acceptance and comfort with high-dose opioid-based analgesia (76,77).

Opioids can be subclassified as pure agonists (e.g., morphine, oxycodone), partial agonists (e.g., buprenorphine) based on their interaction with endogenous opioid receptors. With few exceptions, only pure μ agonists (e.g., morphine, methadone, oxycodone, hydromorphone, fentanyl) should be used to manage pain in patients with advanced diseases. Physiologic withdrawal syndromes may be precipitated by the use of agonist-antagonists in patients who are physically dependent on opioids (78).

Administration of μ-receptor agonists has become increasingly refined over the past decades as new synthetic formulations and routes of administration have become available. Irrespective of these advances, general clinical practices have not been significantly altered (79). Algorithms described in numerous guidelines, chapters, and review articles endorse the following steps:

1. Establish opioid analgesic requirements in “naïve” patients with liberal, as-needed dosing of an immediate-release (IR) (also referred to as normal-release) formulation.

2. Once patients’ use of the IR formulation has stabilized, generally after a 1 to 2 week interval, their total daily IR consumption is converted to an equivalent total daily dose of a sustained-release (SR) opioid preparation. When possible, the same IR and SR opioid should be used. Long-acting or SR formulations are generally IR opioids placed in time-release matrices so that they may be dosed at longer intervals. Methadone may be dosed on a schedule of every 8 or 12 hours and therefore may be included among the SR or long-acting opioids (80).

3. Patients’ SR preparations should be supplemented with “rescue” or demand IR doses, typically 5% to 15% of the total daily dose, for preemptive use in anticipation of pain or when the SR preparation fails to provide adequate analgesia.

For example, a patient consuming an average of sixteen 10 mg IR oxycodone tablets per day would require twice daily 80 mg SR oxycodone tablets with rescue doses of 8 to 24 mg of IR oxycodone. Subsequent upward titration of the SR opioid to optimal effect is based on the frequency of IR rescue use. For patients with activity-related incident pain and minimal pain at rest (a scenario commonly encountered in patients with painful bone metastases), higher IR rescue doses and lower SR doses may provide superior analgesia and reduce cumulative daily opioid consumption.

The pharmacologic management of pain in patients with advanced disease requires flexibility in both the choice of medication and the dosing strategy. Contingency planning for rapid analgesic escalation should pain abruptly worsen is critical to long-term success. Oral medications offer both ease of administration and upward titration. At the time of the writing of this chapter, morphine and oxycodone are the only pure μ-agonists available in both IR and SR oral preparations. Methadone effectively alleviates cancer pain when dosed both “around-the-clock” and as an as-needed rescue medication (81). However, methadone’s erratic pharmacokinetics demand
PART VI | MANAGEMENT METHODS

cautions and use by experienced clinicians to avoid inadvertent, delayed toxicity (82). Further, clinicians should be alert to the possibility of QT prolongation and torsades de pointes with daily methadone doses greater than 300 mg (83, 84).

The transdermal route benefits patients lacking or expected to lose the capacity for enteral absorption. Fentanyl and buprenorphine patches are currently the only transdermal opioid formulations available in the United States. Data support the use of both transdermal fentanyl (85, 86) and buprenorphine (87, 88) to control cancer pain, despite the fact that buprenorphine is a partial μ-receptor agonist (89). Fentanyl can also be administered transmucosally in the form of a lozenge that has rapid onset and a discrete duration of action making it a good choice for patients experiencing short intervals of severe pain with physical activity.

Parenteral opioid administration offers the capacity for rapid upward titration and accurate establishment of patients’ analgesic needs. Parenteral hydromorphone is a potent opioid (7.5 times more potent than parenteral morphine) that can be administered at concentrations of 20 mg/cc making subcutaneous administration feasible. Fentanyl, oxymorphone, methadone, and morphine are also available for parenteral administration. Morphine’s availability in long- and short-acting tablets, elixir (2, 4 mg/cc, 20 mg/mL), long-acting pellets (deliverable via nasogastric [NG] tube), and parenteral formulations, as well as highly concentrated subcutaneous preparations, allows tremendous dosing versatility. Also, morphine is also quite inexpensive.

Opioid-related side effects need not limit upward dose titration (90). Persistent somnolence can be managed with methylphenidate or modafinil after all nonessential centrally acting medications have been discontinued. A reasonable starting dose of methylphenidate is 2.5 mg twice daily (last dose provided no later than 2 p.m.). Typical effective daily doses range from 10 to 40 mg. Common side effects—including nausea, hyperhidrosis, xerostomia, and constipation, among others—can generally be effectively managed and need not limit opioid therapy (90). When optimal opioid side-effect management proves ineffective, opioids should be rotated (40).

Invasive Techniques

Most patients with advanced disease attain satisfactory relief of pain through approaches that incorporate oral or parenteral analgesic therapy with other noninvasive strategies. Anesthetic and neurosurgical techniques are generally reserved for patients who are unable to achieve a satisfactory balance between analgesia and side effects, or inadequate analgesia despite escalating doses with sequential, strong opioid drug trials (91). Commonly employed regional analgesic techniques include intraspinal (e.g., epidural and intrathecal) opioids and local anesthetics, intraventricular opioids, and regional local anesthetic blockade (92). Neurolysis of the celiac plexus, superior hypogastric nerve plexus, and ganglion impar may afford control of visceral pain (93). Discrete somatic pain arising from malignant invasion or compression can transiently be relieved through peripheral nerve blocks. Temporary peripheral neural blockade is frequently used diagnostically to clarify the location of a primary pain generator. Neuroablative surgical techniques, including rhizotomy, neurolysis of primary afferent nerves or their ganglia, cordotomy, and C1 midline myelotomy, may be used for refractory syndromes in patients with limited prognoses (94).

FATIGUE

Fatigue represents a widely prevalent and highly distressing symptom for cancer patients. It is difficult to overestimate its adverse impact on function because fatigue can significantly compromise patients’ incentive and ability to comply with psychiatric interventions. Patients describe fatigue as devastating to many life domains, degrading their vocational, familial, and societal roles (95). Patients with cancer rate fatigue as more distressing than any other symptom including pain (96). Fatigue arguably presents the greatest challenge to functional restoration in patients with terminal illness as it constrains patients’ ability to engage in rehabilitation and adhere to regimens required for success. Fatigue may become so severe that patients interrupt or abandon disease-modifying treatments (96).

Current understanding of fatigue is largely derived from studies involving patients with cancer. All formal definitions reflect the experience of “cancer-related fatigue,” and generalizations to other disease states should be made with awareness of this orientation. The National Consortium of Cancer Centers (NCCN) defines fatigue as “an unusual persistent subjective sense of tiredness related to cancer or cancer treatment that interferes with usual functioning” (97). Diagnosis of “cancer-related fatigue” per the International Classification of Disease (ICD)-10 requires a known tumor and daily persistence of the symptom for ≥2 weeks plus 6 of the following 11 complaints: diminished energy, increasing need for rest, limb heaviness, diminished ability to concentrate, increased interest in engaging in normal activities, sleep disorder, inertia, emotional lability due to fatigue, perceived problems with short-term memory, and postexertional malaise exceeding several hours (98).

Although not included in most formal definitions additional, accepted characteristics of “cancer-related fatigue” include disproportionate tiredness relative to patients’ exertional level. Additionally, “cancer-related fatigue” is not relieved by rest or sleep and engenders subjective weakness (99–101). Fatigue reduces patients’ mental capacity and psychological resilience (102, 103). Patients may report reduced motivation, capacity to attend, concentration, and difficulty acquiring new learning (104).

Epidemiology

The overall prevalence of fatigue is generally greater than 75% (12, p. 75). Among cancer patients who experience fatigue, most surveys report that one half to two thirds describe the symptom as “highly distressing” (124, p. 53). Fatigue occurs most commonly and severely during administration of anticancer therapies, being reported by as many as 99% of patients receiving chemotherapy (105). Over 40% of patients rate their pain as
“severe” or ≥7 on an 11-point numerical rating scale (106). Fatigue is frequently present at the time of diagnosis, increases throughout treatment, and persists for years after the completion of therapy (104, 107). Cancer patients who report fatigue have decreased disease free and overall survival (108).

Pathophysiology

The processes underlying fatigue remain poorly understood. Biological response modifiers (e.g., cytokines) elaborated by tumors and by the body in response to tumors/treatments have been blamed for fatigue due to well-characterized temporal associations between their administration for therapeutic purposes and the onset of severe fatigue. The fact that tumor necrosis factor-α (TNF-α) and interleukin-6 are elevated in some patients with chronic fatigue syndrome, and that synthetic antibodies targeting proinflammatory cytokines reduce fatigue in patients with rheumatoid arthritis further implicate biological response modifiers (109–111). However, levels of circulating cytokines do not correlate with fatigue severity (112). Therefore, biological response modifiers have received less attention in current mechanistic discussions.

The role of serotonin (5-HT) dysregulation in fatigue has been studied due to evidence that 5-HT contributes to pathological fatigue states as well as fatigue experienced by healthy subjects during intense exercise. Tryptophan, a precursor of 5-HT, levels increase significantly during normal exercise within the brain (113) and patients with chronic fatigue syndrome have elevated serum tryptophan levels (114). However, central 5-HT concentrations do not correlate with the presence or the intensity of cancer-related fatigue suggesting that 5-HT dysregulation is neither required nor sufficient to engender fatigue (115, 116).

Data directly link aberrant hypothalamic-pituitary-adrenal (HPA) axis function to fatigue. Breast cancer survivors with fatigue demonstrate blunted stress responses reflected in low salivary cortisol levels when stressed relative to unaffected controls (117). Investigators speculate that irregularities in diurnal cortisol regulation may be highly relevant to the genesis of pathological fatigue (118). Since cortisol, cytokines, and 5-HT levels regulate one another, HPA axis has been proposed as a unifying mechanism by which cytokines and 5-HT may produce fatigue (104).

Fatigued cancer patients demonstrate abnormal sleep-wake cycles and rest-activity patterns. Diminished daytime physical activity and excessive movement during sleep are associated with fatigue in patients with breast and colorectal cancers (119, 120). Fatigued patients with stage IV colorectal cancer, display diminished variation in rest-activity patterns (108, 121).

A definitive mechanism that accounts for all pathological fatigue remains elusive and likely reflects its underlying complexity and nonuniform pathogenesis. Different mechanisms may interrelate and dominate in different patients and disease states.

Causes

Fatigue may relate to myriad factors, both treatment and disease related, which are potentially modifiable. A discrete source of fatigue can be identified in some patients, leading to symptom alleviation. More often, many potential contributors of unclear relative importance can be identified. Endocrinopathies, blood dyscrasias, degraded sleep quality, centrally acting medications, steroid myopathy, and cachexia represent possible engendering and aggravating factors.

Anemia has received consistent emphasis as a fatigue source since roughly 50% of patients with solid tumors are anemic at diagnosis and many anticancer therapies reduce red blood cell count (122). The relevance of anemia to cancer-related fatigue has diminished with the recognition that the time course of fatigue differs substantially from fluctuations in blood counts. Normalization of hemoglobin levels through blood transfusion or erythropoietin administration fails to consistently alleviate fatigue. No specific reduction or increase in hemoglobin levels has been definitively linked to quantitative changes in QOL or performance status.

Endocrinopathies are underdiagnosed and, in general, easily rectified. Disruption of the adrenal axis, thyroid gland, testes, and ovaries by disease- or treatment-related processes occurs commonly in a range of terminal disease states. Appropriate serologies allow identification of deficiencies.

Centrally acting pharmaceuticals are commonly administered to patients with advanced illness and represent a cornerstone of palliative care. Whenever possible, needless or ineffective sedating medications should be eliminated or replaced by less problematic alternatives. Withdrawal trials may clarify the benefit to side-effect ratio of centrally acting pharmaceuticals in helping patients and clinicians appreciate their utility (103).

Treatment

Methylphenidate has been extensively used to alleviate fatigue in patients with cancer. Four of seven clinical trials to examine the efficacy of methylphenidate in alleviating fatigue resulted in benefit (123–126). A small pilot study combining exercise and methylphenidate, also (open-label), detected benefit (125). Results from two controlled, double-blinded studies conflict but trial designs differed in maximal doses, trial duration, and inclusion criteria (126, 127). Based on current evidence, a trial of methylphenidate is justified for patients with debilitating fatigue starting at a dose of 5 to 10 mg/d. Dose-limiting toxicities associated with methylphenidate include anorexia, insomnia, anxiety, confusion, tremor, and tachycardia.

Modafinil, also a central nervous system (CNS) stimulant, has been studied in two open-label trials. Survivors of breast cancer experienced reduced fatigue, as did patients with brain tumors, while taking modafinil (128, 129). Modafinil causes mild, tolerable side effects (e.g., headache, anxiety, nausea) that resolve upon discontinuation. Treatment can be started at 100 to 200 mg/d and titrated to a maximal dose of 400 mg/d.

Antidepressants inconsistently alleviate fatigue in depressed patients with cancer. SR bupropion (100 to 300 mg/d) in two open-label case series alleviated fatigue after two to four weeks.
of treatment (130,131). Paroxetine has been more rigorously studied with double-blind, control trials. Two such efforts detected enhanced mood but no change in fatigue (115,116). The effect of serotonin-norepinephrine reuptake inhibitors on fatigue has yet to be examined.

Patients with advanced, metastatic cancer and fatigue experienced benefit from corticosteroids in two randomized, double-blind, crossover studies (132,133). Prednisone also reduced fatigue in a less rigorous open-label study (134). Side effects of sustained steroid use may limit their utility to patients with advanced illnesses.

L-carnitine dosed at 500 to 600 mg/d alleviated fatigue in three small, open-label trials (135–137). All studies enrolled patients with cancer. These promising results have yet to be replicated with a blinded study design.

Exercise may offer the greatest benefit of all fatigue alleviating therapies, though its feasibility has yet to be researched in advanced illness. Aerobic conditioning consistently reduces fatigue in cancer survivors and in patients receiving adjuvant chemotherapy (138–140). No harm has been associated with aerobic exercise at 75% to 80% maximal heart rate. Resistance training has been more limitedly studied, yet the encouraging pattern of reduced fatigue persists (141). Androgen-deprived patients with prostate cancer experienced improved QOL, fatigue, and strength after performing two sets of 8 to 12 repetitions at 60% to 70% of one maximal repetition over a 12-week period (142). Participants in exercise programs have never reported reduced QOL or increased fatigue, irrespective of their intensity.

Increasing evidence supports the efficacy of psychosocial interventions to mitigate fatigue. Many different types of interventions including support groups, psychosocial nursing interventions (e.g., coping, stress management, problem solving), energy conservation and activity management, psychologist- or self-administered stress management, nurse-administered cognitive behavioral symptom management intervention, and structured psychiatric group intervention have been studied with randomized, controlled study designs (107). Consistent fatigue reduction has been reported across this wide variety of therapeutic approaches. In spite of convincing evidence, programs offering psychosocial treatment are not widely available and patients with advanced illness may actually experience increased fatigue with intensive and demanding programs (143).

**NAUSEA AND VOMITING**

**Epidemiology**

Chronic nausea and vomiting are prevalent and distressing problems in patients with advanced cancers and other diseases. The reported prevalences vary depending on the clinical population surveyed and the assessment method with reported prevalence rates from 40% to 70% in the palliative care setting (144,145). In a case series conducted by Fainsinger et al. (146), 71 out of 100 patients in a palliative care unit required treatment for nausea in the last week of life. Uncontrolled nausea and vomiting have the capacity to degrade patients’ physical and psychological well-being, seriously undermining their QOL (147). Nausea is more common than vomiting and defined as chronic when it lasts more than 1 week in the absence of clear, self-limiting sources, for example, chemotherapy (148). Control of nausea and vomiting can provide patients with a sense of control over their body and life, decrease anxiety and fear, decrease caregiver burden, and better enable patients to autonomously perform ADLs.

**Pathophysiology**

Much of our understanding of the pathophysiology of nausea and vomiting derives from research on patients receiving chemotherapy and radiation therapy. It is critical that a distinction be made between nausea and vomiting. Nausea is a subjective symptom involving an unpleasant sensation experienced in the back of the throat and epigastrium that may not result in vomiting (149,150). Vomiting, in contrast, is the reflexive elimination of gastric contents through forceful contraction of abdominal muscles to expel toxic substances. The mechanisms underlying vomiting are far better understood than those responsible for nausea.

Distinct sites in the brain, namely the vomiting center (VC) and the chemoreceptor trigger zone (CTZ), are responsible for control of vomiting (151). The VC is an interrelated neuronal network in the medulla including the nucleus tractus solitarius and the dorsal motor nucleus of the vagus. Projections to the VC arise from peripheral pathways, the vestibular system, and the CTZ located in the area postrema of the medulla, as illustrated in Figure 69-3 (152). The CTZ lies outside the blood brain barrier enabling it to sample toxins and metabolic abnormalities. The CTZ responds to chemotherapeutic agents, metabolic products, opioid, and bacterial toxins with stimulation of the VC resulting in emesis. Vomiting induced by raised intracranial pressure, taste, smell, and anxiety is mediated through afferent VC projections from the cortex, diencephalon, and limbic system (153). Projections from the vestibular center participate in motion-induced nausea and vomiting via input from the vestibulocochlear nerve (154). Opioids can alter the sensitivity of the
associated nausea and vomiting depends on stimulation of muscarinic acetylcholine and histamine type 1 receptors and may be alleviated by antagonizing these receptors with scopolamine, dephendydramine, or promethazine (152).

Figure 69-4 illustrates redundancies in the distribution of neurotransmitters throughout the pathways that support nausea and vomiting. The limited number of receptor types is clinically advantageous as a single therapeutic agent can target multiple sites. 5HT3 antagonists, for example, block both peripheral pathways and the CTZ.

Causes
Chronic nausea and vomiting in advanced disease is most often multifactorial. Identification of all contributing factors increases the likelihood of successful management. Autonomic failure causing delayed gastric emptying has been frequently demonstrated among patients with advanced cancer. Autonomic failure is more common among cancer patients with poor performance status and malnutrition (156). The precise mechanisms underlying autonomic failure have yet to be fully elucidated. However, poor nutritional status suppresses the activity of the sympathetic nervous system in other clinical populations (157,158). The presence of cardiovascular effects, including postural hypotension, syncope, and fixed heart rate,
should direct attention to autonomic dysfunction as a driving etiology in chronic nausea.

Drugs frequently trigger or aggravate nausea in patients with advanced disease. Among cancer patients, opioid analgesia can cause nausea and vomiting after an initiation or an escalation in the dose. Opioids induce nausea by stimulating the CTZ, gastroparesis, constipation, and by increasing the sensitivity of the vestibular center (152). Many other drugs can cause nausea, including NSAIDs, antibiotics, iron supplements, tricyclic antidepressants (TCAs), selective 5HT₃ receptor inhibitors, and phenothiazines.

Constipation and bowel obstruction are common among advanced cancer patients (155,159). There are many factors that predispose this population to the development of constipation, including immobility, poor oral intake, dehydration, autonomic failure, opioid analgesics, and other medications. Suspicion should be high for undiagnosed constipation in all patients with advanced disease. Bowel obstruction is less common, but nevertheless is an important potential contributor to chronic nausea (159). Most cancer patients experience a gradual progression from mild partial to complete obstruction, making the initial diagnosis challenging (160). Other potential contributors to chronic nausea include biochemical abnormalities (e.g., hypercalcemia, hyponatremia, liver failure), CNS metastases, increased intracranial pressure, and psychological factors (e.g., fear, anxiety). Deranged serum ion concentrations can also induce or aggravate nausea.

Assessment

Characterizing the intensity, frequency, triggering factors, and qualitative dimensions of patients’ experience is an essential initial step in the assessment of nausea and vomiting. Nausea should be distinguished from early satiety, bloating, and reflux symptoms. Determining patients’ ability to keep fluids and solids down may indicate the need for nonoral routes of medication administration. In addition to a thorough history and physical examination including the frequency and quality of patients’ bowel movements, imaging and serological tests may be required to identify all factors contributing to nausea (148). An abdominal flat-plate x-ray can be helpful in determining whether constipation or obstruction is present. CT scans and MRIs of the brain or abdomen can elucidate the presence of an intracranial source or malignant obstruction. Contrast enhancement is needed for detection of leptomenigeal disease. Serological tests to exclude renal impairment, hepatic failure, and other metabolic abnormalities, including hypercalcemia, hypokalemia, and hyponatremia, should be initiated.

Treatment

For the majority of patients, treatment of nausea will be pharmacological. Causative factors that can be feasibly eliminated or treated should be addressed. Efforts to treat potential contributing factors should not delay the delivery of appropriate pharmacological therapy. There are currently nine principal classes of drugs used as antiemetics in palliative care: butyrophenones, prokinetic agents, cannabinoids, phenothiazines, antihistamines, anticholinergics, steroids, 5HT₃ receptor antagonists, and benzodiazepines. As stated previously, the pathway(s) triggering nausea should be identified and pharmacological interventions chosen in a rational manner. Once the appropriate receptor has been identified, the most potent antagonist should be selected. Route of administration must be considered in patients for whom enteral administration is not feasible. Table 69-4 (152) lists commonly utilized agents in each of these classes.

*Butyrophenones are dopamine antagonists including haloperidol and droperidol. As D₂ antagonists, they powerfully inhibit the CTZ particularly when used in combination with other drugs such as 5HT₃ receptor antagonists. Since these agents do not affect GI motility, they can be safely used in bowel obstruction (148). Additionally, haloperidol can be administered subcutaneously and is safe in renal failure.*

<table>
<thead>
<tr>
<th>Table 69.4 Commonly Used Antiemetic Agents</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Class</strong></td>
</tr>
<tr>
<td>Butyrophenones</td>
</tr>
<tr>
<td>Prokinetic agents</td>
</tr>
<tr>
<td>Cannabinoids</td>
</tr>
<tr>
<td>Phenothiazines</td>
</tr>
<tr>
<td>Antihistamines</td>
</tr>
<tr>
<td>Anticholinergics</td>
</tr>
<tr>
<td>Steroids</td>
</tr>
<tr>
<td>5HT₃ receptor antagonists</td>
</tr>
<tr>
<td>Benzodiazepines</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
Prokinetic agents include metoclopramide and domperidone. Metoclopramide is the most commonly used drug in this category and antagonizes dopamine both centrally at the CTZ and peripherally in the GI tract. In addition, metoclopramide demonstrates weak antagonism at the 5HT₃ receptor. In addition to blocking CTZ receptors, metoclopramide helps to bring normal peristalsis in the upper GI tract. Metoclopramide has been advocated as an effective first-line agent for patients with chronic nausea (161).

Several studies have demonstrated the efficacy of the cannabinoid dronabinol as an antiemetic agent for the treatment of chemotherapy-induced nausea and vomiting (162–166). It is commonly dosed at 2.5 mg twice daily. Cognitive side effects including somnolence, confusion, and perceptual disturbance are frequent, particularly in patients with borderline cognitive impairment related to their disease or other medications. Dronabinol is generally considered a second- or third-line antiemetic.

Phenothiazines, such as chlorpromazine and thiethylperazine, have survived as a mainstay of antiemetic therapy. Phenothiazines exert their therapeutic effects by antagonizing dopamine. Their tranquilizing effects may or may not be desirable, contingent on each patient’s symptom complex. One advantage is their availability in oral, rectal suppository, parenteral, and SR oral preparations. Their mechanism of action is distinct from the 5HT₃ antagonists and dexamethasone, so they can be effectively combined with these agents to augment therapeutic efficacy. Vigilance for extrapyramidal side effects should remain high throughout the duration of therapy. Twenty-five to fifty mg of diphenhydramine can be given to prevent these effects.

Antihistamines, including cyclizine, promethazine, and dimenhydrinate, are useful antiemetics that act to block histamine receptors in the VC and on vestibular afferents. They can be particularly useful if a vestibular afferent input is believed to contribute to the nausea. They are rarely used for antiemetic monotherapy in palliative care. Anticholinergics (e.g., scopolamine) are seldom used but have the advantage of sublingual, subcutaneous, and transdermal routes of administration. They are anticholinergic near the VC and additionally reduce peristalsis and inhibit exocrine secretions, thereby palliating nausea from GI obstruction (167,168).

Corticosteroids, particularly dexamethasone, are a common and highly effective component of aggressive antiemetic therapy. They have powerful nonspecific antiemetic effects that are currently not well understood. Corticosteroids can decrease edema surrounding tumors, thereby reducing intracranial pressure, which may contribute to nausea. They are extremely effective when used in combination with other agents, including metoclopramide and 5HT₃ receptor antagonists (169,170).

Serotonin antagonists have become increasingly accepted as first-line antiemetic agents. The 5HT₃ receptors have been discovered in the CTZ, VC, and in the terminals of the vagal afferents in the gut. Ondansetron, dolasetron, and granisetron, the principal commercially available 5HT₃ receptor antagonists, may be prohibitively priced for many patients. They have been widely used and researched for the prevention and management of acute chemotherapy-induced nausea and vomiting (170,171). To date, no trials have been published evaluating the efficacy of 5HT₃ antagonists in the management of chronic nausea and less expensive D₂ antagonists are equally effective at the end of life (172,173). However, mirtazapine, an antidepressant that antagonizes the 5HT₁A receptor, has shown promise in alleviating nausea and other distressing, cancer-related symptoms (174,175). It is important to appreciate that 5HT₃ receptor antagonists may reduce the incidence of emesis without alleviating nausea.

The benzodiazepine lorazepam exerts its antiemetic effects within the CNS. It is highly effective when used in combination with other antiemetic agents. Additional desirable therapeutic effects include anxiolysis and amnesia. A trial conducted by Malik et al. suggested that lorazepam is highly effective in decreasing the incidence of anticipatory vomiting (176).

A variety of additional pharmacological agents are used to alleviate chronic nausea and vomiting. Thalidomide has enjoyed renewed interest in recent years for its anticachectic, antipyretic, sedative, and antiemetic properties (177). Thalidomide is believed to exert its antiemetic effects via a central mechanism. The pregabational agent megestrol acetate has been predominantly studied for its anticachectic properties. However, several studies reported significant reductions in nausea following the initiation of megestrol therapy (178,179). Octreotide, a somatostatin analogue, effectively reduces nausea in patients with intestinal obstruction by reducing GI secretions (180,181). The neurokinin-1 antagonist aperitant alleviates chemotherapy-induced nausea and vomiting and may control nausea of alternate causes (182,183). Olanzapine, an atypical antipsychotic, has shown promise in the treatment of nausea (184–186) presumably via blockade of a variety of receptors including dopamine, acetylcholine, histamine, and serotonin receptors (152).

Despite growing insight into the neurotransmitters responsible for nausea and vomiting and enthusiasm for a rational, mechanistic approach to palliation, several authors advocate empirical initial treatment with D₂ receptor antagonists such as metoclopromide, prochlorperazine, or haloperidol irrespective of the implicated pathways (161,187–189). There have been no head-to-head comparisons of mechanistic and empirical approaches. However, proponents of mechanistic approaches acknowledge that first-line use of D₂ receptor antagonists makes sense as these agents block CTZ activity (152). It has also been suggested that multiagent regimens such as the combination of a corticosteroid with either a 5HT₃ receptor antagonist (167,190) or metoclopromide (191) offer the best chance of effectively managing chronic nausea. However, a recent trial showed no benefit over placebo when dexamethasone was added to metoclopromide for patients with nausea related to advanced cancer (169).

When a first-line agent fails to adequately alleviate nausea and vomiting, the adequacy of dosing should be ascertained. Often, as-needed dosing proves insufficient and patients require around-the-clock dosing (144). Adding a second agent is often preferred to changing antiemetics agents since multiple neurotransmitters act on the relevant sites (152).
Multiple nonpharmacologic approaches to controlling nausea and vomiting have been proposed and their efficacy has been demonstrated in clinical trials (192,193). Simple measures such as eliminating strong odors and other emetic triggers, consuming small frequent meals, and reducing oral intake during intensely emetic intervals are beneficial (194,195). Self-hypnosis, progressive muscle relaxation, biofeedback, guided imagery, cognitive distraction, music therapy, and systemic desensitization are behavioral techniques that have been used with varying degrees of success. The combined use of guided imagery and progressive muscle relaxation or music is highly effective in reducing chemotherapy-induced nausea and vomiting. Such integrated regimens may be most effective at reducing nausea, although further trials are required before a definitive statement can be made.

Acupuncture and acupressure have been widely used in the treatment of chemotherapy-induced nausea, motion sickness, and hyperemesis gravidarum. Several rigorously designed clinical trials have found that acupuncture can mitigate nausea associated with chemotherapy administration (196,197). Studies have focused on stimulation of pericardium or master of the heart 6 point located on the volar surface of the forearm approximately three finger breadths proximal to the wrist crease. Few studies have investigated acupuncture for chronic nausea in the terminally ill. An isolated study evaluated the use of acupuncture wrist bands in advanced cancer patients and found them to be ineffective (198).

ANOREXIA AND CACHEXIA

Anorexia describes the lack or loss of appetite leading to reduced oral intake and weight loss. Cachexia, in contrast, is a maladaptive state of metabolic dysfunction characterized by (a) progressive wasting of fat and muscle tissue, and (b) disruption of preferential fat loss induced by starvation (199). Wasting occurs out of proportion to reduced oral intake, is resistant to conventional nutritional support, and preferentially affects skeletal muscle over cardiac or smooth muscle (200). Vital organs remain relatively preserved even in advanced cachexia. Anorexia and cachexia are important clinical problems in advanced disease because they associate with reduced performance status, as well as increased risk of treatment failure, side effects, and mortality (201,202). Cachexia is the most common single cause of death documented in cancer patients being directly responsible for 25% deaths and affecting about half of all patients (203). Functional debility associated with cachexia may be profound due to impaired respiratory muscle function when total weight loss exceeds 15% (204). Death generally occurs when patients lose 25% to 35% of their total body weight (205,206).

Epidemiology

The prevalence of cachexia depends on the stage of the inciting disease process. For example, although over 50% of all patients with cancer lose body weight, weight loss occurs much more commonly among patients with stage IV cancer with the prevalence rising as high as 86% in the last 1 to 2 weeks of life (207). As many as 20% of all cancer deaths are caused directly by cachexia (208). Cachexia is a common characteristic of the end stages of many chronic illnesses with estimated prevalences as follows: congestive heart failure (CHF), 16% to 36% (209,210), COPD, 27% to 33% (211,212), and ESRD, 30% to 60% (213).

Pathophysiology

Disequilibrium between oral intake and muscle wasting results from interactions between the host and the tumor or disease processes, in the case of noncancer cachexia, which remain partially understood. Tumors elaborate local proinflammatory cytokines (214,215) that activate host inflammatory cells, which pass through the tumor, circulate systemically, and ultimately induce the hepatic acute phase protein response (APPR) (214–216). Interleukin-6 has been implicated in APPR induction though a host of other cytokines may participate as well (216–218). Tumors also secrete 2 well-characterized pro-cachectic factors: proteolysis-inducing factor (PIF) (219,220) and lipid-mobilizing factor (LMF) (221–223). The complex systemic interactions that may underlie cachexia are illustrated in Figure 69-5.

Host factors contribute to cachexia through the tumor-induced APPR and the neuroendocrine stress response. The APPR is a part of the body’s systemic reaction to tissue injury, neoplastic growth, and infection (203). Up to 50% of patients with solid tumors may present with APPR, manifest in elevated serum C-reactive protein levels, at the time of diagnosis (224). Although elevated acute phase reactants are strongly associated with cachexia, causal mediators between these phenomena have not been definitively identified. One hypothesis suggests that the liver becomes a protein sink competing with skeletal muscle for dwindling amino acid supplies (225). Mechanisms through which the neuroendocrine stress response may contribute to cachexia are also unclear. Increased circulating levels of the anabolic skeletal muscle mediators insulin, growth hormone, and testosterone coupled with decreased circulating levels of the catabolic mediators cortisol and myostatin implicate the neuroendocrine stress response. Further, autonomic dysregulation, another manifestation of the stress response, is more prevalent in cachectic patients (226,227). That such associations may reflect epiphenomena rather than direct causal linkages has been a disappointing realization for cachexia researchers as is evident in their study of TNF-α.

Brief mention of TNF-α is warranted as it has been a focus of intense cachexia research for decades and has been the focus of recent treatment trials. The discovery that TNF-α induces cachexia in animal models, with the typical increased protein degradation that plagues affected patients, led investigators to hope that TNF-α predominantly mediates human cachexia and could be a therapeutic target (228). However, clinical trials with TNF-α synthesis and receptor blockade have been disappointing and therapeutic interest has shifted to PIF (222,223,229).
At the cellular level, the importance of reduced protein synthesis relative to increased protein degradation in cachexia is a source of debate with results differing across studies (199). Several models of cachexia suggest that both are important and must be targeted for therapeutic efficacy. Events underlying reduced protein synthesis are complex but relate to dysregulated phosphorylation of factors that initiate protein translation (230). Efforts to understand accelerated protein degradation have focused on the ubiquitin-proteasome pathway that appears to be of major importance in conditions associated with significant loss of muscle mass (231). The pathway has been comprehensively reviewed (232). In brief, multiple ubiquitin molecules are serially conjugated to proteins destined for degradation that distinguishes them from the 26S proteasome complex that confines proteases preventing indiscriminate degradation. Several enzymes that ligate ubiquitin to myosin heavy chain proteins and to troponin 1 are highly expressed in muscle wasting conditions and are believed to play an important role in cachexia (233–235). The literature is mixed regarding the precise impact of various cytokines and procachectic factors on the ubiquitin-proteasome pathway. TNF-α, IL-6, and PIF influence transcription of ubiquitin pathway proteins in animal models; however, whether results apply to clinical cachexia is unclear (199).

Of significance for physiatrists is the growing recognition that physical activity may play an important mediating role in the induction and progression of cachexia. Historically, the suspicion that increased host resting energy expenditure (REE) contributed to cachexia raised concern over exercise. However, the recognition that host REE varies as reflected in reports of reduced (236), normal (237), or elevated (238) energy expenditure, coupled with reports that cancers with the highest prevalence of cachexia (gastric and colorectal) are associated with reduce REE has put fears to rest. A study of cachectic pancreatic cancer patients revealed their REE and total energy expenditure was comparable to home-based spinal cord injured patients (239), levels that may aggravate sarcopenia (240). Clinical efforts to treat cachexia through physical activity are described in the treatment section as follows.

**Treatment**

Cachexia treatment options have become increasingly limited as evidence demonstrates poor efficacy in previously standard therapies. Nutritional approaches are now deemphasized since dietary counseling and total parenteral nutrition fail to attenuate cachexia (241). As discussed more fully in the nutritional section, unless patients are malnourished, strategies to enhance dietary intake do not improve QOL, functional performance, or survival.

**Appetite stimulants:** Pharmacologic interventions have survived empirical scrutiny little better than dietary approaches. Appetite stimulants, formerly the standard of care and still widely used, target anorexia and may induce weight gain but do not increase lean muscle mass or extend survival in cancer cachexia. Progestational agents are synthetic derivatives of progesterone for oral use with anabolic effects that are independent of their antitumor properties. Progestational agent megestrol acetate, believed to downregulate cytokine synthesis and the subject of recent reviews and meta-analyses, has proven disappointingly limited in its capacity to decelerate cachexia (242,243). Megestrol acetate improves appetite in roughly 30% of treated patients leading to short-term weight gain.

**FIGURE 69-5.** Systemic interactions mediating cachexia.
stabilization but similar to other nutritionally based approaches does not enhance QOL or survival (244–246). Loss of lean muscle mass remains unaffected by megestrol acetate (247). In light of their potential toxicities, venous thrombosis, peripheral edema, vaginal bleeding, impotence and adrenal insufficiency, the cost-to-benefit ratio of progesterational agents may not be defensible for most patients (241,248,249). The usual dosage of megestrol acetate is 80 to 160 mg orally four times daily.

Corticosteroids have also been widely utilized to stimulate appetite and thereby attenuate cachexia. Glucocorticoids suppress TNF-α production and inhibit prostaglandin synthesis. Multiple studies evaluating the effect of dexamethasone and prednisolone have failed to demonstrate significant weight gain (132,250–252). Although corticosteroids may improve patients’ subjective well-being, they have a concerning side-effect profile and hence are not suitable for long-term use (241). A steroid trial in cachectic patients may be warranted for enhanced sense of well-being and to treat associated adverse symptoms. However, there is no current evidence to suggest the reversal or mitigation of cachexia can be expected.

Cannabinoids are marijuana derivatives that act either by interfering with cytokine activity or by acting on endocannabinoid receptors in the hypothalamus, intestine, and adipose tissue and have putative orexigenic effects (245,253). However, a well-conducted trial of delta-9-tetrahydrocannabinol versus cannabis extract versus placebo failed to demonstrate significant differences between the treatment arms with respect to appetite or QOL (254). In light of a side-effect profile that may limit use of analgesics and other symptom-modifying agents (e.g., vomiting, drowsiness, mental clouding) cannabinoids cannot be currently endorsed for management of cachexia.

The upstream approach: Increasing implications of proinflammatory cytokines in animal cachexia models has spurred researchers to evaluate therapies targeting events “upstream” of intracellular proteolytic pathways. Drugs of interest are those capable of inhibiting the elaboration of cytokines (pentoxifylline, melatonin, cyclo-oxygenase-2 inhibitors, and thalidomide) and those that interfere with cytokine actions (anticytokine antibodies, suramin, and anti-inflammatory cytokines) (241,255). Similar to the promising results in murine models, most clinical trials in humans have also been encouraging. Randomized, controlled designs have been used to evaluate the following agents: pentoxifylline (256), melatonin (257,258), ibuprofen (259), etanercept (260) (TNF-α inhibitor), and thalidomide (261–263) with results being mixed. Only thalidomide reversed loss of lean body mass in both open-label and controlled trials (261–263). Further, thalidomide treatment was associated with improvement in a broad array of adverse symptoms (261–263). In mixed patient populations, 300 mg daily produced an average weight gain of 4.5% compared with 0.9% in patients receiving placebo (264). Of note, given thalidomide’s complex immunomodulatory and anti-inflammatory properties with promiscuous effects on sundry physiological systems, it remains unclear whether these positive trials, in fact, support the upstream approach (261–263).

The “downstream” approach: Growing insight into the biochemical and molecular mechanisms responsible for muscle degradation has targeted specific intracellular pathways for potential therapeutic intervention. Strategically, such intracellular approaches are, as yet, theoretically based and clinically limited to eicosapentaenoic acid (EPA) trials. EPA interferes with proteasome activity through different mechanisms in experimental conditions and may, therefore, inhibit the ubiquitin-proteasome pathway (265). The results of promising pilot studies administering EPA to patients with cancer-related cachexia were not been replicated in large, multicenter trials (266,267). However, in a post hoc analysis of this trial, a significant improvement in physical functioning was detected among patients receiving EPA (268). Adding EPA to megestrol acetate added no incremental benefit with respect to patients’ performance status or QOL (244). A recent Cochrane review concluded that empirical data do not currently support use of EPA for cancer cachexia (269).

Stimulation of muscle protein anabolism: The capacity of exercise to increase muscle mass and endurance has led investigators to study its efficacy in cachexia. Animal models suggest that exercise may be a useful adjunctive therapy. In murine models, spontaneous, semivoluntary, and forced-endurance exercise increased protein synthesis in skeletal muscle (270–273). Translation of these promising results to human trials has been exceedingly limited. A program of combined resistive and aerobic training in male AIDS patients significantly enhanced strength, but did not alter lean body mass (274). Concurrent testosterone use during a progressive resistive exercise program afforded more promising results by significantly increasing extremity cross-sectional areas in patients with AIDS-related wasting (275).

Anabolic androgenic steroids administered without prescribed physical activity have proven beneficial in a variety of disease states: severe burns (276), HIV infection (277), COPD (278), and cancer (279). However, findings remain insufficiently robust to support routine use of these agents. Nandrolone decanoate and oxandrolone have been most widely studied with neither agent emerging as superior for the management of cachexia (265).

At present, few definitive conclusions can be drawn regarding the treatment of cachexia. Nutritional supplementation alone is of little benefit. Pharmacologic interventions, excepting thalidomide, have been generally disappointing. A trial of conservative exercise with or without pharmacological manipulation is indicated in most patients.

DYSPNEA

Dyspnea has been defined as an uncomfortable awareness of breathing. Although most people can intuit the experience of this symptom, there is no universally sanctioned definition. For many patients, “the description of, “an unpleasant sensation of difficult, labored breathing,” adequately characterizes their experience. Dyspnea may undermine functional performance
and serve as a deterrent to physical activity making it an important therapeutic target for physiatrists.

**Epidemiology**

Prevalence rates of dyspnea at the end of life range widely depending on the population studied and the mode of assessment. In the National Hospice Study, 70% of 1,754 patients experienced dyspnea during the last 6 weeks of life (280). More than 28% of these patients rated the severity of their symptom as moderate or worse. Prevalence rates of between 24% and 60% have been reported in cancer cohorts, with unexpectedly high rates among patients with extrathoracic malignancies (281–283). Dyspnea is also commonly experienced by patients with noncancerous illnesses. The prevalence of dyspnea varies across diseases: COPD, 95%; CHF, 61%; stroke, 37%; amyotrophic lateral sclerosis, 47%; and dementia, 70% (284,285). Reported prevalences among patients with AIDS range from 11% to 62%, although changes since the advent of modern HIV therapies have not been well characterized (286–288).

Despite epidemiological inconsistencies, dyspnea is clearly a common symptom in patients with advanced disease and prevalence rates increase near the end of life. Reports suggest that dyspnea is more problematic in men, older patients, and those with lower self-reported QOL (289,290).

**Pathophysiology**

The mechanisms that regulate normal breathing participate in the experience of dyspnea through complex interactions between breathing abnormalities and the perception of those abnormalities in the CNS. Regulation of breathing depends on interplay between the CNS respiratory center, chemoreceptors, and afferent input from respiratory muscles. Figure 69-6 illustrates a proposed schema for the interplay between these elements. The medulla houses the respiratory center, which receives input from central and peripheral chemoreceptors, as well as peripheral mechanoreceptors. The respiratory center also receives input from the cerebral cortex (291,292). Chemoreceptors, both centrally and in the periphery, respond to levels of oxygen and carbon dioxide in the blood. These chemoreceptors stimulate the respiratory center and possibly the cerebral cortex, leading to an increase in respiratory rate (293,294). The sensation of dyspnea may be intensified through peripheral mechanoreceptor-induced stimulation of the cerebral cortex in response to tachypnea driven by the respiratory center. Mechanoreceptors sense altered muscle tension within the abdominal wall, diaphragm, and accessory muscles of respiration in tension within the lung.

Dyspnea depends on cortical stimulation. Stimulation of mechanoreceptor can produce the sensation of dyspnea in the absence of increased respiratory activity (292,294). Three principal abnormalities give rise to dyspnea. These include an increase in respiratory effort to overcome a certain load (e.g., pleural effusion), an increase in the proportion of respiratory muscle required to maintain normal workload, and an increase in ventilatory requirements (e.g., hypoxemia) (292). Often all three abnormalities coexist. A striking inconsistency between symptom intensity and the level of functional abnormality often characterized dyspnea. For example, severe asthmatics may not report subjective dyspnea despite marked airflow obstruction (295). Similar findings have been noted in COPD patients with severe respiratory abnormalities (296). Psychological factors can strongly influence the intensity with which dyspnea is experienced and expressed (289,297).

**Causes**

Abnormalities that directly and indirectly affect the cardiopulmonary system can give rise to dyspnea. Premorbid cardiac or pulmonary disease can leave patients exceedingly vulnerable to dyspnea with minimal physiological derangement.

Cancerous sources of dyspnea abounds. Tumor within the lung parenchyma, pleura, or pulmonary lymphatics can significantly reduce oxidative capacity. Cancerous spread may arise from primary intrathoracic tumors or from metastatic disease. Tumorous caking of the pleura as seen in mesothelioma can cause restrictive lung disease with resultant dyspnea. Malignant pleural and pericardial effusions may produce a subjective sense of air hunger. Tumor can also extrinsically compress portions of the airway.

Abdominal ascites and hepatic distension can limit diaphragmatic excursion. Severe anemia may produce dyspnea. Cachexia with associated muscle wasting will increase the percentage of maximal muscle capacity required for normal breathing rendering affected patients vulnerable to rapid respiratory muscle fatigue. Infectious etiologies (e.g., pneumonia) must be considered, particularly in neutropenic patients. Despite the variety of potential causes, dyspnea often persists after reversible causes have been addressed.

**Assessment**

Dyspnea varies in intensity according to patients’ activity levels. Patients may attempt to limit their experience of...
dyspnea by curtailing physical activity (298). Consequently, symptom intensity ratings alone may not adequately capture symptom severity. Comprehensive assessment should characterize dyspnea-related activity avoidance (299). Assessment should focus on potential contributing clinical factors, particularly those that may be reversible. This mandates a comprehensive history and physical examination. Attention should be paid to associated signs and symptoms such as chest pain, fever, rigor, cough, cachexia, abdominal distension, and pleurisy. Associated symptoms, particularly pain and anxiety that correlate with dyspnea intensity, may prove to be critical therapeutic targets (289,290). Helpful diagnostic tests include chest radiography, electrocardiography, echocardiography, pulmonary function tests, arterial blood gases, serum electrolytes, and complete blood count. Abdominal and chest computed tomography scans can elucidate the extent of tumor spread within these body cavities.

Treatment
Symptomatic treatment of dyspnea should occur in parallel with efforts to manage specific pathophysiological contributors. For patients with cancer, the initiation of a novel chemotherapy regimen may offer greatest relief if dyspnea is related to tumor effects. Irradiation, including brachytherapy, also offers the capacity to rapidly reduce local tumor burden and associated obstruction (300). Drainage of pleural or pericardial effusions or abdominal ascites may afford dramatic and abrupt relief. Empiric antibiotic therapy may be appropriate if pneumonia is suspected. Pulmonary consultation may clarify the utility of bronchoscopic approaches to address luminal obstruction or narrowing (298). Rigid bronchoscopes, balloon dilatation, lasers, electrocautery, stents, argon plasma coagulation, and cryoprobes have been utilized to address airway obstruction, generally for intraluminal lesions (298,301).

Symptomatic management is based on three principal elements: oxygen, pharmaceuticals, and supportive measures (302). Randomized, controlled studies have reported dyspnea relief in hypoxic patients with cancer receiving supplemental oxygen (303–307). Patients without rest or exertional hypoxia did not derive benefit. Helium-enriched air may be superior to either room or oxygen-enriched air for patients with cancer or COPD experiencing exertional dyspnea (308,309). A majority of well-designed trials in patients with COPD have shown supplemental oxygen to be beneficial (310–312). However, Liss and Grant found that the delivery of different oxygen concentrations was not superior to room air on resting dyspnea in COPD patients (313). A recent systematic review found little evidence that oxygen enhances functional capacity in COPD patients (314). Use of noninvasive positive-pressure ventilation during walking may enhance the benefits of oxygen among patients with severe COPD (315). Results are mixed in controlled, randomized trials evaluating the effect of supplemental oxygen in patients with CHF (316,317). A trial of supplemental oxygen is warranted in all dyspneic patients given the ease with which it can be conducted and its benign side-effect profile.

Pharmacotherapy of dyspnea is based on a limited number of drug classes: opioids, benzodiazepines, corticosteroids, and bronchodilators. Opioids have been extensively used to palliate dyspnea. A majority of studies evaluating the efficacy of opioids in the treatment of COPD-associated dyspnea have reported benefit (318–326). Efficacy was demonstrated using hydrocodone, dicyclomine, hydromorphone, codeine, and morphine. All trials of opioid pharmacotherapy for cancer-related dyspnea using subcutaneous morphine reported success (273,327–332). Nebulized morphine is not effective (333–335). Morphine should be administered with caution in patients with no prior history of opioid exposure. Increased levels of PCO2 were found following continuous morphine infusion in opioid naïve patients (331). Studies of morphine use in patients with prior histories of opioid exposure did not result in increased PCO2 (38,327–331). At present, there are no definitive guidelines for dosing opioids to treat dyspnea. One report found 25% of patients’ usual 4-hourly dose was sufficient to significantly reduce the intensity of dyspnea (327).

Benzodiazepines are routinely used for the symptomatic relief of dyspnea despite convincing evidence of their efficacy. Four of five controlled studies evaluating the effect of benzodiazepines for the treatment of COPD- and exercise-induced dyspnea failed to show any benefit (336–340). The use of benzodiazepines is warranted in the management of dyspnea associated with anxiety or a panic disorder. Otherwise, current literature does not support their use. However, the combination of morphine and midazolam alleviated dyspnea to a significantly greater degree than either morphine or midazolam alone (341).

Corticosteroids are commonly used to treat dyspnea associated with cancerous infiltration of pulmonary lymphatics and superior vena cava syndrome. Further, they have a long history of reducing bronchospasm associated with asthma and COPD (342,343). A trial of systemic or inhaled corticosteroids may be warranted if excessive mucus production or inflammatory changes in the airway contribute to dyspnea (298). The cost-to-benefit ratio of corticosteroid therapy should be carefully evaluated given their potential adverse impact on diaphragmatic function, particularly in cachectic patients.

Bronchodilators, both orally administered and nebulized, have proven to be extremely useful in the management of bronchospasm associated with asthma and COPD (342). Many cancer patients with dyspnea have a history of smoking or COPD, and dyspnea may be associated with airflow obstruction in patients with lung cancer (344). A trial of bronchodilator therapy is warranted when airflow obstruction may be playing a role in dyspnea. Aminophylline, theophylline, and caffeine improve diaphragmatic contractility in healthy individuals (345). Cachectic patients with generalized muscle weakness may benefit from xanthishes.

Nebulized furosemide has been proposed as a treatment for dyspnea in patients with advanced cancer (346,347). However, an isolated trial detected mild worsening of dyspnea in patients receiving furosemide (347). In the absence of
further evidence, nebulized furosemide cannot be endorsed as a treatment option.

General supportive measures can significantly enhance functional status and relieve distress in dyspneic patients. Patients able to tolerate aerobic conditioning may experience enhanced functional capacity, although the benefits of exercise remain largely theoretical in the palliative setting. Clearly, exercise interventions must be conducted in a manner sensitive to clinical context, prognosis, and patient preference. Pacing strategies and provision with assistive devices can reduce the work involved in activities of daily living, thereby affording dyspneic patients increasing autonomy and comfort. The benefits of continuous and intermittent positive pressure ventilation for patients with COPD, CHF, and neuromuscular respiratory dysfunction are discussed elsewhere in this text.

Patient and family education can have a powerful impact in reducing anxiety. Patients often fear dying of asphyxiation. Provision with a plan for emergency intervention may assuage, if not eliminate, this fear. Nursing-led interventions emphasizing coping with the psychological aspects of breathlessness reduce dyspnea (348–350). The benign nature of such interventions and their consistent efficacy argues that they should be integral to the palliation of dyspnea.

**HICCUPS AND COUGH**

Uncontrolled hiccups and cough can demoralize patients, and interfere with sleep, hydration, and nutrition. If persistent, these symptoms can render patients exhausted, malnourished, and vulnerable to medical morbidities. Hiccups, or singultus, are defined as repeated, involuntary contractions of one or both sides of the diaphragm (generally unilateral) and the intercostal muscles, terminated by abrupt closure of the glottis (351). This produces the characteristic sound of “hic.” In contrast to hiccups, which have no known purpose, cough is a natural defense of the body to prevent entry of foreign material into the respiratory tract.

**Epidemiology**

The prevalence of hiccups in different patient populations in the advanced stages of disease has not yet been characterized. Cough has been anecdotally described as a distressing symptom in patients with CHF, uncontrolled asthma, AIDS, and various cancers. A systematic review of symptom prevalence in patients with incurable illnesses revealed a cumulative cough prevalence of 28% among 12,000 patients in 24 studies (207). In a series of 289 patients with non-small cell lung cancer, cough afflicted more than 60% of the patients at presentation and prevalence increased to 80% before death (283).

**Pathophysiology**

The physiology of hiccups is poorly understood. A proposed hiccup reflex arc consists of afferent impulses transmitted via the phrenic or vagus nerves or the sympathetic ganglia, with efferent signals transmitted via the phrenic nerve to the glottis, diaphragm, and accessory muscles of respiration (351). Connections between the limbs occur at the levels of the spinal cord, a possible hiccup center in the cervical cord between C3 and C5, and higher CNS centers (352,353). Irritation, inflammation, or distension of viscera, primarily esophagus, liver, and stomach may trigger hiccups. Toxic and metabolic factors may also initiate and sustain hiccups including uremia, hyponatremia, hypokalemia, hypocalcemia, ethanol, and some chemotherapeutic agents (351).

Cough is similarly generated through a reflex arc that produces respiratory flow rates high enough to dislodge mucus and foreign substances from the lining of the airways. The cough reflex consists of afferent fibers housed within the vagus, trigeminal, and phrenic nerves. These fibers carry impulses generated by “cough” receptors at the termini of Aδ- and C-fibers found within the larynx, respiratory epithelium, tympanic membranes, esophagus, pericardium, and sinus mucosa. As with nociceptive afferents, Aδ-fibers convey impulses produced by mechanical stimuli, while C-fibers convey impulses produced by chemical stimuli (354,355). Inflammatory neuromodulators, such as prostaglandins, may influence receptor thresholds for depolarization (356). The impulses travel to cough centers within the medulla. Efferent impulses travel via the vagus, phrenic, and spinal nerves to the glottis and respiratory muscles (356,357). The resultant cough occurs in three phases: first an initial inspiratory phase, second a compressive phase, and last an expulsive phase (358).

**Causes**

In theory, many pathophysiological processes may engender hiccups with the most common being an overdistended stomach (359). Other common causes in the palliative setting include esophagogastric irritation, ileus, and peridiasphragmatic metastases. Causes of hiccups associated with advanced disease are listed in Table 69-5.

Cough often results secondary to local effects of the disease process. Tumors can irritate adjacent thoracic, abdominal, and neurologic structures to produce cough as occurs with endobronchial tumor, pericardial disease/effusion, or vocal cord paralysis. Infection may physically strip away epithelium, exposing sensory nerves and increasing the sensitivity of these structures to mechanical and chemical stimuli (360). Patients with both cancer and noncancer illnesses frequently cough consequent to aspiration indicating a need for clinical evaluation (361). Upper airway cough syndrome due to rhinosinus conditions, asthma, and gastroesophageal reflux disease is the most common source of cough in the general population and may also afflict patients in the palliative setting (362,363).

**Assessment**

History from the hiccupping patient should include data on recent bowel movements, oral intake, medication changes, and any associated abdominal or epigastric symptoms. Physical examination should focus on evidence of gastric distention. Neurologic examination may identify potential malignant sources of neural irritation triggers for both hiccups and...
cough since symptoms are mediated by reflex arcs. Serologic evaluation for metabolic derangements must be considered when routine evaluations are unrevealing. Imaging can localize tumor compression involving the neural reflex arcs or intrathoracic and peridiaphragmatic structures.

Potentially useful evaluations specific to the patient with cough include chest or sinus x-rays, and spirometry with pre- and postbronchodilator and histamine challenge. Upper GI endoscopy and 24-hour esophageal pH monitoring may also be useful. Such tests can be highly distressing and possibly dangerous for patients with advanced disease. Therefore, their utility must be weighed against how the information will alter management. An empiric medication trial may be the more clinically appropriate and humane approach.

Symptom-specific questionnaires allow subjective assessment of patients’ experience and associated distress. Both the Leicester Cough Questionnaire (364) and the Cough Specific Quality of Life Questionnaire (365) yield valid and reproducible results.

**Treatment**

Proposed treatments of hiccups are extensive. Nonpharmacological strategies include simple respiratory maneuvers (e.g., breath holding, rebreathing into a bag, compression of the diaphragm), nasal and pharyngeal stimulation (e.g., pressure on the nose, inhalation of a stimulant, traction of the tongue), vagal stimulation (e.g., ocular compression, carotid massage), psychiatric treatment (e.g., behavioral therapy), gastric distension relief (e.g., fasting, use of an NG tube, lavage), and phrenic nerve disruption via anesthetic injection (353,359,366,367). The efficacy of these approaches has not been rigorously assessed in clinical trials.

More empirical data support pharmacological hiccups treatments but should always be combined with definitive, disease-oriented therapies if feasible. If tumor effects are suspected, cytoreductive therapies, for example, radiation, may achieve symptomatic relief. The medication classes commonly used to treat hiccups include phenothiazines (chlorpromazine, proclorperazine), prokinetic agents, and anticonvulsants. Medication choices should reflect patients’ prognoses and functional levels. Chlorpromazine is the only FDA-approved agent for hiccups with dosing as follows: 25 to 50 mg orally tid or qid, or parenteral infusion of 25 to 50 mg in 500 to 1,000 mL normal saline over several hours. Haloperidol, 2 to 5 mg orally or subcutaneously to load followed by 1 to 4 mg orally tid, may have other desirable symptom control properties (chlorpromazine, proclorperazine), butyrophenones (haloperidol), prokinetic agents, and anticonvulsants. Medication choices should reflect patients’ prognoses and functional levels.

**TABLE 69.5 Neurocognitive Rationale for and Against Supplemental Hydration in the Palliative Setting**

<table>
<thead>
<tr>
<th>CON</th>
<th>PRO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comatose patients do not experience symptom distress</td>
<td>Dehydration can lead to pre-renal azotemia, which in turn can lead to accumulation of drug metabolites leading to delirium, myoclonus, and seizures</td>
</tr>
<tr>
<td>Parenteral fluids may prolong dying</td>
<td>There is no evidence that fluids prolong the dying process</td>
</tr>
<tr>
<td>With less urine, there is less need to void and use catheters</td>
<td>Providing hydration can maintain the appearance of “doing something,” even though there may be no medical value, and thus ease family anxiety around the time of death</td>
</tr>
<tr>
<td>With less GI fluid, there can be less nausea and vomiting</td>
<td></td>
</tr>
<tr>
<td>With less respiratory tract secretions, there can be less cough and pulmonary edema</td>
<td></td>
</tr>
<tr>
<td>Dehydration can help reduce distressing edema or ascites</td>
<td></td>
</tr>
<tr>
<td>Dehydration may be a “natural” anesthetic to ease the dying process</td>
<td></td>
</tr>
<tr>
<td>Parenteral hydration can be uncomfortable (e.g., needles/catheters) and limit patient mobility</td>
<td></td>
</tr>
</tbody>
</table>

The anticonvulsants gabapentin (300 to 400 mg orally tid), phenytoin (200 mg IV push to load followed by 300 mg orally qd), and valproic acid (5 mg/kg orally, then increase by 250 mg/wk until hiccups resolve or subcutaneously every 12 hours) may be preferred for neurogenically mediated hiccups. In contrast, hiccups attributed to gastric distension may preferentially respond to metoclopramide (10 mg orally qid) (368,369). Baclofen administered 5 mg orally tid failed to eliminate hiccups but provided symptomatic relief in a blinded, randomized trial (370). Nifedipine (10 to 80 mg orally bid) represents a third-line agent that may be preferred in cases of dose-limiting CNS sedation. Agents from different classes may be cautiously combined given the potential for...
discussion of whether supplemental feeding and hydration are beyond clinical concerns. To be effective in mediating the discomfort and nutritional support may assume significance that extends beyond clinical concerns. To be effective in mediating the discussion of whether supplemental feeding and hydration are appropriate, a clinician must appreciate the subtleties that may influence communication. The ethics of providing artificial nutrition and hydration (ANH) have been fiercely debated in the palliative care literature. A thorough discussion of the pertinent issues is beyond the scope of this chapter. The interested reader is referred to several excellent position papers (384–387). Although the majority of consensus statements hold that ANH is an interventional therapy that can be refused like any other, some contenders maintain that ANH constitutes basic care and must be provided. The authors acknowledge that prognoses, comorbidities, and etiologies differ widely and may dictate futility of ANH. Hence, broad generalizations are of little use.

Empirical evidence supports the use of ANH to extend survival among patients: in a permanent vegetative state (388), with extreme short-bowel syndrome (389), with amyotrophic lateral sclerosis-related dysphagia (390), in the acute phase of stroke or head injury (391,392), receiving short-term critical care (393), and with proximal obstruction of the bowel (394).

**Epidemiology**
The majority of patients who die in the United States and Canada do so while receiving either supplemental nutrition or hydration (395). Data on the prevalence of weight loss and dehydration vary widely across disease states and stages. Weight loss is associated with greater medical morbidity and shorter survival for the majority of cancers. Detecting weight loss among dying patients can be confounded by the presence of anasarca. Nonetheless, it has been reported that nearly all cancer patients have significant weight loss at the time of death (204).

**Causes**
Sources of dehydration are common in terminally ill patients. Probable causes depend on whether the patient has hypoproteinemic, hypernatremic, or isotonic dehydration. Hyponatremic dehydration arises when sodium loss exceeds water loss and occurs most frequently with water, rather than food. Common causes include volume depletion, anorexia, nausea, and vomiting. Overuse of diuretics, salt-wasting renal conditions, third spacing, and adrenal insufficiency are also common sources of sodium loss (396). Hyperonic dehydration occurs if water losses are greater than sodium losses. Fever is a common cause, as well as water loss through lungs, skin, and insufficient oral intake. Isotonic dehydration results from balanced loss of water and sodium. It has been suggested that the majority of patients with advanced disease have this type of eunatremic dehydration (397).

Nutritional deficiency may arise from anorexia as previously discussed. In addition to suppressed appetite, altered mental status, malignant GI obstruction, immobility, dysphagia, odynophagia, and uncontrolled cough and hiccups can interfere with adequate nutritional intake.

**Assessment**
Clinical assessment for dehydration should include screening for mental status changes, thirst, oral/parenteral intake, urine...
output, and fluid loss. Patients may or may not experience thirst. On physical examination, evidence of reduced skin turgor, dry mouth, and postural hypotension should be sought. Laboratory evaluations for hematocrit, serum sodium, blood urea nitrogen, and creatinine should be obtained if the workup is contextually appropriate. Urine osmolarity may also be useful.

Evaluation for reduced nutritional intake should include questions regarding diminished appetite, pain with eating, choking or coughing with swallowing, early satiety, altered or diminished sense of taste, and regularity of bowel movements. Oral pain can be associated with thrush, graft versus host disease, osteoradionecrosis (in head and neck cancer patients), mucositis, osseous metastases to facial bones, herpes zoster infection, and neuropathic pain syndromes. Imaging studies may be required to confirm the presence of malignant involvement of the GI tract.

**Treatment**

There is no evidence to support enhanced comfort through the use of supplemental hydration (398). Clinical studies have produced mixed results (399,400). Proposed reasons for and against providing supplemental fluids related to control of neurocognitive symptoms are listed in Table 69-6 (401). Reduced intravascular volume can lead to the accumulation of opioid and other pharmacological byproducts, which may aggravate undesirable side effects such as confusion, sedation, myoclonus, and seizures. Dehydration can place patients at increased risk for bedsores, constipation, and xerostomia. However, there are equally compelling reasons to question the benefits of supplemental hydration (see Table 69-6).

A standard goal for fluid is 1,500 to 3,000 mL, or eight to ten glasses of water daily. Offering fluid at regular intervals merits a trial. Sports replacement fluids may be preferred by some patients and can correct hypotonic dehydration. Nonoral routes of fluid supplementation include parenteral administration, proctoclysis, and hypodermoclysis. Use of an NG tube for the introduction of enteral fluids is discouraged because patients may inadvertently extubate themselves, and NG tubes can be a source of considerable discomfort (402). Rehydration by proctoclysis involves the use of an NG tube placed rectally (398,402). Tap water or saline can be instilled at a rate of 100 to 400 cc/h. Treatment should be initiated at a rate of 50 to 100 cc/h and gradually increased with close monitoring for rectal leakage. Although this approach has been shown to be safe, economical, and effective, the majority of patients prefer hypodermoclysis. Hypodermoclysis refers to subcutaneous fluid infusion through needles inserted in the tissue of the abdomen or thigh (398,402). Hypotonic or isotonic solutions are generally used with or without hyaluronidase or corticosteroids. An infusion of 100 mL/h is usually well tolerated. Once 1,500 mL have been instilled, an alternate administration cite should be chosen.

Nutritional therapy is aimed at improving metabolic status, body composition, functional status, and QOL. The literature offers insufficient evidence for enhanced survival in terminally ill patients (403). Artificial nutrition can be delivered enterally or parenterally. Whenever possible, the enteral route of feeding is preferred. Slow initial rates of 10 to 20 mL/h should be used. If tolerated, rates can be increased by 10 mL every 8 to 12 hours until full flow rates are reached. For terminal patients incapable of enterally absorbing or assimilating nutrients, total parenteral nutrition can be prescribed. Whether the enteral or parenteral route is used, basal metabolic requirements can be calculated using the following Harris-Benedict equations (404).

\[
\begin{align*}
\text{BMR for men} &= 66.4730 + (13.7516W) + (5.0033H) - (6.7550A) \\
\text{BMR for women} &= 65.5095 + (9.563W) + (1.8496H) - (4.6756A)
\end{align*}
\]

where \(W\) denotes the weight, \(H\) the height, and \(A\) the age.

Ongoing vigilance must be maintained for evidence of infection in patients receiving total parenteral nutrition. High prevalence rates have been reported, particularly in patients with advanced disease (405,406).

**EXERCISE AS PALLIATION**

There are excellent reasons to believe that structured exercise may significantly benefit patients with advanced disease and decelerate their functional decline. First, exercise ameliorates a

---

**TABLE 69.6 Causes of Hiccups in Patients with Advanced Disease**

<table>
<thead>
<tr>
<th>Cause of Hiccups</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intracranial neoplasms (primary and metastases)</td>
</tr>
<tr>
<td>Uremia</td>
</tr>
<tr>
<td>Alcohol</td>
</tr>
<tr>
<td>Hyponatremia, hypokalemia, hypocalcemia</td>
</tr>
<tr>
<td>Fever</td>
</tr>
<tr>
<td>Diaphragmatic irritation (diaphragmatic tumors, pericarditis)</td>
</tr>
<tr>
<td>Pleuritis</td>
</tr>
<tr>
<td>Esophageal obstruction</td>
</tr>
<tr>
<td>Esophageal irritation (graft versus host disease, radiation induced)</td>
</tr>
<tr>
<td>Pericarditis</td>
</tr>
<tr>
<td>Hepatomegaly</td>
</tr>
<tr>
<td>Subphrenic abscess</td>
</tr>
<tr>
<td>Esophageal cancer</td>
</tr>
<tr>
<td>Mediastinal tumors</td>
</tr>
<tr>
<td>Herpes zoster</td>
</tr>
<tr>
<td>Gastric distention</td>
</tr>
<tr>
<td>Gastric cancer</td>
</tr>
<tr>
<td>Pancreatic cancer</td>
</tr>
<tr>
<td>Intra-abdominal abscess</td>
</tr>
<tr>
<td>Bowel obstruction</td>
</tr>
<tr>
<td>GI hemorrhage</td>
</tr>
<tr>
<td>Short-acting barbiturates</td>
</tr>
<tr>
<td>Dexamethasone</td>
</tr>
<tr>
<td>Diazepam, chloridiazepoxide</td>
</tr>
<tr>
<td>Infections (meningitis)</td>
</tr>
<tr>
<td>Grief reaction</td>
</tr>
<tr>
<td>Psychosis</td>
</tr>
</tbody>
</table>

---
variety of symptoms that constrain patients’ function including dyspnea and fatigue (407). Controlled trials of aerobic conditioning at 55% to 70% VO$_2$peak in cancer patients both who are receiving and who have completed chemotherapy demonstrate significant benefits not only in symptom burden but also psychological well-being and QOL (408–411). A strong association between exercise and enhanced mood has been reported in many disparate cohorts from a wide range of disease states (412–414). Benefits across multiple symptoms and QOL domains have been reported in cohorts of cancer and AIDS (415) patients, as well as those with advanced lung (416) and heart disease (417).

Despite the many compelling reasons to consider exercise in symptom alleviation, skepticism has been raised regarding the ability and willingness of patients with cancer and other types of advanced stage illness to consistently participate in exercise programs (418,419). The encouraging work of Yoshioka (11) suggests that even hospice patients can make substantial gains in mobility through conventional physical therapy (PT) approaches, yet the low-level mobility activities employed in PT are generally distinct from conditioning activities delivered for symptom alleviation. Limited pilot data suggest that patients with advanced cancer and limited life expectancy benefit from circuit training through enhanced physical performance and reduced fatigue (420). However, although the investigators selected patients with Karnofsky Performance Status more than 60 and estimated prognoses of 3 to 12 months, only 63% of patients approached agreed to participate and 54% of participant were able to complete the study (421). A literature review demonstrated that disappointing, approximately 50%, uptake and completion rates characterize many exercise studies targeting cancer populations, not solely those with advanced disease (422). Neuromuscular electrical stimulation has been suggested as a possible alternative to active participation in exercise protocols as a means of preserving functionality in patients with advanced cancer (421). Supportive data are limited to a promising case study (423).

Efforts to assess receptivity to function-oriented therapies among patients with metastatic cancer suggest that interest is high and correlates with self- and clinician-rated functional limitations (11,424,425). Various barriers have been identified as constraining patients’ ability to actualize their interest and begin regular structured exercise including transportation, cost, winter weather, and the demands of cancer treatment (425,426). Clark et al. reported that fatigue was the most frequently listed barrier to physical activity among patients with advanced cancer receiving chemotherapy (425).

The importance of structure and support is suggested by reduced adherence levels following termination of experimental exercise protocols in cancer patients (427,428). It is reasonable to assume the patients with advanced diseases may be in considerably greater need for ongoing support and supervision in order to adhere to exercise. Thus far, creative approaches to providing structure for exercise that avoid heavy time and travel burdens have not been developed. Since patients identify time as a significant barrier to exercise, home-based programs may be a promising option for future development (425). One study failed to detect a difference between exercise interventions delivered in supervised and self-directed fashions (429).

CONCLUSION

One factor that may foster the integration of rehabilitation and palliative services is the striking parallel between their general therapeutic paradigms. Both disciplines encourage creative problem solving, whether for functional, psychological, or physical difficulties in the interest of improved function and QOL. The desire to help patients live well to the fullest of their abilities galvanizes palliative and rehabilitation teams in a unified purpose. Both disciplines acknowledge the importance of a holistic integration of disparate disciplines so that the unique case-by-case challenges may be met. It is hoped that this common ground will serve as a foundation for building a body of clinical and academic rehabilitative expertise unique to the palliative patient.

REFERENCES


329. Mazocchi C, Buelin T, Rapin CH. The effects of morphine on dyspnea and ventilatory function in elderly patients with advanced...


Biofeedback (BF) can be defined as the use of instrumentation to reveal covert physiological processes via user detectable cues, such as visible light and audible tone, for appropriate response shaping. The BF loop consists of a BF machine and user, and as necessary an instructor/trainer/clinician. Ultimately, the machine functions as an extension of human sense to record and display the internal physiological signals or events; the user, armed with feedback information, acquires the skills to control the physiological response toward desired state through mind-body self-regulation.

Biofeedback became recognized as an alternative or adjunct medical tool in the 1960s and has been applied to psychotherapy, physical medicine, incontinence treatment, pain management, and more recently, in the management of other behaviors associated with pediatrics and oncology. This presentation is directed toward rehabilitation medicine in general and the control of movement in particular.

In physical medicine and rehabilitation, the initial BF application arose from diagnostic and research work in electromyography (EMG) (1), which became the most practical BF form in the field. From the 1960s to the 1990s, most clinical and experimental BF applications used EMG, joint angle, position, force, or pressure to reeducate the control of muscles, joint, and balance in patients with various neuromotor deficits. The outcomes provided concrete evidence that objective neurological signs and symptoms can be altered, particularly in patients with upper motor neuron paralysis and spasticity resulting from brain damage (2–5). However, the initial BF applications did not correlate well to patients’ motor function improvement. This observation may be partly due to the early “static” BF training paradigm, in which patients received BF training in a static posture or during nonfunctional related movement. Since the mid-1990s, several experimental studies, inspired by the concept of task-oriented training for motor functional recovery, have transformed BF interventions into functional task training that might employ feedback principles (6–11). In addition, recent technological advances have further promoted this new direction. The emergence of novel sensors, advanced signal processing and control, remote communication, and three-dimensional (3D) displays in BF applications will further leverage the influence of BF therapy in physical medicine and rehabilitation.

The purposes of this chapter are to (a) provide a general review of the early experimental and clinical BF applications in physical medicine and rehabilitation and (b) introduce recent developments in rehabilitation technologies and applications toward task-oriented BF interventions. Readers are encouraged to keep in mind that some of these new technologies and applications are still in their infancy; continuous efforts are demanded to establish effective training protocols and outcome measurements and in evaluating these new BF technologies and interventions before reaching definitive conclusions.

TRADITIONAL BIOFEEDBACK IN REHABILITATION

Methods of Biofeedback
Electromyographic Biofeedback
The electromyographic signal or EMG is an electrical manifestation of muscle activity and an effective window to inspect neuromuscular control system. In BF retraining, EMG is the most used form to down-train hyperactive muscles or up-train flaccid or weak muscles in patients with various sensorimotor deficits (12–14), thus further improving patients’ control over joint. Usually, bipolar surface EMG electrodes are placed on one or two targeted muscles. The sensed signals are digitally sampled at 1,000 Hz or higher rate. The raw signal, the integrated EMG, or the frequency of EMG is then translated into simple acoustic and visual signals (e.g., lights and audio cues) or graphic computer displays (Fig. 70-1). Patients receive the feedback in a quiet environment and mostly in a static posture. Noise is held to a minimum, and visual distractions avoided. Special uses of intramuscular EMG (IMG) BF include attempts at training deep inaccessible muscles, paralytic muscles, muscles separated from the skin by considerable adipose tissue, or muscles that are not easily isolated by surface electrodes. IMG signals are commonly recorded by invasive, indwelling fine-wire or concentric needles and sampled at 10 kHz or higher rate.

Joint Angle Biofeedback
Joint angle BF can be efficient for improving joint movement control (15), even more than EMGBF (5). When the active joint motion is presented but limited in patients with neuromotor deficits, compared to EMGBF, joint angle BF might
be promising for effective and expeditious recovery of joint control. In addition, angle BF is indicated when the goal of training is the regulation of joint movement, such as correction of genu recurvatum (16,17) or the control of movement with appropriate timing and coordination (5,6). Moreover, joint angle BF may be used when the muscle that must be monitored is inaccessible or difficult to isolate. Electrogoniometers reliably reproduced the clinical measurements of joint angle (11); they have been widely employed in angle BF devices. Other applied sensors include mercury tilt switches and gyroscopes. The quantified joint angle is fed back to the patient during single joint movement for targeted angle tracking (15), or during multijoint coordinated movement, such as gait (6,11).

**Pressure or Force Biofeedback**

Force or pressure monitoring may be indicated when information concerning the amount of force being transmitted through a body segment or assistive device is desired. Force/pressure sensitive platforms whose applications are described in more detail later in this chapter (see “Balance Rehabilitation”) are often used for retraining of balance (18). As shown in Figure 70-2, a patient with balance control deficits stands on a force plate that measures the ground reaction force or pressure and/or moments in three orthogonal directions under the feet. The derivative of force or pressure measurements such as center of force (COF) (18), center of pressure (COP) (19–22), or center of gravity (COG) (23–25) is displayed as a cursor projected on a two-dimensional (2D) screen in front of patients. The goal of the patients is to move the cursor to a desired location or within a targeted area (18). Some commercial force platforms are also equipped with motors that can translate or tilt the platform for balance perturbation (26,27). Force BF training under conditions of posture perturbation have been applied to permit older adults at risk of fall to experience strategies to abort falls (14,24,28). Most of applications feedback the COP cursor instantaneously and such training has been reported to induce both favorable (the amplitudes of the horizontal motion of the COG are diminished) and unfavorable features (the vertical difference between the COG motion and the COP trajectories are enhanced) for reacquisition of balance control (20). Findings from one study suggested that delayed visual feedback of the COP cursor exceeding 600 ms can significantly suppress the unfavorable features and retain the beneficial aspect of COP BF training (20).

Recent sensor technologies, such as insole pressure sensors (29) and portable pressure walkways (30), allow force/pressure BF training in ambulation. All these devices can display the pressure distribution and the COP in real time. In addition to the force measurement using force plates, the loading force or pressure measured by load cells in assistive devices such as canes (31) or prostheses (32) has been used to train patients’ bodyweight loading on the impaired limbs during ambulation.

**Miscellaneous Techniques**

Beyond EMG, joint angle, and force, many other parameters have been monitored for miscellaneous BF applications in
neuromotor rehabilitation. For example, BF of step length (7), knee-to-knee distance (8), and step timing (9) were applied to correct abnormal gait pattern. Trunk acceleration was the BF parameter for stance balance training (33). The applied BF equipment ranged from a simple mirror to expensive motion tracking cameras.

Some studies conducted BF training using multiple data sources. For example, studies coupled EMGBF with joint angle BF and trained stroke patients to control the joint to a desired position by increasing the recruitment of agonist and/or reducing the muscle activity from antagonist (6,14). The risk in using multisource BF is that multiple quantified feedback cues might overload patients' perception and confuse patients; careful design of feedback cue displays is essential for successful BF applications.

Biofeedback Modalities

Visual Feedback

Visual displays available with BF devices include banks of lights, liquid crystal display (LCD), meters, oscilloscopes, or computer monitors. The visual display can be binary (0/1 in value or light on/off), digital (integral numbers), or continuous (signal waves or value bar).

The sensitivity scale in the visual display should be determined by the goals for the BF training. In EMGBF, if the EMG electrodes are properly applied, it is theoretically possible to obtain accurate readings of activity as low as 1 µV, but such low integrated levels are important only during generalized relaxation training. Retraining patients with musculoskeletal weakness (e.g., the quadriceps after meniscectomy) requires a feedback device with comparatively low sensitivity settings because such patients often learn to generate several hundred microvolts of integrated EMG during isometric or isokinetic contractions. Most contemporary feedback displays have software that adjusts the range of sensitivities as the patient changes his or her muscle output capability with an intent of providing a continuous range that can be visualized and manipulated by the patient within his or her training sessions.

Caution must be exercised when trainees have visual deficits secondary to brain injury. Furthermore, when the BF training involves activities such as gait, visual feedback should be avoided or kept minimal because vision is largely occupied to guide motor coordination.

Audio Feedback

Many commercially available devices offer auditory feedback in the form of a tone, buzzer, click, or a combination of these possibilities. Similar to visual feedback, the audible feedback could be binary, discrete, or continuous. In devices with binary display, a monotone buzzer is heard only when the patient achieves a specific feedback value preset by the therapist. In EMGBF, a low threshold setting may be used in training for recruitment of activity above a given level in a weak or paretic muscle. Once the patient reliably exceeds this level, the threshold is raised progressively. This technique is often referred to as shaping. The reverse shaping strategy to reduce integrated EMG levels (i.e., reduction of resting hypertonus) also may be used. Additionally, binary auditory feedback is useful for BF therapy set in the activities of daily living (ADL); the auditory tone is given only if needed so that users can concentrate on the daily activities most of the time. The device with discrete audible feedback maps more than two physiological states into sounds with different pitch, duration, or loudness. The continuous auditory feedback directly displays the sampled physiological signal. For example, surface EMG was directly transformed into a sound. When the muscle is activated, the audio components increase in intensity and pitch, which can be resolved by patients as effective feedback required to regulate the muscle activation level.

A number of studies compared the influence of auditory display to that of visual feedback during BF training. The auditory cue was reported to have more patient acceptance than the visual feedback when the BF training was carried out in ambulation (11). Another study found that audio BF had larger effects on reducing COP displacement while visual BF had larger effects on reducing trunk sway, which suggests that the two BF modalities may encourage a different type of posture sway strategy (34).

Tactile Feedback

Applied tactile sensation arises from a simple mechanical vibrating stimulator attached to the skin. By modulating the vibration frequency and amplitude, the vibrating stimulator feeds back the sensed physiological signal to the user (35,36). Vibrotactile feedback is safe, frees the user from having to maintain visual attention to the feedback cues, and presents minimum distraction to others. However, results from previous studies indicate that while the user could easily adapt to the tactile stimulation, the tactile biofeedback cues were ignored as participants engaged in ADL (36). These observations may have resulted from the relatively smaller number of reported studies using tactile BF compared to the use of visual or auditory cues.

Scientific Basis of Biofeedback

The mechanisms underlying successful use of BF are still unclear. In physical rehabilitation, BF may enhance sensorimotor integration because this approach highlights utilization of sensory cues that inform patients about consequences of their movements while allowing them to develop adaptive strategies for motor learning and recovery.

Neurological Basis of Biofeedback

Wolf and Binder-MacLeod (37) have asked, “How do patients process feedback information? What factors account for the patient’s ability to gradually become less dependent upon artificially induced information signals after failing to achieve significant gains using conventional rehabilitation procedures?” Previous studies demonstrated that the BF therapy was associated with cortical reorganization (38,39). Specifically, fMRI studies have shown that following biofeedback training to enhance ambulation following stroke, enhanced activation during
controlled knee flexion extension was seen in the ipsilesional primary sensorimotor cortex. The extent to which such changes are manifest with biofeedback neuromuscular reeducation of different joints and for different diagnoses requires considerable investigation. The two possibilities suggested many years ago by Basmajian (12), either an auxiliary feedback loop to recruit existing cerebral and spinal pathways or the development of new pathways, are worthy candidates for further study.

Nonetheless, the central nervous system uses a multitude of internal modulatory networks. The role of somatosensory or other subcortical areas, and the basal ganglia should not be underemphasized; their timely activation make voluntary movements meaningful. Proprioception, tactile, auditory, and visual inputs are harmonized into the controls, as is the cerebellum. Therefore, damage resulting from external and internal trauma to area 4 of the cerebral cortex may spare pathways that are primarily engaged for reacquisition of skilled movements, or they may spare redundant pathways that EMGBF training can bring into play. Wolf (40) has further speculated three detailed conceptualizations: (a) override: visual or audio feedback may activate the somatosensory cortex by entering at a level higher than the level of damage; (b) bypass: an appropriate feedforward system can be established via the brain stem motor nuclei; (c) repetition with existing neural circuitry: central synapses previously unused in executing motor commands may be activated by visual and audio feedback. As such, continued training may establish new sensory engrams and help patients perform tasks without feedback.

Plasticity and relearning by the brain should be thought of as a more complex phenomenon than usurping function by specialized areas of the neighboring cortex. There is evidence from rat stroke models that changes in dendritic architecture or density of synaptic boutons are associated with motor relearning, provided that the task practice is repetitive and challenging (41,42). Whether the BF training employs and in turn reinforces these new neural connections is unclear.

Placebo Effect
Basmajian (43) addressed placebo factors that are powerful in all therapies, including BF alleged “cures.” Therefore, the placebo effect needs to be understood, accepted wisely and gratefully, and rebuffed. In fact, many clinical studies, especially those based upon pharmacological interventions, question the impact of the intervention if improvements are not in excess of 30% beyond control values or have not been powered adequately to determine a “minimally clinical important difference” (44,45). A deliberate placebo may be best used in double-blind research studies, or rarely as a therapeutic test.

Biofeedback Applications in Physical Medicine and Rehabilitation
Stroke Rehabilitation
A major application of biofeedback in rehabilitation hospitals and outpatient clinics lies in the treatment of patients following stroke. The National Institute for Health (46) lists stroke as the number one cause of adult disability in the United States. Approximately 780,000 people experience a new or recurrent stroke each year (47). Motor dysfunction after stroke may be characterized by muscle weakness, abnormal muscle tone, abnormal movement synergies, and lack of coordination during voluntary movement. Restoring the neuromuscular control in patients with stroke is essential to further improve their motor functions. EMGBF is a useful tool for reeducating neuromuscular control in stroke rehabilitation (3,4,48–52). Other forms of BF, such as force and joint angle BF, have also been applied, but with less frequent use.

EMG Biofeedback for Upper- and Lower-Extremity Rehabilitation
Wolf et al. (37,53) investigated patient characteristics that may affect the outcome of EMGBF training of hemiplegic patients. The presence of proprioceptive loss appeared to diminish the probability of making functional gains in the upper limb. Age, gender, hemiparetic side, duration of stroke, previous rehabilitation, and number of training sessions did not have a significant effect; patients should not be excluded from feedback training based solely on either age or length of time since their stroke. However, the optimal time for introducing BF appears to be the same as for any rehabilitation efforts: fairly early.

In EMGBF, a practically standard training scheme has been adopted. Detailed EMGBF training protocols for stroke rehabilitation can be found elsewhere (54). Scientists at Emory University have conducted a series of studies to implement EMGBF for stroke rehabilitation. The belief that neurological patients should first have hyperactive muscle down-trained motivated the application of EMGBF to spastic muscles initially. Relaxation training was applied on patients when they were at rest, during passive movements, during distractive movements, and then during shortening contractions. Once the muscles were relaxed, attention was directed toward the antagonist muscles, the ones that are usually weak and need to be up-trained in EMGBF. Last, the coordination of both muscle groups must be trained for efficient joint control. Ultimately, this BF training paradigm was put into a functional context. While this paradigm did yield significant improvements (3,4,55,56), the greatest predictor for ultimate success in applying EMGBF was in the patients’ abilities to demonstrate small amounts of active extension at the elbow, wrist, and fingers (14). In fact, these observations served as the inclusion criteria for the first forced use and constraint induced movement therapy studies (57–60).

EMGBF for upper-extremity rehabilitation starts from the shoulder joint, followed by elbow, wrist, and hand (54). EMGBF training of pronation and supination of the forearm is difficult because of the cross talk between the EMG activity from the pronators and supinators and other EMG present in the forearm; EMGBF may be combined with angular BF in cases of severe spasticity or flaccidity. Targeted training of the lower limb is simpler than that of the upper limb. Training of the lower limb need not follow the proximal-to-distal progression for the upper limb. The primary functional goals are improved ambulation and to develop a relatively limited
number of stereotyped patterns used during ambulation. One of the important areas of EMGBF training was for the treatment of foot drop caused by paralysis of the ankle dorsiflexors and spasticity of the plantar flexors (2,61). Other training protocols involve training of multiple joints simultaneously to coordinate, such as the training of hip and knee extension critical for the gait stance phase, hip flexion with knee extension important during the terminal swing phase of gait, and hip extension with knee flexion (54).

The current view that the task-oriented repetitive training is essential for motor functional recovery motivates clinicians and scientists to conduct BF training when patients are walking (6–8,17,62–64). BF retraining during ambulation focuses on specific gait timing, because the coordination of muscles or joints depends on gait phase. In this case, a gait event detection system such as footswitch is essential (6). Continuously monitoring the gait performance, such as gait symmetry, weight loading, or muscle recruitment has also been reported (7,62). Auditory warning buzz provides discrete feedback only if the monitored parameter is out of the desired range. The clinical effect of BF training for stroke rehabilitation is still controversial, the detail of which is summarized later in this chapter (see "Randomized Controlled Trials and Meta-Analyses").

**Balance Rehabilitation**

Another major application of BF in stroke rehabilitation is the training of balance control (23,25,65–68). Roughly 40% of stroke patients will experience a serious fall within a year after having a stroke (69). Unsteadiness during stance, asymmetric weight loading, and decreased ability to move within a weight-bearing posture have been reported among stroke survivors. Nichols gave a thorough review and update on force platform BF for balance retraining (18). The training protocols address three components of the function: steadiness, symmetry, and dynamic stability. In retraining of postural steadiness, stroke patients stand on a force plate, wearing a fall arrest harness. Patients are required to keep the cursor representing COF or COP within a narrow range while they sway the body weight. To improve the control of postural symmetry, the force BF training progresses from static standing to dynamic movements, such as sit-to-stand transfers and stepping in place. The percentage of weight bearing on the nonparaparetic and paraparetic leg is quantified and displayed visually and audibly (19,65,70). The goal of trainees is to equalize the weight loading on each leg. The training of postural steadiness and symmetry resulted in the reduction in weight-bearing asymmetry (65,70–72); however, stroke patients failed to reduce their spontaneous sway amplitude and did not improve on functional measures (65,70). The dynamic postural control will be last trained. The patients are instructed to voluntarily move the COP/COF cursor from one target to another in different directions accurately without falling (21,23). The capability to transfer body weight and adopt a different stance position is a prerequisite for safe mobility (73). Hence, training of dynamic stability is thought to have important links to the function (18), but no strong evidence has been found to support this contention.

**Spinal Cord Injuries**

Spinal cord injury (SCI) results in varying degrees of weakness and sensory loss at and below (i.e., caudal to) the site of injury. The motor deficits may be represented as muscle weakness and limb paralysis, spasticity, lost of normal bladder control, and breathing difficulty.

The primary goals for interfacing patients with SCI with EMGBF are much the same as those outlined previously for stroke patients. First, attempts are made to reduce hypermotor responses to induced length changes in spastic muscles. Such hyperactive behavior of spastic muscles may occur during spontaneous episodes of clonus or during induced clonic seizures, when the lower or upper extremity responds to various tactile stimuli. Once the patient can reduce such responses in supine, sitting, and ultimately standing postures, efforts are directed toward recruitment of weak muscles.

In patients with paraparesis following SCI, EMGBF training at the Emory University enabled many patients to increase the speed at which they ambulate and to reduce the number of required assistive devices. For patients with tetraparesis and obvious residual voluntary movement, feedback combined with an exercise program facilitated active range of motion (ROM) and improved upper-extremity function. Feedback may be beneficial for these patients because the modality may be easily incorporated into exercise programs with immobilized patients during the acute phase of injury; it provides immediate information to the patient concerning the level of voluntary muscle activity; and, by so doing, this modality may help patients obtain spatial and temporal summation of muscle potentials leading toward increased contractility, and therefore preparing the patient for a more vigorous therapy program. However, in cases in which little activity could be observed among chronic SCI patients, feedback provided little or no significant benefit toward restoration of function (74). Other clinical trials demonstrate controversial results on the effects of EMGBF for patients with chronic, tetraparesis secondary to SCI. One study using a pre-post experimental design found that EMGBF was helpful in increasing voluntary muscle recruitment in long-term spinal injured patients (75), while another similar study found no significant improvement (76).

In addition to the regular EMGBF training described above, the study groups at the Emory University have conducted a series of studies to investigate whether EMGBF can be applied to reduce the spinal stretch reflex in healthy and neurological patients (77–79). After 2-week training, all the SCI patients exhibited significantly reduced spinal stretch reflex. However, the value of this observation in SCI rehabilitation and clinical application requires further investigations.

**Cerebral Palsy and Traumatic Brain Injuries**

Early EMGBF applications monitored the spastic muscles in patients with cerebral palsy (CP) or traumatic brain injury (TBI) (80–84). In the study of Neilson and McCaughey (80), a progressive muscle relaxation procedure was adopted for the biceps brachii muscle. EMGBF was employed to down-train the sensitivity of tonic stretch reflex. Four young
Adults with CP reduced involuntary muscle activity and stretch reflex sensitivity. However, only one athetotic patient improved voluntary joint control as a consequence of reducing the amount of involuntary arm movement. EMGBF was also applied to down-train the gain of stretch reflex for correction of muscle contracture (81,82). Although patients with CP significantly decreased the stretch reflex, the contracture was not altered (82); thus, EMGBF might only be useful for preventing the progress of muscle contracture.

With the view that the deficits of motor function in CP patients are more related to deficits in strength and control than spasticity (85), the recent applications of EMGBF shifted attention to the up-training of muscle activity, muscle synergy, and joint control. A successful application of EMGBF in patients with CP is seen in the retraining of ankle control (86–89). EMG from either plantarflexor (86,89) or dorsiflexors (87,88) was measured and displayed. Patients were trained to increase the muscle activity or to improve the muscle or joint control during gait cycle. The results of these studies were encouraging: the voluntary muscle control, the ankle active ROM, and the functional gait parameters were positively affected by EMGBF training. Bolek (90) developed a multiple-site performance-contingent feedback to treat motor dysfunctions in patients with CP. The applied EMG sites were customized to each patient who learned the synergy of multiple muscles for improved joint control. Fourteen of the sixteen children with CP improved the control of multiple muscle recruitment patterns during the treatment.

Head position control using positional BF (91,92), control of drooling using EMGBF (93), and trunk sitting posture control by angular (36) or pressure BF (94) have been used with CP patients, but formal studies are sparse.

**Multiple Sclerosis**

Multiple sclerosis (MS) causes a variety of sensorimotor dysfunction, including muscle weakness, abnormal muscle tone, difficulties in coordination and balance, problem in speech and swallowing, fatigue, change in sensation, and bladder and bowel control difficulties (95,96). Involuntary muscle activity was measured and displayed. Patients were trained to prevent involuntary muscle activity and to improve the muscle or joint control during gait cycle. The results of these studies were encouraging: the voluntary muscle control, the ankle active ROM, and the functional gait parameters were positively affected by EMGBF training. Bolek (90) developed a multiple-site performance-contingent feedback to treat motor dysfunctions in patients with CP. The applied EMG sites were customized to each patient who learned the synergy of multiple muscles for improved joint control. Fourteen of the sixteen children with CP improved the control of multiple muscle recruitment patterns during the treatment.

Head position control using positional BF (91,92), control of drooling using EMGBF (93), and trunk sitting posture control by angular (36) or pressure BF (94) have been used with CP patients, but formal studies are sparse.

**Dystonias and Dyskinesias**

Ignoring the many conflicting theories of etiology that provide no clear guides to therapy, we briefly discuss here a number of movement disorders that have responded well to behavioral therapy featuring EMGBF. They present themselves mainly in isolated muscle groups; spasmodic torticollis is the best example (103–106), but also included are the rarer blepharospasm (107–109), hemifacial spasm (110), writer’s cramp (111–115), and severe torsional dystonias of the torso (i.e., malignant dystonia musculorum deformans). Many brief case studies have been reported. The most thorough studies of many patients successfully treated with EMGBF combined with other behavioral techniques were reported by Brudny et al. (116) and by Cleeland (103).

**Peripheral Nerve Denervation**

Facial palsy causes weakness or paralysis of the facial muscles, accompanied by other complications (117). Synkinesis is one of the complications and is an abnormal involuntary associated movement during blinking. The underlying mechanism of synkinesis is inappropriate reinnervation of the regenerating facial nerve fiber to the facial muscle (118).

EMGBF is an efficient rehabilitation intervention for patients with facial palsy (118–125). Results from studies have shown that after EMGBF training, there was substantial improvement in facial symmetry and voluntary functions. Moreover, EMGBF was reported to be useful to further improve the facial function in patients undergoing facial anastomosis surgery (123–125).

**Polyneuropathy and Peripheral Neuropathy**

Patients with diabetes might develop sensory neuropathy that compromises proprioception. BF devices function as a bypass to close the motor control loop for sensorimotor integration (24,126). Patients with sensory neuropathy demonstrated dysfunction on balance control and increased rates of fall (127). A controlled trial tested BF of COG for balance retraining against conventional therapy and found that BF is more efficacious (24).

Walker et al. (126) developed a pressure BF device with audio feedback, which was applied to warn the excessive pressure on the user’s foot and therefore prevent neuropathic ulcer.
Pain Management and Biofeedback
Related disciplines, especially psychotherapy and pain clinics, widely use EMGBF for relaxation, both general and specific (54). In addition, skin temperature feedback training is used for treating pain conditions normally not seen by rehabilitation clinicians (e.g., migraine) with generally good results (54). While these approaches are exciting, the voluminous literature available is overwhelming and beyond the scope of this chapter. Interested readers are encouraged to explore some contemporary but excellent references (128,129).

Acute and chronic back pain treatment with targeted surface EMGBF, as an adjunct to conventional excises, has shown positive effects on the strength of lumbar paraspinal muscles (130,131). Recently, advances in the ultrasound technique make the noninvasive measurement of deep muscle activity in real time possible. Studies employed ultrasound image-based BF to down-retrain the deep lumbar muscle of patients with patients post back pain and found promising effects of the intervention (132,133).

Orthopedic and Prosthetic Rehabilitation
Joint angle BF devices have been successfully used in retraining of hand function after reparative surgery of traumatized fingers and their motor nerves (134). The feedback goniometers have been developed that are worn on the hand during exercise and therapeutic activities. These devices supply a threshold feedback signal when a predetermined joint angle is attained. Clinical trials indicate that many patients in a comprehensive hand rehabilitation program after injuries or corrective surgery will improve somewhat more in active ROM if they use feedback goniometers while exercising. Similar joint angle BF goniometers have also been used for training knee ROM in patients post total knee arthroplasty (11).

EMGBF can be helpful in the treatment of musculoskeletal disorders (135–137). For example, a controlled study demonstrated that BF alone is more effective than electrical stimulation alone in the recovery of the peak torque or force of the lower back muscles (130,131). Recently, advances in the ultrasound technique make the noninvasive measurement of deep muscle activity in real time possible. Studies employed ultrasound image-based BF to down-retrain the deep lumbar muscle of patients with patients post back pain and found promising effects of the intervention (132,133).

Randomized Controlled Trials and Meta-Analyses
Randomized controlled trials (RCTs) are the gold standard prerequisite to obtain a statistically validated conclusion because RCTs eliminate all forms of spurious causality (40). Previously conducted controlled trials tested BF plus conventional physical therapy (CPT) against CPT alone (4). For chronic stroke patients, EMGBF, when used as an adjunct to CPT, resulted in improvement in upper-limb ROM and muscle strength. The authors emphasized that EMGBF is an adjunctive and not a total therapy.

Meta-analyses of BF studies by three separate study groups have yielded conflicting conclusions, ranging from “effective tool for neuromuscular reeducation in the hemiplegic stroke patient” (138) to “[do] not conclusively indicate superiority” (139) and “do not support the efficacy of BF in restoring the ROM of hemiparetic joints” (140), with a caution that a type II error may be masking “an important clinical benefit.” As is the case with so many meta-analyses, the reason for the diverse results might reside in use of different criteria for data inclusion and incongruities in the specification of end point measurements. Schleenbaker and Mainous (138) included studies with a non-RCT, while other studies only considered RCTs (139–141). Some meta-analyses selected studies that assigned no therapy to the control group (138,140), which resulted in the increase of the effect size of biofeedback therapy. The end point measurements in meta-analysis studies were diverse: some focused on functional scores and the others compared non-functional related parameters such as passive joint ROM, which contribute to the divergent conclusions among the meta-analysis studies. In general, EMGBF yielded positive effects if the outcome measurement was related to control of specific muscle and joint that had been trained; however, the effects of EMGBF did not correlate well with motor function recovery. Another confounding factor contributing to the inconclusive meta-analyses is the combination of upper-limb results (less favorable) with lower-limb results (more favorable).

Another recent study reviewed 26 BF studies for upper-extremity rehabilitation (142). The selected trials were too heterogeneous with regard to the forms of BF (EMG, kinematics, kinetic BF, and feedback of knowledge of result) and the applied patient populations (stroke, SCI, CP, and Parkinson’s disease). In addition, the selected studies investigated BF therapy either alone or as an adjunct to different therapy (CPT, functional electrical stimulation (FES), or therapeutic robotic training). The review found neither the evidence supporting the effectiveness of augmented biofeedback on motor function of the upper extremity nor the relationship between the reported effects and patient characteristics or type of BF.

Recent Task-Oriented Biofeedback Application in Rehabilitation
New Concept: From Static to Task-Oriented Biofeedback
The major body of BF research from 1960s to 1990s trained patients with motor deficits in a static position or nonfunctional related movement. Patients were comfortably seated with arm or leg support while receiving visual or auditory feedback of EMG signals. Such a training paradigm elicits localized motor improvement specific to the trained muscles or joint, but it does not transfer to the generalized motor function of paralyzed limb.

As is well known, one major goal of physical rehabilitation is to encourage motor functional recovery in patients with neuromotor deficits, which in turn will improve their quality of life. Contemporary opinion on motor control principles suggests that improvement in functional activities would benefit from task-oriented biofeedback therapy (14,18,52,143–145).
In physical rehabilitation, task-oriented training encourages a patient to explore the environment and to solve specific movement problems. Therefore, effective biofeedback therapy should also reeducate the motor control system during dynamic movements that are functionally goal-oriented rather than relying primarily upon static control of a single muscle or joint activity (14).

Here, we define “static biofeedback” as the BF training performed when patients are in a static position or nonfunctional related joint motion, and “task-oriented biofeedback” as the BF training conducted when patients are performing movement within functional context. The BF training paradigm should focus on “static” first to improve the motor control of specific muscles or joints, followed by “task-oriented biofeedback” for motor functional recovery.

Several studies have performed BF training in a functional context (6–9,14,146), most of which are for gait retraining. The results of existing task-oriented BF studies demonstrate the promise of the task-oriented training concept; however, a task-oriented BF approach faces several difficulties. First, in the functional task training, it is essential to choose the best information or variable to feedback because muscle activity is not always superior (5). The choice of a biofeedback vehicle depends upon the motor control mechanism, training task, and therapeutic goal. Assume that the training task is to reach for and grasp a cup of coffee using only the affected arm. Motor control models suggest that the brain may control limb kinematics by shifting the equilibrium points (147) or creating a “virtual trajectory” of the end point (148), instead of scaling individual muscle activity patterns. Hence, hand trajectory may be a more viable feedback variable than muscle activity (149). Successful reaching and grasping actions also require the control of the finger grip aperture and a hand orientation (150). Therefore, multiple variables should also be considered for biofeedback. Secondly, feeding back multiple indices to patients whose cognition and perception may also be impaired might overload their perception. If the variables are displayed with traditional abstract and quantitative cues, patients may be overburdened by their ability to process multiple sources and become confused and distracted. Designing a BF system that overcomes the “information overloaded” obstacle for task retraining is both a technical and conceptual challenge.

**New Technologies and Applications for Task-Oriented Biofeedback Training**

**Information Fusion**

An effective task-oriented biofeedback system requires orchestrated feedback of multiple variables that characterize the task performance without overwhelming a patient’s perception and cognitive ability. A usable system of biofeedback for repetitive task training requires sophisticated technology for sensory fusion.

In physical rehabilitation, a therapist visually observes patients’ motor skills and senses the stiffness of joint by stretching a patient’s limb. The therapist then integrates these observations as well as his/her professional knowledge to determine the augmented feedbacks that can enhance the patient’s task performance. Information or multisensor fusion is a technology analogous to the therapist’s integration process, which enables the BF machine to be intelligent for determining the state of patient’s task performance with multiple sensed data and some prior knowledge. Therefore, an information fusion is defined as technologies integrating dynamic and volatile flows of information from multimodal sensing sources to determine the state of the monitored system (151–153). The fusion technologies improve the robustness of the machine perception and decision making to monitor or control dynamic systems or those with uncertain states.

Information fusion is essential to narrow the feedback information to patients during task performance. Indeed, it can be as simple as a classifier in BF applications (9). The complex fusion systems are being developed and applied on real-time movement tracking (154); however, the application of these systems to task-oriented BF training is very limited. Here, we want to elicit the awareness of the technology. Although information fusion is a potentially powerful tool for advanced biofeedback systems integrating multisensor information, the challenge of determining the most appropriate and effective means to provide feedback remains.

**Virtual Reality**

Task-oriented biofeedback therapy might be more widely effective if the biofeedback cues are multimodal. Plasticity is considered one of the neurological mechanisms of biofeedback therapy; the BF training involving multimodal processes (i.e., vision, auditory, proprioception, etc.) may be efficient and beneficial for motor recovery (155). Additionally, the attractive and motivating cue design is always preferred in the success of therapies because motivation and attention are two key factors to induce neuroplasticity (156). Furthermore, to relieve a patient’s burden for processing multiple biofeedback parameters, design of easy-to-understand biofeedback cues is essential (157).

Virtual reality (VR) is a technology that can be used to design biofeedback cues possessing these desired features. First, VR consists of computerized graphics/animation, sound, and/or haptic stimulation; it not only constructs a virtual environment simulating real functional context but also presents multimodal sensory feedbacks to retrain the user’s task performance. Secondly, VR display is immersive so that patients will be engaged in the biofeedback training. Third, VR-based BF cues are intuitive and easy to understand, because the feedback parameters are displayed via a scene resembling a situation in the real world. For example, compared to the display of quantified joint angles in upper extremity, a virtual arm is an intuitive, integrated presentation that is more easily perceived by neurological patients than multiple abstract, quantified presentations. Additionally, VR and training tasks are easily customized by the computer program. The augmented feedbacks, that is, knowledge of result (KR) and knowledge of performance (KP), are easily incorporated in the VR for efficient motor learning (158).
In VR, visual feedback is accomplished via computer graphics. A stereovision can be created with head-mounted display (HMD) or 3D monitors (159,160). In such systems, the computer generates left and right eye perspective images. Each image is connected to the corresponding monitor in the HMD system or to the corresponding half of the 3D monitor. Some systems generate stereovision by using field sequential stereo in combination with liquid crystal shutter glasses. When the perspective scene generated on the screen is for the right eye, the right lens is clear while the left is opaque. Then the left eye's view is displayed, and the glasses switch the clearness of the lens, creating a clear left lens and opaque right lens (160). The scene can be projected onto walls or a desktop. This kind of system is also called “cave” VR or “fish tank” VR. When applying these immersive systems to neurological patients, however, caution must be exercised because motion sickness, dizziness, and visual problems can occur. A less immersive 2D display that employs screen projected image/animation with depth reference frames can be an alternative choice. Figure 70-3A demonstrates an example of 2D interface designed by Arizona State University for retraining of target-oriented reaching task in stroke patients (161).

Surround sound can provide an immersive environment for auditory biofeedback. Sound is an effective feedback source for temporal information; visual feedback works better for spatial information. Auditory feedback can take the form of pleasant and captive music pieces rather than the simplistic and often annoying tones or beeps in older biofeedback studies. An example of musical feedback design for reaching task is shown in Figure 70-3B (157,161). Music notes within a phrase are distributed spatially along the specified path. When the hand reaches a point along the path where a musical note is located, the corresponding note starts to play. The rhythm of music playback depends on the movement speed. Therefore, patients could “compose” different melodies based on movement pattern and quality. If the movement is spastic, for example, the musical phrase could be distorted by multiple repetitions of one note. Studies have shown that musical feedback can improve the motor coordination of Parkinson patients (162) and enhance motor learning in a patient with large-fiber sensory neuropathy (163).

Haptic feedback provides users the kinesthetic sensation that is important for task performance (155,164). The position of an end-effector in the haptic device is sensed at very high rates while a user manipulates the end-effector with the hand. The position is sent to a computerized controller, which then decides the feedback force in a haptic interface simulating the interaction between the user’s hand and the virtual environment. For example, when simulating the feel of squeezing a rubber ball with a force feedback haptic glove, motors within the glove apply forces to fingers. The farther the user flexes the fingers, the harder the motors push back to feedback the squeeze force. Haptic interfaces allow the patient to interact with and to manipulate a virtual object; such interfaces also encourage patients to immerse themselves in the virtual environment.

The major application of VR-based biofeedback to treat sensorimotor deficits has focused on upper-extremity exercise. Preliminary studies on VR-based biofeedback showed promise for motor functional recovery. Holden et al. (149) utilized VR to train reaching and hand orientation in stroke patients. The virtual “mail” fed back the user’s hand position and orientation. The goal of the patient is to reach the slot of a mailbox with correct hand orientation. One of two recruited stroke patients improved Fugl-Meyer (FM) upper-extremity assessment test scores, the performance in a real mailing task, and some motor functions. No improvement was observed in the other patient, however. Later, nine more participants were recruited for further testing (165). All participants showed significant improvement in their FM score, the Wolf Motor Function Test (WMFT) score, and selected strength tests as compared to before the training. The inconsistent results in one patient suggest that a single VR design may not be effective for all stroke survivors. Another study investigated the VR-based BF training for hand rehabilitation in stroke survivors. The training system sensed the finger joint angles from the Cyberglove.

**FIGURE 70-3.** Multimedia-based biofeedback system designed by Arts, Media, and Engineering (AME) program at the Arizona State University. (A) visual feedback of end-point trajectory in VR and (B) auditory feedback of end-point trajectory via music for retraining of target-oriented reaching task in stroke patients. (Reprinted from Huang H, Wolf SL, He J. Recent developments in biofeedback for neuromotor rehabilitation. J Neuroeng Rehabil. 2006;3:11, with permission.)
and the applied finger force and the fingertips position from a haptic glove (166,167). The finger joints were displayed using a virtual hand, and the applied grasping force was fed back via a haptic glove (Rutgers Master II). After 2 weeks of biofeedback therapy, patients improved grasping force, finger joint ROM, and movement speed. Three participants showed hand functional improvement. This study focused on training of hand grasping movement and force; however, the major impairment to hand function in stroke patients is motor incapability for hand opening (extension of metacarpophalangeal joints). To improve the effectiveness of hand functional recovery in patients with brain injury, future designs of VR-based BF should emphasize motor tasks that encourage hand opening and wrist extension rather than retraining hand closure.

Biofeedback using VR for the training of lower extremities has been reported recently. An ankle haptic device and a VR system were integrated and tested on individuals post stroke (168). In the training session, patients sat on a medical chair with the impaired foot strapped to the haptic device. The orientation of a virtual airplane or boat fed back the position of ankle, and the haptic device fed back applied force of foot. The training engaged users to allow for repetitive intensive practice. Preliminary result demonstrated that use of VR to augment rehabilitation of individual post stroke merits further study (168). The number of reported VR-based BF studies on gait retraining is relatively small at present, possibly due to the technical challenge of how to process a multitude of information and then present it to patients. A recent study tracked human stepping motion using motion analysis cameras and displayed lower-limb motion in real time via a virtual human, superimposed upon a virtual instructor (169). The goal of the user was to track the movement of the virtual instructor, both the position of joints and the cadency. Any movement error was displayed by the mismatch of two virtual figures. Recruited able-bodied subjects adapted to the virtual instructor while performing the step-in-place task; however, no testing on patients was reported.

**Biofeedback Therapy Combined With Robot-Assisted Task Retraining**

One inherent limitation of task-oriented BF therapy is that patients with more severe motor deficits cannot participate optimally due to an inability to initiate any functional movement, thus preventing utilization of BF to improve performance. Traditional skateboard facilitated motion using joint angle BF is one of the solutions to ease the limb motion by reducing resisting friction. Recently, rehabilitation robots provide another solution to the problem by providing mechanical assistance to patients’ limb movement (170–173).

Robot-assisted rehabilitation training can also benefit from biofeedback. The problem for robot-aided training is that patients with robotic assistance gradually rely on robots and reduce the active involvement in the task learning. Biofeedback encourages self-regulation and motivation and therefore is a useful, complementary tool for the robot-aided therapy. Several research groups have added biofeedback features into therapeutic robots (170,171). In gait training with Lokomat, which can predefine the pattern of lower-limb kinematics, estimation of self-generated joint torque was displayed in order to motivate patients to produce more force during gait training (171). The combination of robot-aided therapy with task-oriented biofeedback retraining is such an attractive approach for sensorimotor rehabilitation that many new studies can be anticipated.

**Wearable Sensors**

Recently, a technology called wearable sensors or sensor network has been introduced into rehabilitation applications, which permit implementation of task-oriented biofeedback anytime and anywhere (174–176). Jovanov et al. (174) designed a wireless body area network that connected data from multisensors on the body to a personal server such as a cell phone or personal digital assistant (PDA). These data can be sent to other computers through a wireless network. This equipment could be used to provide biofeedback during ambulatory settings. Similarly, Sung et al. (176) developed a flexible wearable platform for ambulatory health monitoring with real-time feedback and context awareness. The physiological and motion data were detected by multisensors attached to the user and were processed in the PDA. Based on the quantitative physiological symptoms and behavioral response, the PDA provided warning feedback through sound or vibration. Wearable sensor technology is still in its early design stage and has not been applied in clinics.

**Telerehabilitation**

Telerehabilitation makes use of advanced technical developments in communication, including wireless vehicles and internet use, to deliver physical therapy to patients at home. Several studies presented in-home VR-based BF systems (177–180). At the user end, patients attached multiple sensors, accessed the BF training programs on the server via a telemedicine network, and received the training with visual or audible feedbacks. Patients may also wear other assistive device such as robots or neuroprostheses to aid the movement of impaired limb. At the other end, therapists could intervene, such as adjustment of task difficulty in the training program. The sensed data were quantified and stored in the database for later retrieving.

Current research that applies telerehabilitation in rehabilitation medicine focuses on investigation of the equivalence of assessment and therapy to in-personal assessment and therapy. The clinical application of telerehabilitation in the practice of physical therapy is very limited, however. One of the reasons is the affordability of assistive device for home use. Patients with sensorimotor deficits very likely do not possess the mobility to perform tasks alone; the assistive devices such as robots are essential for telerehabilitation. Nevertheless, most of therapeutic robots are for research or clinical setting; design of low cost, wearable assistive robots is highly demanded to promote telerehabilitation for rehabilitation (104). These demands may well transform into reality, as costs for health care rehabilitation
delivered to catastrophic injured patients are perceived to escalate. One obvious and potentially cost-effective vehicle to foster the rehabilitation process while channeling more of the treatment and management responsibilities onto caregivers is more extensive use of telehealth.

Future Research and Application on Task-Oriented Biofeedback

The message contained in this chapter should, by no means, imply abandonment of classic EMG BF. Rather, the more effective biofeedback training protocol may start from localized muscle or joint biofeedback retraining, followed by task-oriented training to enhance the motor function. In addition, although task-oriented biofeedback shows promise for functional motor recovery, strong evidence is still lacking. For example, the number of patients in many of the studies cited was small, and some reports that benefit from task-oriented biofeedback were not consistently observed among all participants. Clearly, more clinical RCTs are needed to quantify the efficacy of task-oriented biofeedback therapy.

Our efforts to introduce engineering designs are driven by a need to promote the awareness of existing rehabilitation technologies. Recent advances in technologies are encouraging; however, these technologies are immature, and their clinical value is unclear. In the future, therapists and clinical researchers are urged to study carefully those factors contributing to motor deficits in different patient populations; establish clear training goals; assign necessary biofeedback indices in task-oriented biofeedback training; and work closely with rehabilitation engineers to assist the clinical evaluation of these new techniques.

SUMMARY AND CONCLUSION

Biofeedback remains an important adjunct to the tools of the rehabilitation therapies; it has been subjected to intensive scientific scrutiny with controlled studies of varying quality pervading its history, and ineffective procedures are being extricated. In addition to traditional, static BF training strategy, further BF study may shift attention from static to task-oriented biofeedback training, which may enhance motor functional recovery. Moreover, novel rehabilitation technologies such as VR, therapeutic robots, and telehealth are exciting and have been introduced independently or combined with BF applications. Collectively, the totality of advances among these technologies may bring BF-based rehabilitation into a new era.

ACKNOWLEDGMENT

This chapter is dedicated to the late John V. Basmajian, M.D., often referred to as the “FATHER OF EMG biofeedback” who mentored many rehabilitation scientists and clinicians and who authored preceding versions of this chapter.

REFERENCES


INTRODUCTION

The clinical application of electrical stimulation (ES) to treat neurologic disease may be broadly classified as therapeutic or functional. The purpose of therapeutic ES is to improve the health or voluntary function by inducing physiological changes that remain after the stimulation is used. Therapeutic application of ES may either decrease impairment or prevent further impairment associated with immobility or disuse of a limb or organ system. The use of ES for functional purposes is distinct from therapeutic ES in that the benefits persist only during the period of intervention. The term functional electrical stimulation (FES) refers to application of electrical current to activate paralyzed muscles in precise sequence and intensity as to supplement or replace lost function. In FES applications, stimulation must be “on” to achieve a desired function; therefore, FES systems are usually designed to be worn by the user and are called neuroprostheses because they substitute for lost function. As an intervention in neurorehabilitation, specific applications of FES may have both neuroprosthetic and therapeutic benefit. This chapter is limited to the therapeutic applications of ES to treat the neurologically based disorders most commonly encountered in the clinical practice of Physical Medicine and Rehabilitation. Neuroprosthetic or FES applications of ES is reviewed in the chapter entitled Functional Neuromuscular Electrical Stimulation.

HISTORY

The history of therapeutic ES dates back to 3,000 BC; archeological evidence suggests that Egyptians knew of the powerful electrical shock emitted by an electric catfish found in the Nile River. In ancient Egyptian tombs, hieroglyphics have been discovered which depict the Malapterurus Electricus (1), a species of catfish now known to possess electric organs innervated by a pair of electromotoneurons located within the cervical spinal cord. In 43 AD, Scribonius Largus, a physician to the Roman emperor Claudius, provided the first written documentation of the therapeutic application of ES (2). In the book entitled, “Compositions Medicamentorum,” he detailed the use of the electric torpedo fish for the transient relief of pain associated with headaches and gout. The torpedo fish, which could produce electric shocks between 100 and 150 V, was applied either directly on the patient or placed in a pool of water in which the patient stood to provide therapeutic benefit.

It was not until the middle of the 18th century that potential medical applications of electricity became a focus of study. The Leyden jar, invented in 1745 by Pieter van Musschenbroek (1700 to 1748), was an early device for the storage of electricity which helped early researchers to try various therapeutic applications of electricity. In the 1750s, Benjamin Franklin noted the amnestic effect of electricity applied to the brain (3) and experimented with ES in an attempt to restore motor movement in stroke survivors (4). Luigi Galvani’s description of his “animal electricity” experiments in 1786 and the invention of the voltaic cell by Alessandro Volta (1745 to 1827) significantly contributed to the foundation for modern electrophysiology. Michael Faraday (1791 to 1867) in 1831 performed a series of experiments, establishing the concept of electromagnetic induction. His characterization of the relationship between a DC electric current and a magnetic field further laid a foundation for the therapeutic use of electricity. Duchenne De Boulogne (5) (1806 to 1875) in the mid-1800s used ES both as a therapeutic agent as well as a diagnostic agent to map functional anatomy. Despite these advances, medical skepticism of the therapeutic value of ES persisted through the end of the 19th century.

The advent of the 20th century heralded significant advances in the therapeutic applications of ES. ES of the human sensory cortex was first reported by famed neurosurgeon Harvey Cushing (1869 to 1939) in 1909 (6). Walter Rudolph Hess (1881 to 1973) reported “affective responses” to hypothalamic stimulation in 1928 (7). In 1969, D.V. Reynolds described the analgesic effect of ES on the periaqueductal gray matter of the brain in an animal model (8). Today, transcutaneous electrical nerve stimulation (TENS) and neuromuscular electrical stimulation (NMES) systems are commercially available and routinely employed for the treatment of neuropathic pain and paralysis. Spinal cord stimulation (SCS) and deep brain stimulation (DBS) have been shown to be safe and efficacious for the treatment of various neurological symptoms including neuropathic pain and movement disorders. The burgeoning field of Neuromodulation encompasses the reversible therapeutic modulation of central, peripheral, and autonomic nervous system activity via ES and proposes to provide
practical, clinical treatment options for neurologically disabled patients in the 21st century.

**TENS FOR PAIN**

TENS is a widely used form of electroanalgesia based on the "gate control theory of pain" postulated by Melzack and Wall in 1965 (9). The "gate theory" of pain control states that stimulation of large somatosensory fibers in peripheral nerves or at the dorsal column can effectively inhibit the central transmission of nociceptive signals carried by small, unmyelinated C-fibers from the periphery. Activation of low-threshold myelinated fibers essentially "closes the gate" at the presynaptic level of the dorsal horn of the spinal cord to produce transient pain relief. In 1967, Wall and Sweet (10) demonstrated the electroanalgesic effect of transcutaneous sensory stimulation applied at 100 Hz to patients with chronic cutaneous pain.

The proposed mechanism of neuromodulation via TENS remains controversial. The literature supports a primary effect via presynaptic inhibition at the level of the dorsal horn of the spinal cord as evidenced by changes in neurotransmitter levels associated with application of TENS. In an animal model, high-frequency TENS has been shown to activate delta-opioid receptors and thus reduce the release of glutamate and aspartate at the level of the spinal cord which would otherwise be elevated in association with joint inflammation (11). TENS has also been shown to reduce production of substance P in the dorsal root ganglion and thus suppress nociception via C-fibers in the peripheral nerve (12). Central mechanisms have been proposed to explain the effect of TENS including a neuromodulating effect via descending supraspinal sites such as the rostral ventromedial medulla (13). In a study examining the effect of sensory input on corticospinal excitability, short-term high-frequency TENS was found to decrease human motor cortex excitability (14). Lastly, research has suggested that TENS-mediated hypoalgesia may, in part, be the result of a direct peripheral effect (15).

The first modern TENS unit (16) received a U.S. patent in 1974. A TENS unit is a battery-powered electronic pulse generator that provides low-voltage electrical current through transcutaneous electrodes. A constant voltage stimulator allows for fluctuations in current intensities that may occur with variations in impedance of the electrodes, skin, and tissues. A TENS unit is programmable and optimal settings are determined based on the symptomatic response of the individual patient. The parameters for prescription of TENS include stimulation intensity (mA), pulse frequency (range 2 to 250 Hz), and pulse duration (50 to 1,000 μs). Conventional TENS is the most commonly applied mode of stimulation for rapid-onset, localized anesthesia. Conventional TENS is characterized by high-stimulation frequency (>100 Hz), low intensity (10 to 30 mA), and short pulse duration (50 to 100 μs). Acupuncture-TENS is characterized by high-intensity stimulation (sufficient to evoke muscle contraction) delivered at low frequency (<4 Hz) and longer pulse duration (>200 μs). While acupuncture-TENS may produce a more sustained analgesia, tolerance is less than with conventional TENS. Lastly, pulsed (burst) TENS is essentially a combination of conventional and acupuncture-TENS, and is characterized by bursts of high frequency stimulation (>100 Hz) delivered in bursts at lower frequency (<4 Hz).

No specific advantage is established for pulsed TENS as compared to conventional TENS. The clinician determines the placement of transcutaneous electrodes based on the pain location and may be influenced by trigger points, cutaneous nerve paths, and location of traditional acupuncture sites. For low back pain, for example, transcutaneous electrodes may be placed several centimeters apart either on or adjacent to the painful region, with stimulation being delivered with a pulse frequency of 80 Hz, pulse duration of 100 μs, and amplitude set to the submotor threshold, which varies widely from patient to patient. The onset of analgesia is immediate and may persist beyond the interval of time that the TENS unit is stimulating.

Indications for the use of TENS include both neurogenic pain (deafferentation pain, phantom pain, sympathetically mediated pain, postherpetic neuralgia, trigeminal neuralgia, atypical facial pain, brachial plexus avulsion) and musculoskeletal pain (joint pain associated with rheumatoid arthritis, osteoarthritis, acute posttraumatic pain). Placement of TENS electrodes is contraindicated on or near the eyes, in the mouth, transcerebrally, over the carotid sinuses (vasovagal reflex), anterior neck (laryngospasm), on areas of decreased or absent sensation, or over the trigeminal nerve (if history of herpes zoster induced trigeminal neuralgia). TENS should not be used in epilepsy or pregnancy, or in patients with an implanted cardiac pacemaker or defibrillator due to risk of interference or failure. Skin irritation and electrical burn are the most commonly reported complications, but in general, TENS is well-tolerated.

Despite widespread usage and a 40-year history as an adjunctive treatment to pharmacologic pain management, the therapeutic efficacy of TENS remains controversial. A recent Cochrane review (17) analyzed 25 randomized clinical trials (RCTs) (128 subjects) to determine the efficacy of TENS for chronic pain. Trials were included that compared high-frequency TENS with (a) sham TENS (control), (b) no treatment (control), and (c) low-frequency TENS. In 13 of 22 controlled studies, there was a positive analgesic outcome in favor of TENS. No difference in analgesic efficacy was found in seven of nine studies which compared high- and low-frequency TENS. A second Cochrane review (18) analyzed four RCTs (n = 585 subjects) to determine the efficacy of TENS as compared to placebo for the specific treatment of chronic low back pain. A qualitative synthesis was completed as clinical heterogeneity prevented the use of meta-analysis. The authors noted that there was conflicting evidence about whether TENS was beneficial in reducing back pain intensity and back-specific functional status. They concluded that the evidence from the small number of placebo-controlled trials did not support the use of TENS in the routine management of chronic LBP. In summary, TENS is a noninvasive, widely used modality for...
pain that is easy to apply with relatively few contraindications. However, high quality, randomized, controlled clinical trials remain necessary to definitively assess the therapeutic role of TENS in the management of pain.

**NMES FOR MOTOR RELEARNING**

Motor relearning is defined as the reacquisition of motor skills following central nervous system injury. Research on central motor neuroplasticity supports a primary role of goal-oriented, active repetitive movement training to enhance motor relearning. Asanuma and Keller demonstrated that ES of the somatosensory cortex alone or in conjunction with thalamic stimulation in an animal model induces long-term potentiation (LTP) in the motor cortex (19). They hypothesized that proprioceptive and cutaneous afferent impulses associated with repetitive movements induce LTP in the motor cortex, which then modifies the excitability of specific motor neurons and facilitate motor relearning (20). Consistent with this hypothesis, nonhuman primate research has demonstrated that after local damage to the motor cortex, goal-oriented, active repetitive movement training of the paretic limb shapes subsequent functional reorganization in the adjacent intact cortex, and that the undamaged motor cortex plays an important role in motor relearning (21). Repetitive movements that require the development and practice of new motor skills are the kinds of behavioral experiences that induce long-term plasticity in motor maps (21). When animals are trained to perform new tasks such as retrieving food pellets from a small well (22–24) or a rotating well (25), there is evidence of task-specific cortical reorganization. However, repetitive movement tasks that do not require skill acquisition (i.e., motor tasks that are already mastered and, therefore, are easy to carry out and require minimal or no cognitive effort) are not associated with any significant changes in the motor cortex (24,25).

The use of NMES-mediated goal-oriented repetitive movement therapy for motor relearning is predicated on repetitive movement therapy research. Acute administration of ES to a peripheral nerve activates both sensory and motor structures in the brain (26,27) and reduces the intracortical inhibition (28,29). Functional MRI (fMRI) studies show cortical responses to NMES-facilitated movement in both the upper (30) and lower (31) extremities, which suggests that repetitive movement therapy mediated by NMES has the potential to facilitate motor relearning via cortical mechanisms.

NMES may also facilitate motor relearning via spinal mechanisms. Rushton theorized that the corticospinal-anterior horn cell synapse is a Hebb-type, modifiable synapse and that it can be modified by NMES (32). Originally proposed in 1949, “Hebb’s rule” (33) states: “When an axon of cell A… excite(s) cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells so that A’s efficiency as one of the cells firing B is increased.” The synapse is thought to be strengthened by the coincidence of presynaptic and postsynaptic activities. Under normal circumstances, neural activity in the pyramidal tract easily discharges the anterior horn cells and the strength of the presumed Hebb-type pyramidal tract/anterior horn cell synapse is maintained by this traffic. However, following brain injury, neural activity in the pyramidal tract is significantly reduced. Failure to restore this traffic leads to “decorrelation” of presynaptic and postsynaptic activities, which weakens the synapse. Rushton suggested that NMES-mediated antidromic impulses provide an artificial means to synchronize presynaptic and postsynaptic activity in the affected population of anterior horn cells. Accordingly, he predicted that combining NMES with simultaneous voluntary effort is an effective means of facilitating motor relearning.

Regardless of cortical or spinal mechanisms, experimental and theoretical considerations suggest that the necessary prerequisites for NMES-mediated motor relearning include repetition, novelty of activity, concurrent volitional effort, and high functional content.

**Types of NMES for Motor Relearning**

Three NMES paradigms have been used for motor relearning: cyclic NMES, EMG-mediated NMES, and as a neuroprosthesis. Cyclic NMES provides repetitive activation of muscles without input from the stroke survivors. The activity is novel in that the stroke survivor has difficulty performing the task without the NMES; however, the task is not functionally relevant. EMG-mediated NMES couples cognitive intent and NMES-mediated muscle contraction. This approach may be applied to patients who can partially activate a paretic muscle but are unable to generate sufficient muscle contraction for adequate exercise or functional purposes. The approach incorporates novel tasks and includes cognitive investment, but the task is not functionally relevant. Neuroprostheses provide repetitive movements in the context of functional tasks that are novel; thus, neuroprostheses have a theoretical advantage over both cyclic and EMG-mediated NMES for motor relearning.

**Upper Limb Applications**

RCTs have investigated the efficacy of cyclic NMES in enhancing upper-limb motor relearning (34–38). Transcutaneous electrodes are typically placed over the motor points of the finger and wrist extensors. Four studies reported improved outcomes in motor impairment at the end of treatment, with mild to moderately impaired subjects benefiting most. Two acute stroke studies reported sustained effect (36,37) and two studies reported improvements of activity limitation at the end of treatment (37,38).

EMG biofeedback, position, or EMG-triggered NMES have been evaluated for upper limb motor relearning in six controlled trials (39–44). All six studies demonstrated improved outcomes in motor impairment at the end of treatment; one study showed a sustained effect at 9 months post-treatment (40). Two studies demonstrated decreased activity limitation (41,44). Finally, three studies reported evidence of central mechanisms using neurophysiologic assays including reaction time and fMRI (42–44).
The initial review of these studies suggested that NMES is efficacious in reducing motor impairment but not activities limitation (45). The authors suggested that the effect is more significant for those with milder impairments. A subsequent review by the same group concluded that EMG-mediated NMES may be more effective than cyclic NMES (46). However, a more recent meta-analysis concluded that EMG-mediated NMES was no more efficacious than “usual care” in facilitating motor relearning (47). They noted that most studies were with chronic stroke survivors and results might be different among acute stroke survivors. Consistent with this conclusion, two small RCTs failed to demonstrate the superiority of EMG-mediated NMES over cyclic NMES (48), or usual care (49) among chronic stroke survivors. While acute studies are ongoing, at present it remains uncertain whether cyclic and EMG-mediated NMES are efficacious in facilitating upper limb motor relearning among chronic stroke survivors.

There is emerging evidence that hand neuroprostheses may facilitate motor relearning in the poststroke patient population. Two recent RCTs using a hybrid brace-NMES device that incorporates surface electrodes into a brace for hand grasp and release showed significant improvements in motor impairment (50,51). A third RCT of multichannel NMES also resulted in significant improvement in upper limb motor function (52). Several novel neuroprosthesis approaches with encouraging preliminary results are presently under investigation, including injectable microstimulators (53), contralaterally controlled surface FES (Fig. 71-1) (54), and the incorporation of work stations (55). Additional studies are needed to confirm the motor relearning effect of hand neuroprostheses, especially with respect to translation to functional abilities.

**Lower Limb Applications**

Studies of stroke patients treated with cyclic lower limb NMES have demonstrated enhanced walking ability, increased maximal isometric contraction of the ankle dorsiflexors and plantarflexors, increased dorsiflexion torque, increased agonist EMG activity, and decreased EMG co-contraction ratios (56–58). In one controlled RCT, acute stroke survivors treated with lower extremity NMES and on standard rehabilitation for 3 weeks were more likely to return to a home setting at discharge from rehabilitation compared to controls who did not receive the NMES (58).

The potential motor relearning effect of lower limb neuroprostheses has been evaluated. After using a single channel peroneal nerve stimulator (PNS), Lieberson et al. commented, “On several occasions we observed, after training with the electrophysiologic brace…patients acquire the ability of dorsiflexing the foot by themselves” (59). Since then, case studies using either implanted or transcutaneous electrodes have described similar post-FES observations of improved ambulatory function, more normal EMG muscle activation patterns, emergence of EMG signals in previously silent muscles, increased strength of EMG activity, and decreased co-contraction of antagonist muscles (60–66). A recent meta-analysis of the effect of nonimplanted FES devices on gait speed concluded that FES was effective in improving poststroke gait speed (67). However, studies with implantable PNSs have yielded conflicting results (68,69).

Since gait deviation is not limited to ankle dysfunction, multichannel transcutaneous stimulation systems have been investigated in hemiplegia (70–72). However, the therapeutic application of transcutaneous systems becomes more difficult due to limited muscle selectivity, poor reliability of stimulation, and pain of sensory stimulation. Therefore, percutaneous stimulation systems, which use fine wire electrodes, have been explored for lower-limb motor relearning (73–75). In a single-blinded RCT, chronic stroke survivors receiving gait training with percutaneous NMES demonstrated significant improvements in gait components as measured by the Tinetti Gait scale and knee flexion coordination relative to controls who did not receive NMES (76).

**Summary and Future Directions**

Repetitive movement therapy, mediated by NMES, has the potential to facilitate motor relearning via both cortical and spinal mechanisms. The necessary prerequisites for NMES-mediated motor relearning include repetition, novelty of activity, concurrent volitional effort, and high functional content. Although earlier studies suggested that cyclic and EMG-mediated NMES reduces motor impairment, more recent data, especially among chronic stroke survivors, raise doubts regarding their therapeutic benefits. These approaches may be efficacious among acute stroke survivors, but there are insufficient data to confirm this. While the efficacy of cyclic
and EMG-mediated NMES in facilitating motor relearning remains uncertain, the implementation of neuroprostheses will likely have significant clinical impact due to the higher functional content.

SPASTICITY

For over a century, researchers have attempted to quantitate the effect of ES on spasticity. Duchenne (5), in 1871, noted that stimulating the muscle antagonist to a spastic muscle resulted in a relaxation of the spastic muscle. In 1952, Levine et al. (77,78) published two articles on the inhibitory effect of ES on the spasticity of the antagonist muscles. More recently, Dewald et al. (79) has studied the effects of 10 minutes of low-intensity skin stimulation applied over the spastic biceps muscle of stroke survivors. In seven of nine subjects, ES resulted in both a reduction in peak reflex torque responses in elbow flexors and extensors that lasted for 3 minutes and a significant increase in mean threshold angle for the onset of reflex torque such that a greater angular torque was required to initiate the stretch reflex response. The authors suggested that the changes in reflex torque may reflect synaptic plasticity of spinal circuitry outside the stretch reflex loop.

The effects of TENS on spasticity in patients with various upper motor neuron (UMN) dysfunction have been described. TENS in spastic hemiparesis is associated with an increase in soleus H reflex latencies (80,81) and F-wave latencies (81), decrease in H/M and F/M ratios (81), and decrease in EMG co-contraction ratios (57). Clinical evidence of spasticity reduction associated with TENS applications in stroke include decrease in Modified Ashworth Scores (82), magnitude of plantar flexion stretch reflex (57), and resistive ankle torques (83). Underlying mechanisms for spasticity reduction secondary to TENS have been proposed and include both an enhancement in presynaptic inhibition of the spastic plantarflexor and a “disinhibition” of descending voluntary commands to the parietal dorsiflexor motor neurons (57). A sustained therapeutic effect of TENS on spasticity has not been demonstrated (35).

The effect of TENS on spasticity may be related to stimulation parameters. In a study of 32 patients with spasticity of spinal origin, high frequency (100 Hz) but not low frequency (2 Hz) stimulation was effective for the short-term amelioration of spasticity (84). The antispastic effect induced by high-frequency ES was partially reversed by a high dose of naloxone, indicating that the effect may be mediated in part by endogenous opioids. Clinical evidence of a therapeutic modulation of spasticity associated with NMES is equivocal. Early case series suggested short-term benefits of NMES on spasticity in hemiparesis (85) and spinal cord injury (SCI) (86,87). Stimulation of the tibialis anterior (TA) muscle, prior to a mechanical stretch of the soleus muscle, has been shown to inhibit the stretch reflex of the soleus (88). In a study of spastic hemiparesis, subjects treated for 1 month with lower extremity ES showed a trend toward reduced spasticity based on $F_{\text{max}}/M_{\text{max}}$ ratio, H-reflex latencies, and H-reflex recovery curves (89). As compared to a Bobath inhibitory program, a program of NMES of the dorsiflexor muscles resulted in an increase in passive ankle range of motion and reduction in modified Ashworth score (90). Other researchers have suggested that interaction of electrocutaneous stimulation with an impaired human motor control system may result in unstable reflex loops and cause excessive spastic reactions. In a study of SCI, onset times of spastic activity during an electrically elicited dorsiflexion contraction were noted to shorten with increased stimulation frequency though a stimulation burst did have a spasticity reduction effect on a subsequent burst, suggesting a potential short-term therapeutic effect of stimulation (91).

A therapeutic effect of upper (92) and lower extremity neuroprostheses on spasticity has been suggested. Several studies (93–95) have demonstrated that a neuroprosthesis can effectively open a spastic hemiparetic hand; however, adjustments of stimulation intensities and conditions of upper extremity positioning were necessary for maximal functional response. Short- (96) and longer-term (97) applications of PNS-assisted ambulation have been associated with a decrease in amplitude of the Achilles tendon reflex (96), and a decrease in the tonic activity of the TA and triceps surae muscles (97).

The efficacy of adjuvant ES combined with botulinum toxin A for treatment of spasticity has been proposed (98–101), based on animal studies demonstrating enhanced toxin uptake and accelerated onset of paralytic effect by ES. Frasson et al. (99) reported that low-frequency ES (4 Hz), but not high frequency ES (25 Hz) accelerates the effect of botulinum toxin on neuromuscular blockade. In two small controlled studies, Botulinum toxin combined with NMES reduced spasticity in both the upper (102) and lower (103) extremities. However, a recent single-blinded study found no difference in efficacy of low-dose Botulinum toxin (100 U) plus short-term ES versus high-dose Botulinum toxin (400 U) to treat equinovarus spasticity (104).

In summary, there is evidence of short-term electrophysiologic effects of TENS on spasticity in patients with various UMN disorders. The functional relevance of short-term decreases in spasticity secondary to TENS is unclear; long-term modulation of spasticity secondary to TENS has not been established. A therapeutic effect of NMES on spasticity has not been established; variable outcomes may be influenced by stimulation parameters, application methods, quantification measures, and etiology of UMN dysfunction. There is evidence that NMES may enhance the antispastic effect of Botulinum toxin; however, the superiority of NMES to other adjuvant interventions has not been established.

BONE DENSITY

The physiologic alterations of bone associated with immobility in the neurologically impaired patient are well-described. Following SCI, weight-bearing trabecular-rich sites such as the distal femur and proximal tibia show the greatest amount of
demineralization. Pathologic fractures following minor trauma can be caused by reduced bone mass in association with modified bone matrix composition (105,106). Capacitively coupled electrical fields have been shown to prevent and reverse osteoporosis in animal studies (107–111). Various applications of ES have been shown to stimulate osteogenesis (112), positively affect bone mineral density (BMD) (113,114), and enhance fracture healing (115) in animal models.

Studies have described variable short-term effects of NMES (116) and FES-cycling (117–121) programs on bone density in SCI. However, several studies have found site-specific effects of FES-cycling on neurogenic osteopenia. In a controlled study (122), 6 months of FES-cycling resulted in a partial reversal of BMD loss in the distal femur and proximal tibia in SCI; however, the effect was not sustained 6 months after completion of the FES-cycling program. Belanger et al. (123) measured the BMD of the distal femur, proximal tibia, and mid-tibia in SCI following a 6-month FES-cycling program. At baseline, the BMD of SCI subjects was lower than that of controls. After training, the distal femur and proximal tibia had recovered approximately 30% of the bone loss. Shields and Dudley-Javorski (124) applied unilateral NMES to the plantar-flexor muscles in SCI using each subject as his own control. A 2-year training program resulted in a 31% higher distal tibia trabecular BMD in trained limbs as compared to untrained limbs (Fig. 71-2).

In summary, the literature supports the use of ES, in the form of lower extremity FES-cycling, to enhance bone mineral density at the distal femur and proximal tibia in SCI. However, the therapeutic effects on BMD are short-term and have not been shown to persist beyond the period of the FES-cycling intervention. The clinical efficacy of ES to prevent osteopenia, enhance BMD, or prevent fractures in non-SCI UMN paralysis has not been established.

**FATIGUE, MUSCLE ATROPHY, AND STRENGTHENING**

NMES has been demonstrated to have therapeutic benefit for muscle weakness and atrophy following central nervous system injury. Loss of muscle mass may occur as a result of spinal motoneuron injury resulting in denervation atrophy or central pathway injury resulting in disuse atrophy. Muscle fiber alterations at the cellular level form the basis of present understanding of NMES-associated muscle strength enhancement. Skeletal muscle contains type I (slow-twitch) fibers and type II (fast-twitch) fibers. Type I fibers have high oxidative capacity. Type II fibers have high glycolytic capacity with the ability to metabolize adenosine triphosphate (ATP) more quickly and are thus associated with short, forceful contractions. NMES reverses the transformation of type I fibers to type II fibers, commonly seen following an UMN injury (125). In an animal model (126), changes in fiber population and myosin light chain patterns, at the level of the mitochondria and sarcoplasmic reticulum, are associated with intermittent or continuous stimulation. Ljubicic et al. (127) noted numerous biochemical and physiological adaptations induced by chronic low-frequency ES in an animal model including: alterations of calcium dynamics and myofibrillar proteins, changes in mitochondrial enzymes, myoglobin, and induction of angiogenesis. Training-induced changes in fatigue resistance seem to arise, in part, from an improved oxidative capacity. In a controlled study that compared selected morphological and metabolic properties of single fibers from the TA (128) SCI subjects who received 6 months of ES increased both the proportion of type I fibers and activity of succinate dehydrogenase (SDH) in both fiber types. Muscle fatigue has also been correlated to alterations at the cellular level form the basis of present understanding of NMES-associated muscle strength enhancement. Skeletal muscle contains type I (slow-twitch) fibers and type II (fast-twitch) fibers. Type I fibers are fatigue-resistant with high oxidative capacity. Type II fibers have high glycolytic capacity with the ability to metabolize adenosine triphosphate (ATP) more quickly and are thus associated with short, forceful contractions. NMES reverses the transformation of type I fibers to type II fibers, commonly seen following an UMN injury (125).

of low-frequency ES on muscle atrophy including variable muscle loading conditions, effects of both animal gender and anesthesia on hormone levels (cortisol, testosterone, growth hormone), and conversion of fast to slow muscle fibers associated with higher-frequency stimulation-induced increases in basal cytosolic calcium levels.

A limited number of studies have evaluated the efficacy of NMES on muscle strength when applied to the upper extremity. Peckham et al. (137) demonstrated an increase in contractile force resulting in functional gain with the application of ES to the forearm finger flexors of ten quadriplegic patients. However, Seeger et al. (138) studied the effect of FES-assisted exercise versus isotonic exercises and found no increase in maximum voluntary upper extremity force in SCI. In separate studies, NMES applied to the wrist extensors in hemiplegia has also been demonstrated to be associated with higher isotonic wrist extension torques (37), greater isometric finger extension (44), and improved grip speed (139).

The ability of NMES, used in various applications, to enhance the strength of specific lower extremity muscles post-SCI has been demonstrated. Cyclic NMES to the quadriceps has been shown to enhance quadriceps muscle strength (123,140). Lima (141) compared NMES to isotonic exercises of the knee and found that flexor torque, though not extensor torque, was increased in the NMES group. Similarly, FES-cycling programs (142,143) and lower-extremity FES-exercise programs (144) have been shown to be effective in increasing the muscle strength in SCI. Sloan et al. (145) reported improvements in voluntary isometric strength, stimulated isometric strength and stimulated isometric endurance of the quadriceps, muscle grading of the quadriceps and biceps femoris and the cross-sectional areas of the quadriceps and total thigh muscle associated with a 3-month FES-cycling program in SCI. Guidelines for FES training in SCI were proposed by Petrofsky et al. (146) who studied 90 paraplegic patients and determined that the optimal protocol for FES-isokinetic exercise of the quadriceps was a three time per week regimen for 30 min/day with equal alternating periods of work and rest. Kern et al. (147) determined that the striated appearance of normal skeletal fibers may be maintained up to 20 years post-SCI and thus proposed that there was no upper limit for the initiation of a training program based on FES in SCI. Several reviews outline the use of NMES to increase strength and endurance in the SCI (148–150), stroke (151), and able-bodied populations (152).

In summary, on a cellular level, NMES has the ability to reverse the transformation of type I fibers to type II fibers that is commonly seen following an UMN injury. NMES can play a preventative role in minimizing the amount of muscular atrophy seen following an UMN injury. The ability of cyclic NMES or FES-cycling to reverse muscle atrophy and enhance lower extremity muscle strength in SCI is supported by the literature and accepted in clinical SCI care. While NMES may be clinically used to prevent muscle atrophy and enhance strength in other UMN disorders including stroke, the clinical and functional relevance of resultant strength changes has not been established.

### Tissue Oxygenation and Wound Healing

Early studies (153,154) proposed the therapeutic benefit of ES in the treatment of soft tissue injury and wound healing. Subsequent studies have demonstrated that ES affects wound healing by both a direct local tissue effect at the cellular level and via an enhancement of angiogenesis, tissue perfusion, and oxygenation (155–157). Certain chemotactic factors found in wound substrates contribute to tissue repair processes by attracting cells, which carry positive or negative charges, into the wound environment. ES can facilitate the induction of directional cell migration (i.e., electrotaxis) (158) of neutrophils, macrophages, epidermal cells, and fibroblasts into the wound. These cells accelerate wound healing through their specific cellular roles in autolysis, cellular protein synthesis, mitotic cell division, and anti-inflammatory activities (159,160). In a recent review, Kloth (155) presented the relationship between endogenous bioelectric currents and wound repair, and reviewed clinical research applications of ES on wound healing.

ES induces angiogenesis to enhance tissue perfusion and oxygenation. In a study by Zhao et al. (161), applied electric fields of small physiological magnitude directly stimulated vascular endothelial growth factor (VEGF) production by endothelial cells in vitro. They proposed that endogenous electric fields (EFs) might play a role in angiogenesis in vivo by stimulating the VEGF receptor signaling pathway to induce key preangiogenic responses. Other animal studies have demonstrated ES-induced changes including increased capillary-to-fiber ratios, blood perfusion, transformation of muscle fiber types from mixed (oxidative and glycolytic) to primarily oxidative fibers, and improved muscle fatigue (162–166).

In vivo studies support the efficacy of ES for wound healing. In a controlled study, patients with stage IV decubitus ulcers treated with high voltage ES healed at a mean rate of 44.8% per week with 100% closure in a mean of 7.3 weeks whereas ulcer size of the control group increased over the same period of time (167). In a study of 30 patients with sacral or leg ulcers, low-intensity DC significantly accelerated the rate of wound healing in humans 1.5 to 2.5 times faster when compared to conventional wound therapy (168). In a meta-analysis, Gardner et al. (157) analyzed 15 studies, which included 24 ES samples and 15 control samples. The rate of healing per week was 22% for ES and 9% for controls; the net effect of ES was 13% per week, an increase of 144% over the control rate. Factors which appear to be predictive of wound healing by ES include wound size, patient’s age, time elapsed from wound appearance to the beginning of the treatment, width-to-length ratio, location, and type of treatment (169). Additionally, biphasic- and direct-current stimulation may result in faster healing of chronic wounds (169).

An optimal delivery system for wound healing and pressure sore prevention in patients who lack sensation or are unable to perform independent pressure relief has not been established. Electrode design and ES delivery systems must consider both...
temperature (170) and nonhomogeneity of the wound (171) that can affect current delivery and skin responsiveness. In an able-bodied adult, intermittent ES has been shown to increase oxygenation level of the buttock muscles by 6% (172). The role of percutaneous ES (173) and an implanted NMES stimulation system (174,175) consisting of intramuscular electrodes with percutaneous leads has been proposed. In a series of SCI patients, Bogie and Triolo (175) found that 8 weeks use of an implantable NMES device was associated with an increase in the unloaded tissue oxygen level and a reduction in the ischial region pressure. A second study (174) demonstrated an increase in gluteal muscle thickness of 50%, decrease in regional interface pressures of 20%, and an increase in tissue oxygen levels associated with the implanted system.

In summary, ES has a therapeutic effect on wound healing by facilitating the induction of cell migration (neutrophils, macrophages, epidermal cells, and fibroblasts) into the wound and via angiogenesis, tissue perfusion, and oxygenation. The therapeutic efficacy of ES on wound healing is dependent on wound characteristics and ES parameters and delivery systems. Research suggests that biphasic- and direct-current stimulation applications result in faster wound healing rates. Further research is needed to determine the clinical efficacy and viability of implanted NMES systems for pressure wound healing and prevention.

CARDIOVASCULAR AND PERIPHERAL HEMODYNAMIC EFFECTS

NMES can affect specific cardiovascular and peripheral hemodynamic parameters when applied therapeutically to subjects with muscle paralysis. The specific effect of NMES on peripheral arterial blood flow, arterial diameters, and pressures has been variable. In a study of able-bodied adults, a linear increase in femoral arterial blood flow was associated with increasing frequencies of transcutaneous NMES applied to the leg muscles (176). Stoner et al. (177) evaluated the effects of NMES-induced quadriceps exercise therapy on posterior tibial artery resting diameter, flow-mediated dilatation, and arterial range in SCI subjects. Following 18-weeks of NMES training, flow-mediated dilatation and posterior tibial arterial range improved. Sherry et al. (178) studied the effects of low-frequency (2 Hz), burst-mode TENS applied over the common peroneal and tibial nerves in healthy controls. When TENS was applied at 25% above the motor threshold a transient increase in calf blood flow was detected; however, TENS had no effect on arterial pressure regardless of stimulation intensity. Similarly, Sabatier et al. (179) found no change in femoral artery diameter on resting, reactive hyperemic, and exercise blood flow associated with 8 weeks of NMES-facilitated quadriceps resistance training in SCI.

NMES-facilitated exercise may enhance venous return, stroke volume (SV), and cardiac output (CO) in SCI. NMES-facilitated knee exercise resulted in an increase in CO associated with an increase in venous return but no change in heart rate (HR) (180). In a study of orthostatic responses in SCI, Raymond et al. (181) noted a 20% increase in SV, 16% increase in CO, and a 12% decrease in total peripheral resistance associated with NMES-induced muscle contractions. Lastly, an increase in blood flow and decrease in venous pooling was associated with a reduced rate pressure product in SCI subjects who performed arm cramp ergometry while concomitantly receiving NMES-induced leg contractions (182).

Lower extremity FES-treadmill training with body weight support (BWS) has been studied for its therapeutic effect on cardiovascular responses in SCI (183,184). In a controlled study, Carvalho (184) noted significant increases in oxygen uptake, carbon dioxide production, pulmonary ventilation, and systolic blood pressure (SBP) following 6-month FES-treadmill training. In a study of the training effects of FES-assisted amputation in SCI (185), effects consistent with central cardiovascular adaptations, including a posttraining increase in peak oxygen uptake, peak power output, and time to fatigue, were noted. Lastly, in a study of chronic tetraplegic subjects who underwent 3-month FES-treadmill gait training with BWS (186), resting and exercise SBP increased from baseline without change in HR.

FES-leg cycle ergometry (FES-LCE) programs have also been shown to improve the central cardiovascular fitness in the SCI patient population. In a series of studies (181,187–191), Raymond et al. examined the effects of FES-LCE on cardiovascular function. When arm cranking was combined with FES-LCE in paraplegia (188), a significantly higher oxygen uptake, SV, and expired ventilation volume was noted during submaximal, steady-state exercise, as compared to arm cranking alone. A subsequent study (189) demonstrated that CO was higher in paraplegic subjects, using FES-LCE as compared to voluntary leg cycling, at equivalent submaximal oxygen uptake rates. Faghri et al. (192) noted that HR and blood pressure significantly decreased during submaximal exercise, whereas SV and CO significantly increased, following a 12-week FES-LCE program. In a study which compared FES-LCE to a continuous passive motion program in six tetraplegic subjects, an increase in both CO and the arterial-venous oxygen difference accounted for an elevation of oxygen uptake (193). Lastly, Hooker et al. (194) suggested that peripheral circulatory adaptations develop in SCI patients during FES-LCE training. In a study of 18 SCI subjects, significantly higher peak oxygen levels (45%), oxygen uptake (23%), pulmonary ventilation (27%), HR (11%), CO (13%), and significantly lower total peripheral resistance (14%) were found following 16 weeks of FES-LCE.

In summary, various forms of NMES-facilitated lower extremity exercise training have been demonstrated to enhance venous return, SV, and CO in SCI. Beneficial training effects have been demonstrated with the application of NMES in the forms of NMES-facilitated exercise, NMES-assisted treadmill gait training, and FES-LCE. The therapeutic efficacy of NMES-facilitated exercise training programs to enhance cardiovascular and peripheral hemodynamic functioning has not been established for other UMN diagnoses.
DEEP VENOUS THROMBOSIS PROPHYLAXIS

Factors contributing to the development of deep venous thrombosis (DVT) in the rehabilitation patient population include impaired venous return and hypercoagulability. Potential mechanisms for the therapeutic efficacy of ES for prevention of DVT have been proposed. In a study of 49 able-bodied subjects (195), ES resulted in a significant increase in venous femoral and popliteal blood flow in the calf and plantar muscles. In a study by Lindstrom et al. (196), intraoperative ES-induced tetanus of the calf muscles reduced calf venous volume approximately three times more efficiently than ES single impulses. No enhancement of the fibrinolytic activity of venous blood was observed in the postoperative period. A more recent study, however, suggests that ES may have an important antithrombotic effect on arterial and venous thrombosis. Augejouf et al. (197) explored the antithrombotic properties of ES using a laser beam-induced thrombosis animal model. Continuous ES was found to decrease the number of emboli and the duration of embolization both in arterioles and venules and reduce the amplitude and velocity of the ex-vivo platelet aggregation induced by ADP.

A therapeutic role of ES of the calf muscles for the prevention of venous thrombosis and embolism was first explored in 1949 (198). More recent studies that focused on the application of ES in the perioperative period to reduce the incidence of venous thrombosis have found variable degrees of efficacy (199–202). Merli et al. (203) compared the additive effects of ES in combination with low-dose heparin, heparin alone, and placebo in acute SCI. Only the combination of low-dose heparin and ES group had a significant reduction in the incidence of DVT. A second study by Merli et al. (204) reported that heparin, in combination with thrombo-embolism deterrent (TED) hose, was similarly effective to heparin plus ES. Two review articles (205, 206) that focused on the incidence and prevention of DVT in acute SCI recommended the clinical use of ES in combination with low-dose heparin as a prophylactic measure based primarily on these two specific clinical trials.

In summary, sequential ES applied to the calf muscles appears to increase the venous blood flow. The ability of ES to prevent venous thrombosis in the perioperative period has not been established. ES, as an adjunct to low-dose heparin, may decrease the risk of venous thrombosis in acute SCI as compared to heparin alone. As an adjunctive therapy, however, ES provides no greater benefit than TED hose. No study has compared ES, as either a primary or adjunctive intervention, to low-molecular weight heparin (LMWH) which is presently a standard of care for prevention of venous thrombosis in the acute post-SCI period. Further studies are indicated to explore potential antithrombotic properties of ES.

POSTSTROKE SHOULDER PAIN AND SUBLUXATION

Shoulder pain and subluxation are common complications following stroke. Poststroke shoulder dysfunction is characterized by early spasticity and weakness which may progress over time to result in mechanical instability and diminished mobility at the glenohumeral joint. The relationship between poststroke shoulder subluxation and poststroke pain is complex. Shah et al. (207) found a 35% prevalence of rotator cuff, biceps and deltoid muscle tears and 53% prevalence of tendinopathy in chronic stroke patients with shoulder pain. Importantly, the presence of rotator cuff tears, rotator cuff and deltoid tendinopathies were not found to be related to the severity of poststroke shoulder pain. In a study of 107 acute stroke subjects (<30 days poststroke), glenohumeral subluxation was present in 48.6% of patients (208). Presently, transcutaneous and intramuscular NMES of the supraspinatus, trapezius, and deltoid muscles are potential treatment options under investigation to reduce subluxation, improve biomechanical integrity, and reduce pain.

Eighty-one percent of patients may develop subluxation following stroke (209). NMES applied to the shoulder has the ability to reduce poststroke shoulder subluxation. In a study that compared conventional physical therapy (PT) with and without NMES to the supraspinatus and posterior deltoid muscles (210), the NMES group showed significant improvement in arm function, EMG activity of the posterior deltoid, range of motion, and reduction in subluxation. Studies suggest that a therapeutic program of NMES applied to the shoulder may reduce subluxation in both acute (211) and chronic (212) poststroke patients. However, in a controlled study, Wang et al. (213) focused on duration of poststroke and compared the effectiveness of 6 weeks of NMES applied 6 h/day to the supraspinatus and posterior deltoid muscles in the acute (<1 year) versus chronic (>1 year) poststroke periods. Shoulder NMES was only effective in reducing subluxation in the acute stroke group. A meta-analysis (209) determined that transcutaneous ES applied to the shoulder and combined with conventional therapy, prevented on average 6.5 mm of shoulder subluxation as compared to 1.9 mm with conventional therapy alone. The authors concluded that NMES reduced or prevented shoulder subluxation and improved motor impairment in the acute poststroke period, but not in the chronic poststroke period. Lastly, Wang (214) in a review article also concluded that 6 weeks of NMES applied to the supraspinatus and posterior deltoid muscles enhanced the upper limb motor recovery in the early poststroke period. However, a recent study raises questions about the effect of shoulder NMES on motor recovery when applied in the acute poststroke period. Church et al. (215) studied the effect of shoulder transcutaneous NMES applied to stroke patients within 10 days poststroke. Significant differences were seen at 3 months in favor of the control group for specific measures of arm function and impairment including the grasp and gross movement subsections of the ARAT, Frenchay Arm Test, and the arm subsection of the Motricity Index. They concluded that the “routine” use of transcutaneous NMES to the proximal upper extremity after acute stroke could not be recommended. However, this study enrolled all stroke survivors regardless of presence or absence of subluxation, pain at baseline was not an inclusion criterion, they received substantially less NMES compared to previous
studies and outcome measures focused on distal hand function and were not specific to the proximal intervention.

Shoulder pain, distinct from shoulder subluxation, is a frequent poststroke complication that is associated with a reduced quality of life (216) for stroke survivors. The incidence of poststroke shoulder may be 40% within the first 6-months poststroke (217). A 72% incidence of shoulder pain was found in a series of 219 hemiparetic patients followed for 1-year poststroke (218). The majority of patients with poststroke shoulder pain rate the pain to be of moderate to severe intensity (219). The effects of TENS and transcutaneous NMES on poststroke shoulder pain have been explored. Linn et al. (211) found that 4 weeks of cyclic NMES applied within 48 hours poststroke reduced the shoulder pain and subluxation though the effect was not maintained beyond the treatment period. In a review article (220), four randomized trials (34,210,211,221) were evaluated that assessed various transcutaneous ES techniques applied at variable poststroke durations for prevention or treatment of shoulder pain. The authors concluded that ES did not affect pain incidence or intensity though ES improved pain-free range of passive humeral lateral rotation. More recently, Chantraine et al. (222) compared the effect of 5 weeks of transcutaneous NMES, applied within the first month poststroke, to conventional therapy in a nonrandomized controlled study of 120 patients with both pain and subluxation. In this study, the NMES group showed significantly more improvement than the control group on pain relief, reduction of subluxation, and recovery of arm motion.

Although data suggest that transcutaneous NMES reduces shoulder subluxation and improves pain-free range of motion, clinical implementation has been difficult due to the need for skilled personnel to ensure tolerable and reliable stimulation. In order to address this limitation, a percutaneous, intramuscular NMES system which includes helical intramuscular electrodes, a “pager” size stimulator, and connector wires has been trialed for the treatment of poststroke shoulder pain (Fig. 71-3). In a controlled clinical trial, Yu et al. (223) studied 61 chronic stroke survivors with shoulder pain and subluxation. The treatment group received percutaneous, intramuscular NMES to the supraspinatus, posterior deltoid, middle deltoid and trapezius muscles, can reduce shoulder pain for up to 1 year posttreatment. However, given its invasive nature, the superiority of percutaneous NMES over its invasive nature, the superiority of percutaneous NMES over “best practice” and “usual care” in reducing poststroke shoulder pain needs to be demonstrated.

**SPINAL CORD STIMULATION**

SCS has been used for more than 30 years to treat a variety of pain syndromes. Shealy et al. (227) first reported the modulation of pain associated with ES of the dorsal columns of the spinal cord in 1967. The analgesic effect of spinal epidural stimulation was subsequently reported by Shimoji et al. (228) in 1971. Lindblom et al. (229), in 1977, described the effect of single and repetitive ES of the dorsal columns on cells in laminae IV and V of the ipsilateral dorsal horn at S1 in an animal model. They noted that post-dorsal column stimulation suppression of pain appeared to be secondary to spinal mechanisms whereas longer-lasting pain relief appeared dependent on interneurons and suprasegmental loops. A dorsal column stimulator (DCS) is a specific implantable neuromodulation device used to treat chronic, intractable pain of neurologic origin. More recent human studies have demonstrated the clinical
efficacy of dorsal column stimulation in relieving specific chronic pain disorders including failed back syndrome, pain associated with SCI, complex regional pain syndrome (CRPS), and postamputation limb pain.

The proposed mechanism of action of SCS for reduction of neuropathic pain remains poorly understood. In an early animal model (230), DCS applied to the cervical and mid-thoracic levels depressed the activity of spinothalamic tract cells at the lumbar sacral level for about 150 ms. In a study of the morphology of human superficial dorsal and dorsolateral column fibers at the lower thoracic spine (231), the mean A-beta fiber diameter, density and frequency increased significantly medially to laterally and the largest A-beta fibers were activated up to a maximum depth of about 0.25 mm. In a study of DCS in normal humans, Shimoji et al. (232) suggested that the pain-alleviating effect may be at least partly due to orthodromic and/or antidromic activation of descending inhibitory pathways which inhibit the transmission of pain signals by A delta and C fibers from the corresponding dermatome. A review of mechanisms of SCS (233) concluded that multiple mechanisms may include antidromic activation of large-diameter afferent fibers (232), activation of supraspinal loops relayed by the brain stem or thalamocortical systems (234–237), activation of the anterior pretectal nucleus resulting in descending inhibition of pain (238–240), suppression of wide-dynamic-range neurons within the dorsal horn (241–243), and antidromic release of neuropeptides such as calcitonin gene-related peptide (244–246).

In clinical trials, the efficacy of SCS for chronic pain is generally assessed by a postimplantation reduction of pain of at least 50% and reduction of oral pain medication requirements. In a multiyear retrospective study, Kumar et al. (247) reported on the efficacy of epidural SCS in 121 subjects with chronic pain of various etiologies. Postimplantation, subjects were followed for periods ranging from 6 months to 10 years, with a mean follow-up period of 40 months. Forty-eight patients (40%) were able to control their pain by neurostimulation alone; 14 patients (12%) required occasional analgesic supplements, in addition to stimulation, to achieve 50% or more relief of the prestimulation pain. In a multicenter study of 90 patients implanted with a DCS for chronic pain of the back and/or legs, Shatin et al. (248) reported that 70% had “good” or “excellent” pain relief at an average of 14.5 months postimplantation. The percentage of patients requiring opioid medications decreased from 40% preimplantation to 28% postimplantation. Lastly, in a 5-year follow-up study of 50 patients implanted with a SCS for failed back syndrome (249), a minimum of 50% reduction in pain was recorded in 53% of patients at 2.2 years and in 47% of patients at 5-years postimplantation. Eighty-three percent of the subjects continued to use their stimulators at the 5-year follow-up. The percentage of patients using analgesic medication decreased from 74% preimplantation to 12% postimplantation. In a study evaluating patient satisfaction postimplantation (250), 60% of 41 patients considered themselves improved from their preoperative condition and 78.1% of patients would recommend SCS to someone with a similar pain.

Other important outcome measures have been utilized in clinical SCS trials, including measures of impairment, activity limitation, and cost-effectiveness. In a prospective study of patients with intractable lower extremity pain who underwent SCS implantation (251), a statistically significant improvement in isometric lower extremity function 6 weeks postimplantation was found that persisted at both 12- and 24-months follow-up. Kemler and Furnee (252) randomized 54 subjects with CRPS to receive either SCS and PT or PT only. Pain, though not reflecting functional status or depression, improved significantly in the SCS+PT group at 6 months postimplantation. At 1 year, pain and quality of life were improved significantly only in the SCS+PT group. The cost effectiveness of SCS as compared to conventional therapy has been suggested in analyses of the long-term medical costs incurred for treatment of chronic pain (252–254).

The implantation of a DCS is a minimally invasive, reversible procedure (Fig. 71-4). Dorsal column stimulation can only be effective if dorsal column fibers in the spinal cord are preserved and can be stimulated by implantable electrodes. Thus, the implantation of a DCS involves two stages. In the first stage, a temporary percutaneous lead is placed and connected to an external pulse generator. If stimulation is effective in reducing pain by at least 50% over a period of 3 to 7 days, the patient is considered a candidate for the implanted SCS system. In the second stage, one or more epidural electrodes are inserted into the spinal canal over the dorsal aspect of the spinal cord. The implanted electrode may be placed via percutaneous insertion or may be surgically placed via a laminectomy. The spinal level for electrode placement is dependent on the location of the pain. There are several types of SCS implantable units including a conventional implantable pulse generator (IPG) (neurostimulator), rechargeable IPG, and a radiofrequency device. The radiofrequency device consists of an implanted receiver and transmitter device that utilizes radiofrequency to control stimulation. Postimplantation, the SCS system is programmed for the appropriate stimulation patterns and the patient is provided with a remote control to activate the stimulator. Complications of implantation are generally related to the surgical procedure itself and include infection, epidural hematoma, paralysis, and lead migration.

Successful reduction of pain, secondary to SCS, is highly dependent on patient selection; most candidates have previously failed conventional pain treatment modalities. Psychological factors may contribute to DCS failure which underscores the need for specific criteria for patient selection (255). Preoperative psychiatric evaluation and a trial of percutaneous DCS testing may be predictive of the individuals likely to have the greatest benefit (256). Multiple reviews (257–263) have summarized the efficacy of SCS for treatment of chronic pain associated with failed back syndrome, SCI, CRPS, postamputation limb pain, and other chronic pain syndromes. There is clear evidence that ES of the dorsal columns can modulate pain. In the
clinical care of chronic pain patients, SCS is a viable treatment option for select patients with intractable pain who have failed conventional management.

**DEEP BRAIN STIMULATION**

DBS is a nonablative, reversible neurosurgical intervention for the treatment of Parkinson's disease (PD) and other neurologically based movement disorders. DBS of the thalamus was first introduced in France in 1987 for essential tremor and was applied in 1993 to the subthalamic nucleus (STN) (264) for the treatment of PD (265). DBS involves the application of continuous low-voltage electrical impulses to targeted areas of the brain for the purpose of modulating dysfunctional brain signals. The most common sites of electrode implantation are the thalamus, STN, and globus pallidus (GP). For dystonia and symptoms associated with PD (rigidity, bradykinesia/akinesia, and tremor), the electrode is generally placed in either the GP or STN; for essential tremor, the electrode is
is used to adjust the stimulation parameters to provide the appropriate level of stimulation to optimize symptom suppression while minimizing the side effects. The patient can control the duration of stimulation via a handheld control device. The DBS system can be used to provide symptom relief for up to 24 hours a day.

There are potential complications associated with DBS (277). Complications specific to the surgical placement of the device include infection, mild disorientation or sleepiness, intracranial swelling, cerebrospinal fluid leak, or brain hemorrhage. Complications associated with the technical components of the device itself are relatively common. In a retrospective study of 84 patients (278), 25.3% of subjects experienced hardware-related complications with electrode connector issues being the most frequent problem. Similarly, a prospective study (279) found that 26.2% of DBS patients had hardware complications which included lead migration, lead fracture, extension erosion, extension fracture, and IPG malfunction. The potential risks associated with MRI imaging (280,281) and functional MRI imaging (282) in patients with DBS systems include excessive heating, electrode displacement (281), and/or functional disruption of the device. DBS patients may be safely imaged with MRI though specific guidelines for MRI usage in individuals with DBS systems are dependent on the IPG model and components. Neuropsychiatric side effects were described in a controlled

FIGURE 71-5. Soletra neurostimulator (Medtronic). DBS system consists of a neurostimulator for controlled electrical pulse generation delivered by extension and lead(s) to the targeted structures deep within the brain. (Courtesy: Medtronic Corp.)
study of PD patients who underwent bilateral STN DBS (283). Six months postimplantation, declines were noted in verbal fluency, naming, selective attention, and verbal memory. Approximately, 9% of the STN patients had more serious psychiatric complications as compared to 3% of controls. In a more recent controlled study (284), DBS resulted in a selective decrease in frontal cognitive functions at 6 months post-implantation which did not negatively affect the quality of life. Lastly, a recent meta-analysis (285) concluded that there were relatively low rates of depression, cognitive impairment, mania, and behavior change but an increased rate of suicide in patients treated with DBS, particularly patients with STN and GP stimulation.

DBS has been proposed as a therapeutic intervention to promote late functional recovery following traumatic brain injury (286–291). In 2007, Schiff et al. (292) reported behavioral modulation associated with bilateral DBS of the central thalamus in a patient who remained minimally conscious 6 years following a TBI. Outcome measures included specific cognitively mediated tasks, functional limb control, and oral feeding. A statistical linkage between functional improvements and stimulation was determined. DBS is presently under investigation for the treatment of arm tremor in multiple sclerosis (293,294), major depression (295), Tourette’s syndrome (286), epilepsy (296,297), obsessive-compulsive disorder (298,299), and neuropathic pain (300–302).

CORTICAL ELECTRICAL STIMULATION

Several forms of cortical stimulation are being trialed as neuro-modulation interventions to ameliorate symptoms of neurologic disease. The three most common types of cortical stimulation are transcranial magnetic stimulation (TMS), transcranial ES (using pulsed or direct current), and epidural cortical stimulation (ECS) using surgically implanted electrodes (303). In contrast to DBS, the efficacy of cortical stimulation to treat neurologic disorders has not yet been definitively established and many research questions remain unanswered regarding the mechanisms of therapeutic efficacy, modes of stimulation, relevant neurologic diagnoses, and meaningful motor impairment, activity limitation, and quality of life outcome measures.

TMS was developed in 1985 as a noninvasive method of exciting cortical neurons using electromagnetic induction. Rapidly changing magnetic fields are applied transcranially to induce focused electric currents in the underlying cerebral cortex using an electromagnetic coil (round, figure-eight or double-cone configuration) applied directly to the scalp. With stereotactic MRI-based control, the precision of targeting TMS can be approximated to a few millimeters (304). Repetitive transcranial magnetic stimulation (rTMS) is being trialed as a method to induce sustained therapeutic effects in patients with movement disorders (305), stroke (306), neuropathic pain (303,307), depression (308), and various neuropsychiatric disorders. By manipulating stimulation frequency, intensity and coil orientation, rTMS is presumed to influence the excitability of cortical pathways by altering synaptic transmission (309). Low frequency (1 Hz) rTMS decreases brain activation while stimulation at higher frequencies (>5 Hz) increases brain activation. A rare but present risk of TMS is induction of a seizure. In depression studies, no cognitive side-effects including memory loss, concentration difficulties or loss of other cognitive capacities have been reported. Multiple recent reviews (305,310–312) have summarized the theoretical basis of TMS and research and potential therapeutic clinical applications of TMS.

Transcranial direct current stimulation (tDCS) is the noninvasive application of weak direct currents to induce sustained changes in cortical activity and excitability via the hyper- or depolarization of resting membrane potentials (313). Transcranial DCS has been demonstrated to influence the excitability of the human motor cortex in research applications (314,315). In a study of seven healthy adults, Vines et al. (316) found that stimulating a motor region directly, or indirectly by modulating activity in the homologous region on the opposite hemisphere, affected the motor skill acquisition by a proposed mechanism of facilitating synaptic connectivity. Lang et al. (317) studied the effect of tDCS of the primary motor cortex on regional neuronal activity in a controlled study of 16 healthy volunteers. When compared to sham tDCS, 10 minutes of anodal and cathodal tDCS induced widespread changes in cortical and subcortical rCBF as measured by positron emission tomography (PET) scanning. The therapeutic application of tDCS is being investigated for various neurologic disorders including movement disorders (318), motor recovery poststroke (319), epilepsy (320), depression (321,322), and tinnitus (323). Recent reviews (2,313,324) provide an overview of present understanding and potential applications of tDCS. At present, tDCS is limited to research applications but appears to be a promising neuromodulatory technique that may ultimately benefit the field of clinical neurorehabilitation.

ECS is a minimally invasive technique which involves the application of low-voltage, high frequency stimulation to the motor cortex via chronically implanted epidural electrodes. ECS was first described by Tsubokawa et al. (325,326) for treatment of central (thalamic) and neuropathic pain. Since then, ECS has been demonstrated to be effective for pain associated with trigeminal neuropathy, lateral medullary and thalamic infarction, anesthesia dolorosa, postherpetic neuralgia, SCI, and postamputation phantom pain (327). Preoperative fMRI, 3-dimensional volume MRI, neuronavigation and intraoperative neurophysiological monitoring have improved the technique for identifying the targeted motor cortical area and for optimum placement of the electrode array (328). Once the epidural electrode has been activated, stimulation-induced pain relief may occur within minutes, with pain relief lasting up to hours following discontinuation of ECS (327). In a study of the effect of ECS on neuropathic pain, Drouot et al. (329) suggested that ECS acts on neural pathways involved in sensory discrimination to modulate transmission of pain signals. Saitoh et al. (330) found an increase in rRCB in the contralateral thalamus following motor cortex stimulation in a
patient with poststroke pain. Similarly, Brown (327) proposed that the efficacy of ECS in central pain is based on an increase in rCBF in the cingulated gyrus and the ipsilateral ventrolateral thalamus in which corticothalamic connections from the motor and premotor areas predominate. A recent review (328) concluded that ECS is increasingly reported as an effective surgical option for the treatment of various forms of refractory neuropathic pain.

Recently, there has been interest in ECS as a treatment for movement disorders such as PD, particularly for patients who are excluded from the more invasive DBS procedure. The first application of chronic epidural motor cortex stimulation for PD was reported by Canavero and Paolotti (331) in 2000. In an animal model of PD (332), exposure to intermittent, high intensity cortical stimulation of the primary motor cortex via epidural electrodes resulted in a temporary reduction in the severity of PD symptoms. In a human study (333), 16 subjects with PD who underwent ECS via chronically implanted electrodes demonstrated sustained improvement of a broad spectrum of symptoms including tremor, rigor, akinesia, axial symptoms, speech and swallowing, and reduction in levodopa dosage. A recent review of extradural cortical stimulation for movement disorders concluded that potential mechanisms of action of ECS in PD include ‘hyperdirect’ motor cortex-STN input, inhibition, resynchronization, plasticity changes, inter-hemispheric transfer of inhibition/excitation and modulation of other cortical areas (334). Important unresolved issues for broader clinical application of ECS in PD include optimal electrode position and stimulation parameters, the benefit of unilateral versus bilateral stimulation, and refinement of patient selection criteria (335).

ECS is also being studied for its ability to affect neuroplasticity to facilitate poststroke motor recovery (336). In an animal model, Adkins et al. (337) implanted chronic epidural electrodes over the motor cortex of animals that had sustained a subtotal ischemic lesion of the sensorimotor cortex. The animals were divided into two subgroups based on the level of motor impairment and randomized to receive 18 days of rehabilitative training with either 100 Hz cathodal ECS at 50% of movement thresholds or no stimulation. ECS partially normalized the movements, and a significantly greater density of axodendritic synapses was noted in both impairment subgroups in the ECS group as compared to the no stimulation group. The moderately impaired ECS animals had increases in presumed efficacious synapse subtypes (perforated and multiple synapses) as compared to the no stimulation group. Synaptic density was positively correlated with postrehabilitation reaching success. The authors suggested that ECS promoted functional recovery and that ECS-induced functional improvements may be mediated by synaptic structural plasticity in stimulated cortex. Early human studies demonstrated the feasibility of ECS among chronic stroke survivors (338). A follow-up phase II RCT by Levy et al. (339) suggested efficacy with improvements in upper limb motor impairment and activities limitation. However, the Everest Clinical Trial (340), a phase III RCT to test safety and efficacy of ECS delivered during rehabilitation for upper limb motor function, failed to corroborate earlier findings.

CONCLUSION

In this chapter, we have summarized the history, mechanisms, and clinical applications of ES to treat the neurologically based disorders most commonly encountered in the clinical practice of Physical Medicine and Rehabilitation. The application of ES to treat neurologic disease has a long history dating to ancient Egyptian times. Present day clinical applications of ES are broadly defined as providing either a therapeutic or functional benefit. The goal of therapeutic ES is to achieve a sustained effect that persists beyond the period of intervention. Within the field of Physical Medicine and Rehabilitation, therapeutic ES is commonly prescribed to both decrease the impairment and prevent further impairment and activity limitation associated with immobility or disuse of a limb or organ system.

At the dawn of the 21st century, the efficacy and safety of therapeutic ES has been established in a variety of neurologically based disorders including chronic pain, paralysis, and movement disorders. Therapeutic ES has the capacity to treat disease, decrease impairment, and enhance quality of life for patients who have sustained neurologic injury. The establishment of therapeutic efficacy of a broad spectrum of transcutaneous ES devices as well as implanted neuroprostheses and neuromodulators clearly expands the armamentarium of the physiatrist in the clinical care of patients. Future research will undoubtedly focus on further defining the patient populations, applications and mechanisms by which ES may impart therapeutic benefit. Ongoing basic science and clinical research offers promise for therapeutic benefit of ES for an even broader spectrum of neurologically based disorders.

ACKNOWLEDGMENTS

The preparation of this chapter was supported, in part, by the National Institute of Child Health and Human Development (NICHD), grant K24HD054600.

REFERENCES


10. Wall PD, Sweet WH. Temporary abolition of pain in man.


101. Kang BS, Bang MS, Jung SH. Effects of botulinum toxin A therapy with
111. Li M. Electrical stimulation in the treatment of osteoporosis in sciatic
112. Lirani-Galvao AP, Bergamaschi CT, Silva OL, et al. Electrical field stim-
114. Midura RJ, Dillman CJ, Grabiner MD. Low amplitude, high frequency
96. Carnstam B, Larsson LE, Prevec TS. Improvement of gait follow-
115. Park SH, Silva M. Neuromuscular electrical stimulation enhances
1970
125. Peckham PH, Mortimer JT, Van Der Meulen JP. Physiologic and meta-
133. Carvalho de Abreu DC, Junior AC, Rondina JM, et al. Muscle hyper-
120. Eser P, de Bruin ED, Telley I, et al. Effect of electrical stimulation-in-
126. Pette D, Muller W, Leisner E, et al. Time dependent effects on con-
128. Martin TP, Stein RB, Hoeppner PH, et al. Influence of electrical stimula-
132. Baldi JC, Jackson RD, Moraille R, et al. Muscle atrophy is prevented in
136. Dupont Salter AC, Richmond FJ, Loeb GE. Prevention of muscle dis-
137. Peckham PH, Mortimer JT, Marsolais EB. Alteration in the force and
121. Fronzler A, Coupaud S, Perret C, et al. High-volume FES-cycling par-
129. Sanderson JE, Biedert RM, Biedert RM. Muscle atrophy is prevented in
130. Zehnder AM, Biedert RM. Muscle atrophy is prevented in incomplete
131. Carvalho de Abreu DC, Junior AC, Rondina JM, et al. Muscle hyper-
140. Dupont Salter AC, Richmond FJ, Loeb GE. Prevention of muscle dis-
141. Sanderson JE, Biedert RM. Muscle atrophy is prevented in incomplete
142. Carvalho de Abreu DC, Junior AC, Rondina JM, et al. Muscle hyper-
144. Dupont Salter AC, Richmond FJ, Loeb GE. Prevention of muscle dis-
146. Ljubicic V, Adhijerry P, Hood DA. Application of animal models:
chronic electrical stimulation-induced contractile activity. Can J Appl
147. Bright CT, Friedenberg ZB, Black J, et al. Electrically induced osteo-
geneis: relationship between charge, current density, and the amount of
denervation disuse osteoporosis in the rat tibia with capacitively coupled
149. Bright CT, Luessenhop CP, Pollack SR, et al. Treatment of castration-
induced osteoporosis by a capacitively coupled electrical signal in rat
150. Bright CT, Taddini GT, Pollack SR. Treatment of sciatric denervation
disuse osteoporosis in the rat tibia with capacitively coupled electrical
151. Li M. Electrical stimulation in the treatment of osteoporosis in sciatric
152. Lirani-Galvao AP, Bergamaschi CT, Silva OL, et al. Electrical field stim-
ulation improves bone mineral density in ovariectomized rats. Braz J
stimulation on bone mineral density in the paralyzed limbs of the rabbit.
154. Midura RJ, Dillman CJ, Grabiner MD. Low amplitude, high frequency
strains imposed by electrically stimulated skeletal muscle retards the
development of osteopenia in the tibiae of hindlimb suspended rats.
155. Park SH, Silva M. Neuromuscular electrical stimulation enhances
382–387.
156. Clark JM, Jelbart M, Rischbieth H, et al. Physiological effects of
deeper extrema functional electrical stimulation in early spinal cord
injury: lack of efficacy to prevent bone loss. Spinal Cord. 2007;45:
78–85.


292. Schiff ND, Fins JJ. Deep brain stimulation and cognition: moving from discussion 74–76.


296. Schiff ND, Fins JJ. Deep brain stimulation and cognition: moving from discussion 74–76.


299. Schiff ND, Fins JJ. Deep brain stimulation and cognition: moving from discussion 74–76.


301. Schiff ND, Fins JJ. Deep brain stimulation and cognition: moving from discussion 74–76.


INTRODUCTION

Neuromuscular electrical stimulation (NMES) refers to the use of low-level electrical current to produce a muscle contraction. In the field of Physical Medicine and Rehabilitation, NMES is used in cases of neurological injury that result in muscle paralysis or paresis. Clinical applications of NMES provide either a therapeutic or functional benefit. Therapeutic NMES applications are intended to reduce specific impairments or to induce or hasten the recovery of volitional motor function and, therefore, are temporary or periodic interventions. Usually, therapeutic NMES applications do not directly provide motor function; provision of motor function is the purpose of functional NMES applications, which are the focus of this chapter. Functional NMES applications produce coordinated contractions of multiple muscles that enable the motor-impaired individual to accomplish functional tasks directly. A functional NMES device is called a neuroprosthesis, because it substitutes for or replaces lost neuromuscular function and is considered to be an ongoing or permanent intervention. Ideally, a neuroprosthesis should be easy to wear all day so that its utility is available whenever needed. This chapter reviews the principles of NMES and neuroprosthetic systems for several functional NMES applications.

PRINCIPLES OF NMES

Electrical Activation of the Neuromuscular System

Action potentials, and therefore muscle contractions, can be generated with electrical current. Electrical current, applied to excitable tissue through a pair of electrodes, creates a localized electric field that depolarizes the axonal membranes of nearby neurons. If the depolarization reaches a critical threshold, an influx of sodium ions from the extracellular space to the intracellular space produces an action potential that propagates from the site of stimulation and has the same effect as a naturally generated action potential. When the action potential reaches the terminals of the axon, neurotransmitter is released and, in the case of motor neurons, muscle fibers contract.

NMES systems generally operate by activating motor neurons rather than the muscle directly, even though NMES is commonly referred to as “muscle stimulation.” It is easier to stimulate a nerve than a muscle. The stimulus threshold for a neural tissue, the lowest level of charge that will generate an action potential, is much less than the threshold for muscle fibers (1). While it is possible to generate action potentials in muscle fibers directly, it is not usually done because of power consumption considerations. This means that for NMES to be applicable in a given patient, their lower motor neurons must be intact from the anterior horns of the spinal cord to the neuromuscular junctions of the target muscles.

Damage to lower motor neurons prevents the successful application of NMES in patients with polio, late-stage amyotrophic lateral sclerosis (ALS), and peripheral nerve injuries (e.g., brachial plexus injury). In addition, patients with disorders at the neuromuscular junction or muscle tissue (e.g., muscular dystrophies) are ineligible for NMES applications. NMES may be used when the lower motor neurons are excitable and the neuromuscular junction and muscle are healthy, as is usually the case in patients with spinal cord injury (SCI), stroke, traumatic brain injury, cerebral palsy, or multiple sclerosis. To date, most motor neuroprostheses have been targeted toward the SCI population.

Large axons are easier to stimulate than small axons. Large-diameter axons, which innervate the larger motor units, have lower stimulus thresholds than small-diameter axons (2). This is because the wider spacing between the nodes of Ranvier in large axons produces larger induced transmembrane voltage changes. The activation of large motor units before small motor units is known as reverse recruitment order and is the opposite of the physiological size principle, where small motor units are initially recruited, followed by larger motor units. In spite of this reverse recruitment order, it is still possible to produce graded muscle contractions with NMES.

Because large-diameter axons have lower stimulus thresholds, type II muscle fibers (fast-fatiguing) are preferentially recruited over type I muscle fibers (fatigue-resistant). Fatigue resistance is highly desirable for most functional NMES applications. Unfortunately, disuse atrophy tends to convert type I to type II fibers (3). However, long-term use of NMES can reverse muscle atrophy and convert type II fast-fatiguing fibers to type I fatigue-resistant fibers (4,5). Therefore, all current neuroprosthetic NMES applications use some form of muscle conditioning regimen to build and maintain fatigue-resistant muscle.
Axons near the stimulus source are easier to stimulate than axons farther away. The electric field diminishes with distance from the stimulus source; as the distance between the electrode and the axonal fiber increases, the stimulus threshold also increases (1). Less current is required to stimulate neurons in the proximity of the stimulating electrode because the transmembrane potentials generated by the electrical current are largest in the axons close to the stimulating electrode. Therefore, to improve the selectivity of stimulation, electrodes are placed as close as possible to target muscles or nerves.

The strength of a muscle contraction produced by NMES can be modulated by manipulating three stimulus parameters that characterize the wave of current pulses: pulse frequency, amplitude, and duration. If the pulse frequency is too low, the muscle responds with a series of twitches. As the pulse frequency is increased, the twitches begin to overlap and build on one another, producing a smooth or “fused” contraction. This cumulative effect of repeated stimulus pulses within a brief period of time is known as temporal summation. Higher stimulus frequencies produce stronger muscle contractions up to a maximum, but also make a muscle fatigue more rapidly than lower-frequency stimulation. Therefore, the ideal stimulus frequency is the lowest frequency that will produce a fused contraction, and depends on the muscle fiber type and the manner in which the stimulation is being delivered (e.g., surface vs. implanted electrode). In upper extremity applications that use implanted electrodes, a stimulus frequency of 12 to 16 Hz has been found to be the minimum required for producing fused contractions if the muscles have been conditioned to have relatively long-duration twitches. The strength of a muscle contraction may also be increased by increasing the pulse amplitude and/or pulse duration, which effectively increases the electric charge per pulse (6). The greater the electric charge, the larger the electric field and broader the region of activation. In this way, more axons and more motor units are activated, an effect known as spatial summation. In most neuroprostheses, the strength of muscle contraction is controlled by modulating the pulse amplitude or pulse duration while keeping the pulse frequency constant and as low as possible to avoid premature fatigue of muscles.

Safe Stimulation of Living Tissue

Safe stimulation waveforms and electrode materials have been experimentally established. For NMES applications using implanted electrodes, a current-regulated stimulator with a balanced biphasic stimulus waveform should be used. With implanted electrodes it is important to use stimulus parameters (pulse amplitudes and durations) that are appropriate for the dimensions and material composition of the electrode so that the charge density per phase remains within the established safe limits, thereby preventing electrode corrosion or dissolution of metal ions (1). With current-regulated stimulation, the current is directly controlled (current is the charge delivered per unit time). Therefore, the quantity of charge delivered per stimulus pulse can be maintained within safe limits. Improper stimulation can also cause tissue damage by producing irreversible electrochemical reactions at the electrode-tissue interface. Irreversible reactions can be avoided with charge-balanced biphasic stimulus waveforms. Biphasic waveforms consist of a repeating current pulse that has a cathodic (negative) phase followed by an anodic (positive) phase. The first, or primary, phase elicits an action potential in nearby axons, and the secondary positive pulse balances the charge of the primary pulse. The purpose of the secondary pulse is to reverse the potentially damaging electrochemical processes that occur at the electrode-tissue interface during the primary pulse, allowing neural stimulation without causing tissue damage (1).

For NMES applications using surface electrodes, other safety factors need to be considered. For example, the electrode-tissue contact area is often not a constant as surface electrodes often pull away from the skin or dry out. If a current-regulated stimulator is used in this circumstance, the reduction in contact area will cause high current densities and can result in burning the skin and underlying tissue. However, with a voltage-regulated stimulator, the magnitude of current delivered to the tissue is dependent on the impedance at the electrode interface (current = voltage/impedance). If the impedance at the electrode-skin interface increases because the electrode dries or loses strong contact with the skin, then the current delivered will decrease, thereby reducing the possibility of skin burns due to high current densities. Although this may be safer, a major disadvantage of voltage-regulated stimulation is that the muscle contraction produced is more variable because the changes in impedance lead to changes in applied current. Regardless of whether current- or voltage-regulated stimulators are used, even with good electrode-skin contact, burning of the tissue can occur if the current densities are too high. Therefore, surface stimulation should be used with caution and with frequent examination of the skin when applied to patients with impaired sensation or cognition. Table 72-1 summarizes the principles of NMES and their practical ramifications when applying NMES systems.

NMES System Configurations and Electrode Types

At least two electrodes are required to produce a current flow in NMES applications. One electrode, often referred to as the active electrode or cathode, is placed near the peripheral nerve or muscle motor point to be stimulated. The other electrode, known as the return electrode or anode, may be placed in a remote area near less excitable tissue (such as tendon or fascia), or it may be positioned close to the active electrode to confine the electric field to a small region of activation and thereby achieve more selective muscle activation. Additional electrodes are needed to achieve coordinated activation of multiple muscles. Multichannel NMES systems stimulate multiple muscles simultaneously and use either a bipolar or monopolar arrangement of electrodes. Bipolar multichannel systems require a return electrode for each active electrode; so for each muscle or nerve to be activated, a pair of electrodes (cathode and anode) is needed. This bipolar arrangement may allow greater selectivity of activation (7), but requires more electrodes and leads. Monopolar electrode systems reduce the number of electrodes and leads by...
using only one remote return electrode with an active electrode placed near each motor point or nerve targeted for excitation.

Functional NMES systems consist of at least three components: electrodes, a multichannel stimulator, and a controller (Fig. 72-1). The electrodes deliver the electrical current pulses to the excitable tissue, the stimulator generates the current waveforms for multiple cathodes, and the controller regulates the stimulation according to the user’s intent. A controller may be as simple as a switch that the patient hits to turn the stimulation on and off, or may incorporate sensors for recording patient-elicited biopotentials (e.g., EMG or EEG) and use those as signals to regulate the stimulation. Depending on the way these three components are arranged with respect to one another, NMES systems can be categorized as surface, percutaneous, or implanted. In surface systems, the electrodes, the stimulator, and the controller are external to the body. In percutaneous systems, the electrodes are implanted near a muscle or nerve with leads that pass through the skin to an external stimulator with a separate or integrated external controller. In implanted systems, the electrodes and stimulator are implanted and the controller components are either all external or some are implanted and some are external. To date, there is no neuroprosthesis available that has all three components completely implanted, although several research programs are moving in that direction (8,9).

Surface NMES systems (sometimes called transcutaneous systems) use electrodes that are placed on the skin and connected with flexible cables to a stimulator, which may be worn around the waist, the arm, or the leg. Usually, a sensor or switch that controls the stimulation is also connected to the stimulator. Surface electrodes are readily available in a variety of sizes from many manufacturers. The electrodes are placed on the skin over the nerves or over the “motor points” of muscles

<table>
<thead>
<tr>
<th>Principle</th>
<th>Practical Ramification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Action potentials can be produced by NMES. It is easier to stimulate nerve than muscle.</td>
<td>Paralyzed muscles can be made to contract. NMES is generally applicable only to patients with intact lower motor neurons.</td>
</tr>
<tr>
<td>Large axons are easier to stimulate than small axons.</td>
<td>A muscle conditioning regimen is needed to build and maintain fatigue-resistant muscle.</td>
</tr>
<tr>
<td>Axons near the stimulus source are easier to stimulate than axons farther away.</td>
<td>Selectivity of stimulation is improved by placing electrodes close to the target muscles or nerves.</td>
</tr>
<tr>
<td>Applying closely spaced stimulus pulses to a muscle produces an overlapping of muscle twitches, resulting in a “fused” contraction—temporal summation.</td>
<td>The strength of a muscle contraction can be modulated by increasing or decreasing the stimulus pulse frequency. However, very high stimulus frequencies cause rapid muscle fatigue.</td>
</tr>
<tr>
<td>More axons and motor units of a muscle are activated when the size of the electrical field is increased by increasing the electric charge per stimulus pulse—spatial summation.</td>
<td>The strength of a muscle contraction can be modulated by increasing or decreasing the stimulus pulse amplitude and/or duration.</td>
</tr>
<tr>
<td>With implanted electrodes, electrode corrosion can be avoided by maintaining the charge density per phase of stimulus pulse within safe levels determined for the particular electrode material composition and dimensions.</td>
<td>Current-regulated stimulators should be used with implanted electrodes.</td>
</tr>
<tr>
<td>With implanted electrodes, tissue damage can occur if irreversible electrochemical reactions are produced at the electrode-tissue interface.</td>
<td>Charge-balanced biphasic stimulus waveforms should be used with implanted electrodes.</td>
</tr>
<tr>
<td>With surface electrodes, high current densities can cause skin burns.</td>
<td>Surface electrodes that maintain good contact with the skin should be used. Extra caution is needed when applying NMES to patients with reduced sensation or impaired cognition.</td>
</tr>
</tbody>
</table>

TABLE 72.1  Principles of Safe Electrical Activation of the Neuromuscular System and Practical Ramifications
to be activated. The motor point is the site of stimulation that produces the strongest and most isolated contraction at the lowest level of stimulation. Surface NMES systems offer several distinct advantages: (a) the electrodes are generally easy to apply and remove, (b) the stimulation technique is noninvasive and therefore reversible, (c) the use of surface electrodes can be readily learned and applied in the clinic, and (d) stimulators and surface electrodes are relatively inexpensive and commercially available. Stimulation with surface electrodes is the most widely used technique for therapeutic applications, and has been successfully employed to produce standing, stepping, and grasping motions. However, there are several disadvantages to using surface NMES systems: (a) they cannot produce isolated contractions of small or deep muscles, (b) movement of the target muscle under the skin as the stimulated limb moves can change the proximity of the electrode to the target muscle/nerve, resulting in inconsistent muscle contraction and force production, (c) daily donning and doffing the electrodes can complicate use, especially if electrode positions vary slightly from day to day, (d) in many cases, cutaneous pain receptors are stimulated and patients with preserved or heightened sensation may find it difficult to tolerate, and (e) the system may draw unwanted attention, especially if there are multiple stimulus channels and many cables. These disadvantages have motivated the design of systems with implanted electrodes.

Percutaneous systems use electrodes that are implanted into muscles or near nerves, and have leads that pass through the skin and connect to an external stimulator. Percutaneous electrodes can activate deep muscles, can provide isolated and repeatable muscle contractions, and are less likely to produce pain during stimulation because they bypass the sensory afferents in the skin. Percutaneous electrodes are often formed from a multifilament lead within a single insulator that is wound into a helical configuration to produce maximum flexibility (10, 11). Percutaneous intramuscular electrodes are inserted into the target muscle with a hypodermic needle. The needle is withdrawn, leaving the electrode in place and the lead exiting the skin. The exit sites on the skin are protected with a bandage and must be cleaned, dressed, properly inspected, and maintained to reduce the risk of complications (12). A large surface electrode is often used as the return electrode. Percutaneous systems provide a minimally invasive technique for investigating the feasibility of restoring functional muscle contractions without having to prematurely subject research participants to implantable system surgery. Percutaneous systems have served as precursors to fully implanted systems (13) and have provided function in some individuals for periods of years (14). Various studies have reported on the longevity of percutaneous electrodes (11, 12, 15–17). Although the failure rate due to breakage is low during the first few months post implantation, the cumulative 1-year failure rate can vary between 56% and 91% (12, 17), but this depends on many factors including the type of percutaneous lead, lead-routing technique, and the muscles implanted. Other potential complications of percutaneous electrodes include formation of granulomas from retained electrode fragments and electrode-related infections, which are treated with oral antibiotics or minor outpatient surgical procedures (12).

Implanted NMES systems are designed for long-term use. Both the electrodes and the stimulator are implanted, but the controller is external for most implanted systems, to date. Communication between the controller and stimulator is through radio-frequency (RF) transmission; therefore, no leads pass through the skin. Because the stimulator both receives RF commands and generates stimulus pulses, it is often called a receiver/stimulator. Implanted electrodes with insulated lead wires are connected directly to the implanted stimulator with inline connectors, which permit the surgical removal and replacement of individual electrodes, if necessary, without removing the stimulator. The electrode leads are larger than percutaneous leads because they need to be more robust and resistant to failure given the intention to make long-term use of them. In most implanted system configurations the circuitry of the stimulator is sealed in a titanium enclosure, which serves as the return electrode. Depending on the application, the stimulator is implanted in the upper chest or abdomen. The stimulator receives power and control instructions through RF telemetry from an external controller. The RF link allows the stimulator to be fully passive with no active battery, thereby eliminating the need for replacement of the implanted stimulator because of battery failure. The telemetry link requires no wires through the skin; rather, a circular coil (antenna) connected to the controller with a cable is taped to the skin over the implanted stimulator. The controller receives information from the user through switches, sensors, an implanted joint angle sensor, or biopotential electrodes to determine the user’s intention and deliver stimulation accordingly. The controller may be worn on the body or carried on a user’s wheelchair. For long-term clinical application, implanted systems provide major advantages over other systems including improved convenience, cosmesis, reliability, and repeatability.

A variety of electrodes may be used with implanted NMES systems (Fig. 72-2). Epimysial electrodes are sutured directly to the epimysium on the muscle surface (18); intramuscular electrodes are inserted directly into a muscle belly (19); epineural electrodes are sutured to the connective tissue surrounding a motor nerve (20); and nerve helix or cuff electrodes are implanted around a nerve (21–24). Epimysial electrodes have proved to be durable in upper and lower extremity applications (25, 26), and are especially useful for activating broad, superficial, or thin muscles. Intramuscular electrodes allow activation of deeper or smaller muscles, such as the intrinsic muscles of the hand. Nerve-based electrodes are used when it is difficult to access the target muscle directly or when more complete muscle recruitment can be obtained by stimulating the nerve (27).

A unique configuration of implanted NMES system is the injectable microstimulator (BION), which is presently undergoing clinical trials for various therapeutic and neuroprosthetic applications (28). The BION (developed at the A.E. Mann Institute at the University of Southern California) is a small cylindrical (2 mm outer diameter × 16 mm length), injectable,
single-channel unit that functions as a receiver/stimulator and electrode (29). BIONs are implanted with a 12-gauge trochar in target muscles or near target nerves; the quantity to be implanted depends on the application. An external RF coil transmits power and control information from an external control unit. Battery-powered BIONs and EMG-sensing micro-injectable units are presently under development. These technologies and principles of electrical stimulation form the foundation for clinical applications of NMES. The next section will describe neuroprosthetic systems that have been developed for different upper motor neuron impairments.

**CLINICAL APPLICATIONS OF NMES**

This section describes the neuroprosthetic systems for four different clinical applications: upper extremity function, lower extremity function, respiratory function, and bladder and bowel function. These four applications are the most developed, have undergone the most clinical testing, and have several different device options available or in development. For each application, the purposes, indications, and specific neuroprosthetic systems will be described. This is not an exhaustive review of every device that has appeared in the research literature, rather only those systems that have undergone clinical trials and are currently commercially available or devices that are currently in the clinical trial phase of development (i.e., may be available through participation in research) will be discussed and are summarized in Table 72-2 at the end of this section. Finally, for each application, new technology developments and future research directions will be outlined.

**Upper Extremity Function**

The purpose of upper extremity neuroprosthetic systems is to enable individuals to use their paralyzed or paretic arm and hand to perform activities of daily living (ADL) (e.g., eating, personal hygiene, etc.). In particular, these systems provide the ability to grasp and release with the impaired hand and thereby reduce the need of individuals to rely on assistance from others and reduce the time and effort required to perform a task. In addition to hand grasp and release, emerging systems are incorporating elbow extension with shoulder abduction and flexion to provide reaching ability to individuals with proximal and distal upper limb paralysis.

Upper extremity systems may be applicable to people with partial or complete paralysis of the arm and/or hand resulting from cervical SCI or stroke. The hand and forearm muscles targeted for stimulation must not be denervated, that is, the lower motor neurons must be intact. Many persons with C5 and C6 level tetraplegia have preserved lower motor neurons from C7 and C8 neurological segments and may benefit from upper extremity NMES systems that restore grasp and release. Persons with C4 or higher level tetraplegia who have C5 and C6 lower motor neurons preserved may benefit from NMES systems designed to provide shoulder stability and control of elbow flexion and extension, in addition to providing grasp and release. Any joint contractures must be corrected or functional ability will be limited. Spasticity and hypertonia must be under control. The best candidates are individuals with intact cognition, who are motivated, and desire greater independence. In addition, many neuroprosthetic systems require assistance in donning the device, so it may be necessary for the individual to have good attendant support.

The first upper extremity neuroprostheses were developed in the 1960s using surface electrodes, in combination with a flexor hinge splint to open and close the hands of individuals with cervical SCI (30,31). These case studies were followed by work with stroke survivors in the 1970s (32,33). One- and two-channel stimulators activated the finger and thumb extensors and triceps (in some subjects) when the individual elevated their opposite shoulder, a movement that was picked up by a shoulder-mounted transducer. These pioneering efforts have led to the development and clinical testing of several upper extremity neuroprostheses for stroke and spinal cord injured patients (34–41), some of which have been primarily intended for use as temporary therapeutic devices (see Chapter 71). The NESS H200 (formerly known as Handmaster) is the only commercially available upper extremity neuroprosthesis (Bioness Inc., Valencia, CA) at the time of this writing. The H200 was originally developed for tetraplegia (42), but is also applicable to stroke survivors (43). The device is a wrist-forearm orthosis (Fig. 72-3), with five embedded surface electrodes that provide patterned stimulation to the finger and thumb flexors and extensors to produce selected hand movements. The
TABLE 72.2 Summary of Available Functional NMES Systems for Four Applications

<table>
<thead>
<tr>
<th>System Name</th>
<th>Supplier or Research Lab</th>
<th>System Type</th>
<th>FDA Status</th>
<th>CE Mark</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Upper extremity function</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NESS H200</td>
<td>Bioness Inc., Valencia, CA</td>
<td>Surface</td>
<td>510(k)</td>
<td>Yes</td>
</tr>
<tr>
<td>IST12 hand system</td>
<td>CWRU, Cleveland, OH</td>
<td>Implant</td>
<td>IDE</td>
<td>No</td>
</tr>
<tr>
<td><strong>Lower extremity function: footdrop</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NESS L300</td>
<td>Bioness Inc., Valencia, CA</td>
<td>Surface</td>
<td>510(k)</td>
<td>Yes</td>
</tr>
<tr>
<td>WalkAide</td>
<td>Innovative Neuromatics Inc., Austin, TX</td>
<td>Surface</td>
<td>510(k)</td>
<td>Yes</td>
</tr>
<tr>
<td>Odstock dropped-foot stimulator (ODFS)</td>
<td>NDI Medical, Cleveland, OH</td>
<td>Surface</td>
<td>510(k)</td>
<td>Yes</td>
</tr>
<tr>
<td>STIMuSTEP</td>
<td>FineTech Medical Ltd, Hertfordshire, UK</td>
<td>Implant</td>
<td>None</td>
<td>Yes</td>
</tr>
<tr>
<td>ActiGait</td>
<td>Neurodan A/S, Aalborg, Denmark</td>
<td>Implant</td>
<td>None</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Lower extremity function: stand/transfer</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CWRU/VA standing system</td>
<td>CWRU, Cleveland, OH</td>
<td>Implant</td>
<td>IDE</td>
<td>No</td>
</tr>
<tr>
<td><strong>Lower extremity function: walk</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parastep system</td>
<td>Sigmedics Inc., Fairborn, OH</td>
<td>Surface</td>
<td>PMA</td>
<td>No</td>
</tr>
<tr>
<td>CWRU/VA walking system</td>
<td>CWRU, Cleveland, OH</td>
<td>Implant</td>
<td>IDE</td>
<td>No</td>
</tr>
<tr>
<td><strong>Respiratory function</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Avery Mark IV</td>
<td>Avery Biomedical Devices, Commack, NY</td>
<td>Implant</td>
<td>PMA</td>
<td>Yes</td>
</tr>
<tr>
<td>Atrostim</td>
<td>Atrotech Ltd, Tampere, Fineland</td>
<td>Implant</td>
<td>IDE</td>
<td>Yes</td>
</tr>
<tr>
<td>NeuRx DPS system</td>
<td>Synapse Biomedical Inc., Oberlin, OH</td>
<td>Implant</td>
<td>PMA</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Bladder/bowel function</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FineTech Brindley Bladder control system</td>
<td>FineTech Medical Ltd, Hertfordshire, UK</td>
<td>Implant</td>
<td>PMA</td>
<td>Yes</td>
</tr>
</tbody>
</table>

510(k): Also known as a premarket notification, the applicant has demonstrated that the device to be marketed is at least as safe and effective, that is, substantially equivalent, to a legally marketed device that is not subject to PMA, and therefore has been cleared to market the device.

IDE: Investigational device exemption allows an investigational device to be used in a clinical study in order to collect safety and effectiveness data required to support a premarket approval (PMA) application or a Premarket Notification (510(k)) submission to FDA.

PMA: Premarket approval, the FDA process of scientific and regulatory review to evaluate the safety and effectiveness of Class III medical devices. (Class III devices are usually those that support or sustain human life, are of substantial importance in preventing impairment of human health, or which present a potential, unreasonable risk of illness or injury.) PMA is the most stringent type of device marketing application required by FDA.

CE: a mandatory European marking for certain product groups to indicate conformity with the essential health and safety requirements set out in European Directives; indicates that the product may be legally placed on the market within the European Free Trade Agreement (EFTA) and European Union (EU) single market (total 28 countries).

IDE applied to Version 1.0 of the Atrostim device.

The implanted stimulator-telemeter 12 channel (IST12) system is an implanted upper extremity neuroprosthesis that is presently used in clinical trials conducted by researchers at Case Western Reserve University in Cleveland (45). This group has been developing implanted upper extremity neuroprostheses for SCI patients since 1986, when they implemented the first implanted hand grasp system. The IST12 system is an advanced version of the Freehand system, which underwent extensive assessment of clinical outcomes, received FDA approval (premarket approval, PMA) for use in C5 and C6 tetraplegia, and was commercially available from 1997 to 2002. Like the Freehand system, the IST12 system consists of a receiver/stimulator that is surgically implanted in the upper pectoral region and epimysial, and intramuscular electrodes that are implanted at the motor points of hand and forearm muscles with leads that are routed subcutaneously to the stimulator (Fig. 72-4). The external components include a RF transmitting coil that is taped to the chest over the stimulator and connected to a programmable external control unit. This implanted system approach was shown, in the
CHAPTER 72 | FUNCTIONAL NEUROMUSCULAR ELECTRICAL STIMULATION

The IST12 system builds on the clinical success of its predecessor, the Freehand system, by providing additional upper limb function and incorporating implanted EMG control methods. Enhanced function is achieved through additional stimulation channels. The Freehand system had eight stimulus channels (48); the IST12 system has 12. The additional channels are used to activate hand intrinsic muscles, triceps, or pronator quadratus. These additional muscles have resulted in better hand opening, improving object acquisition and release; elbow extension, increasing the user’s workspace; and forearm rotation, improving the user’s ability to optimally orient the hand. In the original Freehand system, the user controlled the opening and closing of their paralyzed hand by moving their opposite shoulder, which required wearing an external shoulder-mounted transducer. With the IST12 system, EMG control strategies are available, where the use of EMG signals are recorded from muscles that the patient still retains the ability to contract and relax. The EMG signals are picked up by implanted EMG-recording electrodes, a strategy that eliminates the need for donning an external sensor and opens the possibility of achieving more natural and intuitive control. Potential control muscles include those that are synergistic to the grasp movement, such as wrist extensors and, to a lesser extent, the brachioradialis, and nonsynergists such as sternocleidomastoid, trapezius, deltoid, platysma, or auricularis. Because the EMG control strategy makes use of muscle signals derived from the ipsilateral side, it allows the neuroprosthesis to be implemented bilaterally, which provides significant additional function to patients with tetraplegia.

The IST12 system is undergoing clinical testing; to date, 15 individuals with tetraplegia have received the system. Four of these participants have two systems implanted, one for each upper limb, in order to enable bimanual task performance. The longest follow-up has been 7 years. The functional results show that the neuroprosthesis provides significantly increased pinch force and grasp function for each subject (45).

multicenter clinical trial of the Freehand system, to be successful in significantly reducing upper extremity impairment and activity limitation (46), and to be well accepted by the recipients (47), with very low rates of infection (<2%) and implant or electrode failure (<1%) (25).

FIGURE 72-3. The NESS H200 system. (Copyright 2007 Bioness, Inc., Valencia, CA.)

FIGURE 72-4. IST12 system for upper extremity function. (Courtesy of the Cleveland FES Center, Cleveland, OH.)
All subjects have demonstrated increased independence and improved function in ADL. Thus far, the IST12 system has been implemented mainly in individuals with C5 or C6 SCI; clinical studies have also begun to evaluate the IST12 in patients with higher spinal cord injuries and in stroke survivors with chronic upper extremity hemiplegia and minimal flexor muscle hypertonia.

**Future Directions**

Current research in upper extremity neuroprostheses focuses on stimulation of additional muscles, evaluation of new control methods, and incorporation of advanced technologies in order to provide additional function, broaden the clinical indications, and facilitate clinical implementation. Patients with C4 and higher levels of SCI need proximal arm control and hand control. Several case studies using percutaneous or surface stimulation of shoulder and elbow, and forearm and hand muscles, have been conducted to test the feasibility of a neuroprosthesis that restores function to the entire upper limb (49–51). Recently, a patient with high level tetraplegia resulting from a C1 Brown Seaward SCI was implemented with two IST12 implants to produce hand, elbow, and shoulder movements that enable basic daily tasks such as eating and grooming (52). Nerve cuff electrodes are used to stimulate elbow and shoulder muscles, and EMGs from neck and shoulder muscles are recorded with implanted electrodes and used as control signals. The cuff electrodes have exhibited excellent and stable performance for more than 1 year. Extensive functional assessment of this system is underway.

As researchers strive to develop NMES systems that are applicable to individuals with higher SCIs, the need for strategies that allow the user to naturally control the entire upper arm and hand becomes even more important. In addition to further exploring the use of multiple EMG signals and signal processing methods for control (53–56), alternative control inputs are being studied, including head orientation and movements (57), eye movements, and signals recorded from the brain (58). In the future, cortical signals recorded from electrodes implanted in the brain will be incorporated into neuroprosthetic systems as control sources, which may allow users to control stimulated limbs more naturally (59,60).

Advances are also being made to eliminate the external control unit and RF coil by implanting the control technology and powering the stimulator with a rechargeable battery, similar to cardiac pacemakers. In development at Case Western Reserve University is a networked neuroprosthesis, a totally implantable modular FES system that can be used for all purposes, for example, upper and lower limbs, trunk support, bladder, bowel, and diaphragm function. The networked neuroprosthesis will be upgradeable to provide additional and advanced functions by adding different components, for example, stimulator and sensor modules. The advantage of the system is that multiple applications can be implemented in a single patient without the need for multiple devices that are specially designed for single applications and are difficult to integrate or upgrade (9).

**Lower Extremity Function**

Historically, there have been three objectives in the application of NMES neuroprosthetic systems to the lower extremity: (a) prevent footdrop in hemiplegia (21,61–63), (b) enable standing and transfer in paraplegia (64–67), (c) enable walking in paraplegia (68,69). Today, these three objectives continue to be the focus of several commercially available systems and FES research programs.

**Footdrop Systems**

In many stroke survivors diminished ankle dorsiflexion, knee flexion, or hip flexion can result in the inability to clear the floor with the affected limb during the swing phase of gait. The purpose of footdrop systems is to produce ankle dorsiflexion and eversion during the swing phase of gait by stimulating the peroneal nerve at the appropriate time in the gait cycle. Clinical indications include footdrop that requires the patient to use compensatory strategies such as circumduction, hip hike or vaulting to clear the toes on the hemiparetic side. The patient should have sufficient balance, endurance, and motor ability to amble at least 30 ft with minimal assistance or better. Peroneal nerve stimulation (PNS) should dorsiflex the ankle to neutral, with balanced eversion/inversion while the patient is standing. PNS provides only limited knee control. Thus, patients who require an ankle-foot orthosis (AFO) to prevent knee flexion collapse or for severe genu recurvatum are not appropriate for PNS. Footdrop systems may also be appropriate for some individuals with incomplete SCI, traumatic brain injury, and multiple sclerosis.

There are three FDA-cleared (510(k)) commercially available surface NMES systems for preventing footdrop: the Odstock dropped-foot stimulator (ODFS by NDI Medical, Cleveland, Ohio), WalkAide (Innovative Neurotronics Inc., Austin, TX), and NESS L300 (Bioness Inc., Valencia, CA). All three systems (Fig. 72-5) use surface electrodes; the cathode is placed over the common peroneal nerve just below the head of the fibula and the return electrode is placed over the tibialis anterior. What distinguishes these systems from each other is the method by which the stimulation is made to turn on at the right time in the gait cycle and the way the components of the system are packaged. The ODFS (70), developed in Salisbury, England, uses a heel switch in the shoe of the paretic side to trigger stimulation (Fig. 72-5A). When the switch is depressed (stance phase), stimulation to dorsiflexors is off; stimulation turns on when the foot is lifted. The heel switch and electrodes connect to the stimulator with small cables. The stimulator is clipped to the waistband or carried in a pocket. WalkAide, developed at the University of Alberta, is a self-contained system worn as a cuff below the knee (Fig. 72-5B), with integrated electrodes and a built-in tilt sensor that detects step initiation and controls the timing of stimulation (71,72). The NESS L300 stimulator and electrodes (Fig. 72-5C) are also packaged as a cuff worn below the knee, but uses a wireless foot-pressure sensor to trigger stimulation appropriately to the gait cycle (73). The results of a meta-analysis of eight studies of surface PNS for hemiplegic gait showed a pooled improvement in walking speed of 38%
with PNS compared to no device (74). Single-channel surface
PNS appears to be effective in enhancing the gait relative to no
device and may be equivalent to an AFO (75).

Two implanted footdrop systems are commercially
available in Europe, but not yet in the United States: ActiGait
(Neurodan A/S, Aalborg, Denmark) and STIMuSTEP
(Finetech Medical Ltd, Hertfordshire, UK). The potential
advantages of implanted systems are that they do not require
daily placement of electrodes, stimulation may be less pain-
ful than with surface electrodes, and convenience, cosmesis,
and repeatability may be improved. Also, the implanted sys-
tems have more than one channel of stimulation, which may
allow a more balanced ankle dorsiflexion. The ActiGait system,
developed at the University of Aalborg in Denmark, uses a
4-channel cuff electrode surgically placed around the common
peroneal nerve (76). The stimulator is implanted on the lateral
aspect of the upper thigh. A wireless external heel switch sends
a signal to an external controller worn at the waist, which then
communicates with the implant to deliver appropriate stimula-
tion. The STIMuSTEP system, developed at the University of

Twente in the Netherlands, consists of a 2-channel stimulator
implanted below the knee, two bipolar intraneural electrodes
implanted under the epineurium of the two branches of the
common peroneal nerve to allow control of inversion and evers-
ion, an external transmitter with a built-in antenna worn over
the implant below the knee, and an external cabled heel switch
(77). Significant improvements in walking speed have been
reported with both the ActiGait and STIMuSTEP systems.

**Standing and Transfer Systems**
The primary objective of some lower extremity NMES systems
is to enable persons with paraplegia due to thoracic or low cervi-
cal level SCI to stand from a seated position and transfer to
another surface independently. The functional goals associated
with standing include reaching for high objects, having face-
to-face interactions with other people, performing tasks that
require standing, and transferring between surfaces (e.g., bed
to wheelchair and back) independently or with only minimal
assistance. The earliest demonstration of standing with NMES
was achieved by stimulating the quadriceps and glutei muscles

**FIGURE 72-5.** Footdrop systems available in the United States.
A: ODFS. (Courtesy of NDI Medical, Cleveland, OH) B: WalkAide (cour-
tesy of Innovative Neurotronics, Austin, TX). C: NESS L300. (Copyright
2007 Bioness Inc., Valencia, CA.)
for knee and hip extension with surface electrodes (64). Later studies showed that standing could be achieved in some subjects with stimulation of the quadriceps alone (65,66), but a more recent study suggests that adding stimulation of the hip extensors could speed up the rising phase, thereby reducing the overall effort (78). A balance aid such as a walker is still required for all NMES standing systems; voluntary control of upper extremities to maintain stability and balance is required. All NMES systems that enable walking in paraplegia also enable standing (see Walking Systems). There is one NMES system that is currently available as part of a clinical trial and that has standing and transfer as its primary aim (79).

The implanted neuroprosthesis for standing and transfer (Fig. 72-6) (79), developed at the Cleveland VA Medical Center and Case Western Reserve University, uses an 8-channel implantable stimulator (48). Epimysial electrodes are implanted bilaterally on the vastus lateralis, gluteus maximus, and semimembranosus (or alternatively, the posterior adductor magnus) for knee and hip extension, and intramuscular electrodes are implanted in the lumbar erector spinae for trunk support. The stimulator is implanted in the anterior lower abdomen (80). The user triggers stimulation with the press of a button. The button can be worn as a ring around the index finger (command ring) and activated with the thumb, or can be mounted on a walker, which is used to assist with balance. The command ring is connected with a cable to an external control unit, which is worn around the waist. The external control unit communicates with the stimulator through an RF-transmitting coil taped over the implant. When the button on the command ring is pressed, a short delay allows the user to comfortably position their hands on an assistive device before stimulation to the trunk, hip, and knee extensors increases to a level sufficient to raise the body from the seated position and maintain continuous standing. Another press of the button reverses the process and lowers the user to a seated position in a controlled fashion.

In a pilot study of 12 subjects with C6-T9 level SCI, the strength of knee and hip extension produced with the CWRU/VA standing system was adequate to achieve standing in all subjects (79). The subjects generally maintained more than 85% of their body weight on their legs and could stand for periods ranging from 3 to 40 minutes. While standing, they were frequently able to release one hand from the balance aid and manipulate objects in the environment. Some were even able to use the system for limited swing-through gait with a walker. The neuroprosthesis also reduced the effort and assistance required to transfer, especially from a lower surface to a higher one. All of the subjects expressed satisfaction with the device and indicated that they would go through the surgery and rehabilitation again to achieve the same results (80,81).

Preliminary results indicate that stimulation thresholds are stable and internal components are reliable with survival rates of epimysial electrodes in the extremities approaching 95% (26). To date, a total of 17 subjects have received the CWRU/VA standing system. Usage data for the last 30 consecutive days surveyed for each individual currently in long-term follow-up (12 to 36 months) show that some form of use was logged on 40% of the days surveyed (82).
Walking Systems

The ultimate objective of lower extremity FES systems is to enable individuals with complete paraplegia to walk again. Some walking systems may also be applicable to individuals with incomplete paraplegia or hemiplegia to assist gait. Pioneering work by Kralj, Bajd, and others in Ljubljana, Slovenia, introduced the technique of eliciting a flexion withdrawal reflex of the hip, knee, and ankle by stimulating the peroneal nerve (68,69). This action, elicited by a single electrode, produces lower extremity motion that can be substituted for the swing phase of gait. Stimulation of the quadriceps for knee extension during the stance phase of gait completes the gait cycle. Hence, it is conceivable to provide ambulation with just four stimulus channels and bracing at the ankles. In patients who cannot voluntarily extend the hips, the gluteal muscles are stimulated as well, for a total of six stimulus channels.

This stimulation technique gave rise to the only FDA-approved (PMA) NMES system for walking. The Parastep system (Sigmedics, Inc., Fairborn, Ohio) uses four or six channels of bilateral surface stimulation of the quadriceps, peroneal nerves (for reflex flexion withdrawal), and, if necessary, the glutei to enable individuals with T4–12 paraplegia to walk with a walker (83). A microprocessor/stimulator unit is worn at the waist, and using a walker with controls built into the handles, individuals with paraplegia can stand and walk with reciprocal gait for limited distances. Parastep has been fitted to more than 1,000 users, with nearly all of them achieving standing and at least 30 ft of ambulation (84,85). Use of the system has additional medical benefits, such as increased blood flow to the lower extremities (86), lowered heart rate at subpeak work intensities (87), increased muscle mass (88), and psychological benefits (89). However, a more recent assessment of the Parastep (90) concluded that in spite of its ease of operation and good cosmetic acceptance, the Parastep approach had limited applications for mobility in daily life because of its modest performance associated with high metabolic cost and cardiovascular strain. Other complications include poor standing posture in some due to hip flexion generated by the rectus femoris when the quadriceps are stimulated, lack of or habituation of an adequate flexion withdrawal reflex, and difficulty controlling the swing limb motion due to the mass flexion response of the reflex. In 2003, Parastep was approved for reimbursement (cost of equipment and training) through the Center for Medicare and Medicaid Services (CMS).

Percutaneous (91–93) and implanted (94,95) NMES walking systems have been tested in clinical trials. Rather than eliciting a reflex to produce stepping, these systems attempt to stimulate multiple individual muscles with appropriate intensities and timing to create walking movements in response to the user pressing switches to initiate stepping. Presently, researchers in Cleveland are extending the functionality of their standing system (see Standing and Transfer Systems) with a 16-channel system. The 16 channels are provided by either two 8-channel stimulators implanted bilaterally (95,96), or more recently, by a single 16-channel stimulator. With this approach, four individuals with paraplegia have been given the ability to stand and walk (using a walker) with standby assistance; three of these have complete injuries. This research group has also recently reported on the use of a single 8-channel implanted system for walking in a patient with incomplete (ASIA C) SCI (97). After 12 weeks of training, with the neuroprosthetic system the participant's functional ambulation category improved from non-functional to limited community ambulation. Practice with the system over 12 weeks significantly decreased the physiological effort required to walk and increased the walking distance and speed. Since that report, a second patient with incomplete SCI has received a 12-channel system (IST12 technology) with EMG control for walking, and more studies are planned.

One major drawback of walking with NMES, regardless of the system design, is the high metabolic cost primarily due to the great mechanical work needed to support the body weight and achieve forward progression. Hybrid systems reduce the energy required for walking by combining NMES systems with mechanical orthoses (98). The bracing supports the user's body weight, and the stimulation provides propulsion. The most widely tested hybrid system is the reciprocating gait orthosis (RGO) (99), which uses surface stimulation of the quadriceps (rectus femoris) and hamstrings to help power hip flexion and extension. Although hybrid RGOs are often successful in allowing users to stand and walk with less energy consumption, they are difficult to put on and are often cosmetically unacceptable, and therefore have low long-term usage rates. Other more recent hybrid system developments include an energy storage orthosis using a pneumatic component to store energy during one phase of the gait cycle and transfer it during the other phase (100).

Future Directions

Advances toward fully implanted footdrop systems are being made. Neurostep (being developed by Neurostream Technologies Inc., Saint-Augustin-de-Desmaures, Quebec, Canada) uses a nerve cuff electrode on the tibial nerve to record afferent signals from sensory receptors in the foot (101). A gait event detection algorithm detects heel contact and toe lift, and triggers an implanted stimulator to deliver pulses to the ankle dorsiflexors through a second 4-channel nerve cuff placed on the common peroneal nerve. An initial feasibility case study reported by Hoffer et al. demonstrated gait event detection success rates ranging from 72% to 90% (8). Clinical trials have begun in Canada and are planned for the United States. Also, the injectable microstimulator (BION) approach is under investigation for the correction of footdrop (102).

For standing and walking systems, researchers are investigating the use of implanted nerve cuff electrodes to provide more complete muscle recruitment, and thereby improve standing posture and extend standing duration (27,103). Future standing and walking systems will aim to reduce or eliminate the need for assistive devices for balance (e.g., walker, crutches, or additional bracing) by incorporating automatic control of balance (104,105). By automatically adjusting for externally applied perturbations, the user's hands will be free to perform reaching tasks or other functional activities. In addition,
automatic control of the stimulation sequence during walking would produce a more efficient and natural gait (106). New implant technology with more stimulation channels could be used to: a) activate synergistic or complementary muscles, b) activate trunk muscles for greater stability when standing or sitting, or c) achieve more complete activation of the targeted muscles. This would improve system performance, simplify surgical installation, and make implanted standing and walking neuroprostheses applicable to a wider user population.

Respiratory Function
Following cervical SCI, a significant number of individuals suffer from chronic respiratory insufficiency because the central neural connections to the diaphragm are disrupted. The diaphragm is innervated by the phrenic nerve, which consists of nerve fibers that derive primarily from C3, C4, C5, and even C6 lower motor neuron and nerve roots, but most of its nerve supply comes from the C4 neurological segment (107). Therefore, an injury to the spinal cord above C4 level will result in respiratory failure requiring mechanical ventilatory assistance, often permanently. A mechanical respirator sustains life by continuously forcing air into and out of the lungs through a tracheostomy at the base of the neck. Unfortunately, mechanical respiration is associated with substantial morbidity, mortality, inconvenience, physical discomfort, fear of disconnection, difficulty with speech, impaired sense of smell, and encumbered mobility (108). A clinical condition that spares the lower motor neurons that innervate the diaphragm (such as a complete transverse lesion at C1 and C2 levels) may allow electrical activation of the diaphragm by stimulating the phrenic nerves. The purpose of a respiratory neuroprosthesis is to restore breathing by rhythmically activating the diaphragm by this technique, and thereby eliminate the need for a mechanical respirator.

Electrical stimulation can be used to activate the diaphragm through the phrenic nerve. This technique, known as phrenic nerve pacing, was introduced in the 1960s by Glenn and colleagues at Yale University (109). Bilateral synchronous stimulation of the phrenic nerves causes contraction and descent of each hemi-diaphragm, which subsequently causes a fall in intrathoracic pressure and inspiration. Cessation of stimulation results in relaxation of the diaphragm, an increase in intrathoracic pressure, and exhalation. The stimulation cycle is repeated 8 to 14 times per minute to produce a normal breathing pattern. The user may adjust the number of breaths per minute and the duration of each breath.

Candidates for phrenic nerve pacing must have proven viability of the phrenic nerves, must be free of significant lung disease or primary muscle disease, have good sitting tolerance, strong desire and motivation, good family support, and a clear understanding of the potential benefits to be achieved. Typically, a significant surgical procedure is required, which brings associated potential complications. Therefore, patients must be carefully screened. Phrenic nerve function is assessed with electrical stimulation of the cervical portion of the phrenic nerve and simultaneous monitoring of the diaphragmatic responses by surface electromyography to measure conduction velocity. In addition, a fluoroscopic examination of the chest should be done to verify that there are adequate diaphragmatic excursions in response to stimulation (110).

Currently there are at least two conventional phrenic pacing systems commercially available, the Avery Mark IV Breathing Pacemaker system (Avery Biomedical Devices, Commack, NY), and the Atrostim Phrenic Nerve Stimulator (Atrotech Ltd, Tampere, Finland). Both these systems have phrenic nerve stimulating electrodes implanted bilaterally in the cervical or thoracic regions and RF receiver/stimulators implanted under the skin on the anterior chest wall with implanted leads connecting the electrodes to the stimulators (Fig. 72-7). The stimulators receive power and commands from an external control unit, transmitted via an antenna taped directly over each implanted stimulator. The differences between the systems are mainly in the type of electrodes and stimulation strategies used. The Avery system, developed by Glenn et al. (109,111), uses either monopolar or bipolar electrodes that the surgeon positions around the nerve and sutures in place. The Atrostim system uses quadripolar electrodes having four contacts spaced evenly around the phrenic nerve (112). Each of the four contacts in turn serves as the cathode, and a contact on the opposite side of the nerve serves as the anode. This multipole sequential stimulation technique is intended to reduce fatigue by activating motor units only one fourth of the activation time that occurs with conventional monopolar stimulation. The Avery system is FDA-approved (PMA), has approval for reimbursement through CMS, and is the most widely used phrenic pacing system worldwide.

FIGURE 72-7. Avery Breathing Pacemaker system. (Courtesy of Avery Biomedical Devices, Commack, NY)
The Atrostim system is commercially available in Europe and has been approved for investigational use by the FDA under an Investigational Device Exemption (IDE) in the United States.

Surgical implantation of phrenic pacing systems may be done by a cervical or thoracic approach (111,113). The thoracic approach requires a thoracotomy to expose the phrenic nerve, which has significant associated risks including hemothorax and pneumothorax, and which requires chest tube placement and intensive postoperative care. Alternatively, the stimulators and electrodes may be implanted without a thoracotomy through an incision in the neck, but with this approach submaximal diaphragm activation may result because additional accessory branches join the phrenic nerve in the thorax in most individuals. The thoracic approach, therefore, has been the preferred method of electrode placement (114). More recently, an endoscopic technique has been developed to allow the electrodes to be implanted without a thoracotomy (115). Although there are a number of acceptable surgical approaches for thoracic electrode placement, the phrenic nerves are usually accessed by incisions through the second intercostal spaces and the electrodes are secured behind or over the nerve. It is critical that the phrenic nerves are manipulated with extreme care to avoid mechanical trauma to the nerve and its blood supply. The stimulator is positioned in a subcutaneous pocket on the anterior chest wall; leads from the electrode are passed through the third or fourth intercostal space and connected to the stimulator.

Postoperatively, the diaphragm is gradually reconditioned to reverse disuse atrophy and to regain strength and endurance (116). Conditioning is initiated approximately 2 weeks after surgery. Phrenic nerve pacing is initially done daily for 10 to 15 minutes every hour and is gradually increased over a conditioning phase that may take 10 to 12 weeks or longer, although it is possible to achieve full-time pacing in some patients within 5 weeks (108). Low-frequency stimulation is used during the conditioning phase to promote conversion of fast-twitch fibers to slow-twitch fatigue-resistant fibers. After pacing is achieved throughout the waking hours, pacing is done during sleep and gradually increased until the pacer is used full time. Generally, the tracheostomy is kept open, but it may be cuffed especially during waking hours. By leaving it open, suctioning and removal of secretions is made easier and mechanical ventilation may be restarted when medically necessary. During sleep, the tracheostomy should be open in order to reduce the risk of upper airway collapse and obstructive apnea.

Phrenic nerve pacing has allowed patients to decrease or even discontinue the use of mechanical respirators and have enabled more normal breathing. Although there may be significant patient variability, most patients describe an improved level of comfort, reduced anxiety and embarrassment, increased mobility, improved speech, and greater sense of well being and an overall health as most important benefits as compared to mechanical ventilation (117–119). While a number of complications have been reported since phrenic nerve pacing was first introduced, technical developments and patient experience have markedly reduced their incidence (118,120). Early failures were mainly due to technical malfunction of the device components or due to insufficient phrenic nerve innervation (a patient selection error). Additional complications encountered include diaphragm fatigue, increases in airway resistance caused by the accumulation of airway secretions (requires suctioning), infection, injury to the phrenic nerve, and upper airway obstruction after tracheostomy closure (108).

There are only a few studies that have evaluated recent success and complication rates. In one long-term follow-up study of 12 quadriplegic patients, 50% continued to use the Avert system full time (mean 13.7 years), one used it part time, three stopped using it, and two were deceased (one had used the system full time, and one had stopped using it) (121). Those who stopped using the pacer did so because of inadequate social or financial support or medical problems associated with the initial injury. All patients demonstrated normal tidal volumes while pacing full time, no patient lost the ability to activate the phrenic nerve, and threshold and maximal currents did not increase over time. In an international study of 64 patients using the Atrostim system, 94% of the 35 pediatric patients and 86% of the 29 adult patients eventually achieved complication-free successful pacing (122). Thirty-four percent of all subjects paced full time, 38% paced only while awake, 14% paced only while asleep, and 2% stopped pacing. The incidences of electrode and stimulator failure were 3.1% and 5.9%, respectively, and the incidences of infection and nerve trauma were 2.9% and 3.8%, respectively.

Alternative electrode placement procedures have been developed to reduce surgical risks and associated costs. One such alternative uses a minimally invasive laparoscopic procedure to position intramuscular electrodes bilaterally near the motor points of each hemi-diaphragm (123,124). With this procedure, four laparoscopic ports provide access to the abdominal cavity for visualization, insufflation of the abdominal cavity, diaphragm mapping, and insertion of the implant tool and four electrodes (124,125). The electrode leads are tunneled subcutaneously to the chest wall where they exit and connect to an external stimulator. This percutaneous diaphragm pacing system is called NeuRx DPS system (Synapse Biomedical, Inc., Oberlin, Ohio); it has recently received FDA approval (PMA) and also has the CE Mark for marketing in Europe. The system has been implemented in at least 55 patients with SCI, with over 90% achieving significant independence from mechanical ventilation and over 50% eliminating mechanical ventilation. Initial experience with the first five patients has been reported (123). The system has also been implemented in at least 70 patients with ALS under an IDE, with preliminary results indicating an improvement in the rate of respiratory decline, improvements in sleep disordered breathing, apparent trophic effects as demonstrated by increasing EMG activity of the diaphragm with conditioning, improvements in respiratory system compliance, and improved survival (mean survival from diagnosis appears to be 61 months for DPS patients).
Future Directions

Unfortunately, many individuals with ventilator-dependent tetraplegia are not eligible for phrenic nerve pacing because of complete or partial injury of one of the phrenic nerves. For patients with only a single functional phrenic nerve, it may be possible to achieve respiration by activating the inspiratory intercostal muscles in addition to the functioning phrenic nerve (126,127). In a study of four patients with SCI between C2 and C4 levels, upper thoracic epidural spinal cord stimulation of the intercostal muscles, in combination with phrenic nerve stimulation, produced inspiratory volumes equal to that typically achieved with bilateral phrenic nerve pacing (127,128). However, intercostal pacing also results in contraction of nonventilatory muscles and therefore is less efficient than diaphragm pacing. Nevertheless, the patients in this study were able to achieve substantial time free of mechanical ventilatory support, and all four reported a level of comfort very near normal breathing. This approach continues to be developed. For patients with significant injury to both phrenic nerves, the development of an intercostal to phrenic nerve transfer procedure (129) may restore phrenic nerve viability and allow these patients the possibility of diaphragm pacing.

Additional avenues of research and development include the use of computer-aided robotic technology to implant electrodes around the phrenic nerves through three 1-cm incisions in the intercostal spaces on each side (130), thereby avoiding the conventional and invasive thoracotomy procedure. Others are working on developing a totally implanted system similar to cardiac pacemakers. Researchers in Japan have used a Medtronic spinal cord stimulator (Itrel 3), an implant with its own battery and requiring no RF communication or external components (except for programming), as a diaphragm pacemaker in six patients with encouraging results (131,132). Systems are also being developed to produce cough and thereby reduce the risk of developing respiratory tract infections as the result of expiratory muscle paralysis (133–135). Recently, a new method has been tested to activate the expiratory muscles via lower thoracic and upper lumbar spinal cord stimulation (136).

Bladder and Bowel Function

SCI above the sacral levels of the spinal cord results in loss of bladder and bowel control. Numerous complications related to this loss of function include frequent urinary tract infections, renal deterioration, bladder or kidney stones, and autonomic dysreflexia. These complications threaten to compromise the health and quality of life of the individual, and can lead to major long-term, and even deadly, consequences. The functional goal of a bladder neuroprosthesis is to produce effective micturition (bladder voiding) and continence, and also improve bowel function.

Micturition and continence are normally performed by the coordinated action of the bladder (detrusor muscle) and the urethral sphincter, two muscles that are controlled by neural circuits in the brain and lumbosacral spinal cord. Micturition normally occurs when the bladder contracts while the urethral sphincter relaxes; continence normally occurs when the bladder relaxes while the urethral sphincter contracts. Suprasacral SCI disrupts the coordination and control of the two muscles, resulting in detrusor-sphincter dysynergia and neurogenic detrusor overactivity (also called detrusor hypertrophy). Detrusor-sphincter dysynergia is a condition when the bladder and urethral sphincter contract simultaneously rather than reciprocally, thereby impairing micturition. Neurogenic detrusor overactivity is characterized by involuntary detrusor contractions during the filling phase, thereby impairing continence. These effects result in elevated bladder pressure, which is the primary risk factor for renal deterioration. Thus, a reduction of bladder pressure is also an important objective of a neuroprosthesis for bladder function.

The use of electrical stimulation for restoring bladder control has been most advanced in the United Kingdom by the work of Giles Brindley (137). The Finetech Brindley Bladder Control system (Finetech Medical, Ltd, Hertfordshire, UK) is an implanted FES system that enables micturition by producing a bladder contraction with stimulation of the sacral anterior spinal nerve roots. The system has been implanted in more than 2,500 patients in at least 20 countries (138,139). The system is FDA-approved (PMA) in the United States and has the CE mark in Europe. The majority of the first 500 patients had SCI; however, patients with other neurological disorders, including multiple sclerosis, spinal cord tumors, transverse myelitis, cerebral palsy, and meningomyelocele have also received the implant (140). Candidates must have intact parasympathetic neurons to the detrusor muscle. The function of these neurons is demonstrated by reflex detrusor contractions on a cystometrogram. Frequent urinary tract infection and problems tolerating catheters or anticholinergic medications are further indications. Patients can receive an implant at any time after reaching neurological stability. They should also have an appropriate degree of emotional and social stability.

The system consists of two pairs of tripolar electrodes (active contact flanked on each side with a return contact), an implanted receiver/stimulator and an external control unit that powers and controls the stimulator through an RF transmission antenna (Fig. 72-8). The electrodes are implanted bilaterally on the sacral spinal nerve roots either intradurally on the anterior (motor) roots in the cauda equine via a lower lumbar laminectomy (141), or extradurally on the mixed sacral nerves in the sacral canal via a laminectomy of S1-3 (142). The advantage of the intradural approach is that it attempts to only activate efferent fibers and therefore minimize activation of reflexes through the posterior sacral roots. However, the extradural approach is reported to carry less risk of trauma to the nerves or of cerebrospinal fluid leakage. Intraoperative electrical stimulation and recording of bladder pressure is used to confirm the identity of the nerves supplying the bladder. The electrode leads are tunneled to the receiver/stimulator, which is placed under the skin on the anterior abdominal wall through a separate incision. Bilateral posterior rhizotomies of the S2-4 spinal nerves are usually performed to eliminate involuntary detrusor contractions (associated with neurogenic detrusor
stimulation (3 to 6 seconds on, 6 to 9 seconds off) leads to sustained contraction of the bladder and relaxation of the sphincter between stimulus periods. Thus, urine flows during the intervals between stimulus periods and the bladder is emptied in spurts, an effect called poststimulus voiding.

A multicenter study has reported that more than 85% of 184 implant recipients use the system as the primary means of bladder emptying (139). Residual volume in the bladder following system use was less than 60 mL in 95% of the users and less than 30 mL in 89%. A substantial decrease in symptomatic urinary tract infections following use of the implant has been reported in several studies (139,143–145). Continence is reported in more than 85% of patients (139,140,145,146), largely because of increased bladder compliance following the posterior rhizotomy (147,148). About 10% to 15% of patients report some stress incontinence of urine following implantation of the stimulator and posterior rhizotomy (147). Several centers in Europe have long-term follow-up experience with the device, particularly with regard to the upper tracts (139,146,149,150). Trabeculation, ureteric reflux, and hydronephrosis tend to decrease in patients who undergo implantation and posterior rhizotomy. There is also a reduction in the incidence of autonomic dysreflexia, an outcome that is particularly beneficial to males with tetraplegia who were formerly dependent on an indwelling catheter, which is prone to blockage and frequent infection. The ability to micturate on demand and improved continence of urine contribute to a reduction in the use of intermittent and indwelling catheterization. Most users become free of catheters and urine collection bags and can discontinue anticholinergic medication, which in turn reduces constipation and other side effects such as dry mouth and drowsiness (143,151). Satisfaction with the system has been reported to be high (143), and the hardware has been reliable (152). Infection associated with the implant is rare, occurring in 1% of the first 500 implants and technical faults in the implanted equipment is uncommon, occurring on an average once every 20 implant-years (152). Studies in Europe and the USA indicate that the use of the implanted stimulator together with posterior sacral rhizotomy results in substantial savings in the cost of bladder and bowel care, particularly from reduction in supplies needed for bladder care, medications, and visits to physicians for the management of complications (153–155).

Bowel evacuation and penile erection are secondary uses of the Finetech Brindley Bladder system. Regular stimulation of the sacral nerve roots not only results in contraction of the detrusor, but also of the external urethral sphincter, an effect that does not seem conducive to micturition. This coactivation occurs because the sacral nerve roots contain both the small-diameter preganglionic parasym pathetic axons innervating the bladder via the pelvic nerve and the large-diameter somatic motor axons innervating the external urethral sphincter through the pudendal nerve. Because the large fibers have a lower threshold for excitation than smaller fibers, it is difficult to produce contraction of the detrusor without contraction of the external sphincter. Nevertheless, micturition can be produced by intermittent stimulation, a technique that takes advantage of the fact that the relaxation time of the smooth detrusor muscle is longer than that of the striated external urethral sphincter muscle. Intermittent overactivity) and thereby provide continence. Postoperatively, urodynamic studies are used to guide the setting of stimulus parameters to give an acceptable voiding pressure and rate and pattern of flow. The stimulus program is checked between 1 and 3 months after surgery since the response of the bladder may change with repeated use; thereafter, review is recommended at least annually, monitoring lower and upper urinary tract function. The external control unit has programs for bladder or bowel control, which the user selects and then turns the device on to deliver the appropriate pattern of stimulation to the sacral nerves.

Stimulation of the sacral nerve roots not only results in contraction of the detrusor, but also of the external urethral sphincter, an effect that does not seem conducive to micturition. This coactivation occurs because the sacral nerve roots contain both the small-diameter preganglionic parasym pathetic axons innervating the bladder via the pelvic nerve and the large-diameter somatic motor axons innervating the external urethral sphincter through the pudendal nerve. Because the large fibers have a lower threshold for excitation than smaller fibers, it is difficult to produce contraction of the detrusor without contraction of the external sphincter. Nevertheless, micturition can be produced by intermittent stimulation, a technique that takes advantage of the fact that the relaxation time of the smooth detrusor muscle is longer than that of the striated external urethral sphincter muscle. Intermittent stimulation (3 to 6 seconds on, 6 to 9 seconds off) leads to sustained contraction of the bladder and relaxation of the sphincter between stimulus periods. Thus, urine flows during the intervals between stimulus periods and the bladder is emptied in spurts, an effect called poststimulus voiding.

A multicenter study has reported that more than 85% of 184 implant recipients use the system as the primary means of bladder emptying (139). Residual volume in the bladder following system use was less than 60 mL in 95% of the users and less than 30 mL in 89%. A substantial decrease in symptomatic urinary tract infections following use of the implant has been reported in several studies (139,143–145). Continence is reported in more than 85% of patients (139,140,145,146), largely because of increased bladder compliance following the posterior rhizotomy (147,148). About 10% to 15% of patients report some stress incontinence of urine following implantation of the stimulator and posterior rhizotomy (147). Several centers in Europe have long-term follow-up experience with the device, particularly with regard to the upper tracts (139,146,149,150). Trabeculation, ureteric reflux, and hydronephrosis tend to decrease in patients who undergo implantation and posterior rhizotomy. There is also a reduction in the incidence of autonomic dysreflexia, an outcome that is particularly beneficial to males with tetraplegia who were formerly dependent on an indwelling catheter, which is prone to blockage and frequent infection. The ability to micturate on demand and improved continence of urine contribute to a reduction in the use of intermittent and indwelling catheterization. Most users become free of catheters and urine collection bags and can discontinue anticholinergic medication, which in turn reduces constipation and other side effects such as dry mouth and drowsiness (143,151). Satisfaction with the system has been reported to be high (143), and the hardware has been reliable (152). Infection associated with the implant is rare, occurring in 1% of the first 500 implants and technical faults in the implanted equipment is uncommon, occurring on an average once every 20 implant-years (152). Studies in Europe and the USA indicate that the use of the implanted stimulator together with posterior sacral rhizotomy results in substantial savings in the cost of bladder and bowel care, particularly from reduction in supplies needed for bladder care, medications, and visits to physicians for the management of complications (153–155).

Bowel evacuation and penile erection are secondary uses of the Finetech Brindley Bladder system. Regular stimulation of the sacral parasympathetic nerves contributes to the transport of stool through the distal colon into the rectum (156). Most users report a reduction in constipation and reduced need for laxatives and stool softeners (156,157). Some users are able to defecate by a pattern of intermittent stimulation similar to that used for micturition but with longer intervals between bursts of stimulation to allow passage of stool (157). This results in a significant reduction in time that is spent in bowel evacuation (143,157). Penile erection has been reported in 60% to 87% of men with intradural electrodes implanted on the anterior sacral roots (140,145,146). In a smaller group

FIGURE 72-8. FineTech Brindley Bladder Control system. (Courtesy of FineTech Medical, Ltd, Hertfordshire, UK.)
of participants with extradural sacral root stimulation, penile erection by stimulation was achieved in only approximately 10% (151).

Future Directions
Methods to induce bladder contractions without activating the external urethral sphincter, and thereby achieve a more physiological voiding pattern, are being investigated (158). For example, recent animal studies show that selective stimulation of the small fibers innervating the bladder may be possible by blocking the action potentials in the large pudendal nerve fibers, innervating the sphincter with biphasic quasitrapezoidal stimulation pulses (159,160) or high-frequency sinusoidal stimulation (161,162). An alternative method for achieving selective stimulation of the detrusor without the external urethral sphincter is to induce a reflex bladder contraction without a reflex sphincter contraction by intermittent stimulation of the pudendal nerve trunk (163).

Researchers are also exploring alternatives to posterior sacral rhizotomy for reducing neurogenic detrusor overactivity. While the benefits of posterior rhizotomy are clear—abolition of reflex incontinence, increased bladder capacity and compliance, prevention of autonomic dystreflexia triggered from the bladder or bowel, and protection of the kidneys from ureteric reflux and hydronephrosis (143)—posterior rhizotomy also abolishes other potential useful sacral reflexes, such as reflex erection, ejaculation, defecation, and sacral sensation if present. One alternative to posterior rhizotomy is to induce reflex inhibition of the detrusor by electrically stimulating large afferent neurons in the sacral dermatomes, a technique known as neuromodulation, which shows some benefit in able-bodied subjects with urge incontinence (164,165). Similar inhibition has been observed following SCI using the Finetech Brindley system without posterior rhizotomy but with stimulation of the posterior sacral afferents to suppress bladder overactivity (166). With this system, neuromodulation increased bladder capacity, but hyperreflexia of the external urethral sphincter persisted and prevented complete emptying in some cases. Another approach is to mimic the function of a posterior rhizotomy only when needed by blocking the action potentials in the posterior sacral afferents with a high frequency electrical stimulation nervous block that is fast-acting and reversible (167,168). Although this approach is still at the level of animal experimentation, its clinical implications reach far beyond the bladder applicantion and may include management of the broader problems of spasticity and pain.

CONCLUSIONS
The principal goal of rehabilitation management of persons with upper motor neuron paralysis is to maximize the quality of life. While quality of life is clearly influenced by a wide range of variables including social, emotional, psychological, vocational, and educational factors, the persistent neurological impairment after injury to the central motor system remains a powerful reminder and determinant of one’s ability to function in society. Functional NMES systems bypass the injured central circuitry to activate neural tissue and contract muscles to provide function to what is otherwise a nonfunctioning limb or organ. Recent advances in clinical medicine and biomedical engineering have made the clinical implementation of functional NMES systems to enhance the mobility and function of the paralyzed person more feasible. Hand neuroprosthesis systems can significantly enhance the upper extremity ADL of persons with tetraplegia. The application of this technology for persons with hemiplegia is in its infancy and must await further technical and scientific development if it is to be applicable to the broader stroke population. Recently, several footdrop systems for hemiplegia have reached the commercialization stage. Several lower extremity systems with and without bracing are being investigated for the purpose of functional transfers and standing, and for ambulation for patients with paraplegia. Phrenic pacing systems can provide artificial ventilatory support for patients with ventilator-dependent SCI. Finally, the bladder neuroprosthesis can provide catheter-free micturition for persons with either paraplegia or tetraplegia.

For more than 40 years, electrical stimulation has been used to restore neuromuscular function in people with paralysis. The principles of safe and reliable activation of neural tissue and the methods of generating stable and controllable muscle contractions have been well established. Electrodes, stimulators, transducers, and sensors have been developed and integrated into neuroprosthetic systems that have benefited individuals with SCI and stroke. Clinical success of several FES interventions has been demonstrated, but commercial success has proven to be more difficult to achieve, especially for upper and lower extremity implanted systems. Expanding the indications both within the SCI and stroke populations, and beyond those populations to other disability groups, will increase the number of people benefited and the potential market size. Increasing the awareness of FES technology and its benefits among rehabilitation practitioners and involving them in the development and clinical testing of neuroprostheses are important components in increasing the number of FES users and in penetrating the market. Costs can be decreased by developing systems that can be more easily manufactured and implemented. Technological advancements will increase the benefit of neuroprostheses, leading to greater function and independence for neurologically impaired individuals.

ACKNOWLEDGMENTS
The preparation of this chapter was supported in part by grants from the National Institutes of Health (NIH), National Center for Research Resources (NCRR) grant KL2RR024990, and the National Institute of Child Health and Human Development (NICHD) grant K24HD054600.
REFERENCES


Assistive Technology

HISTORICAL BACKGROUND AND PERSPECTIVE ON ASSISTIVE TECHNOLOGY

Humans have used tools to accomplish everyday tasks in many cultures throughout history (and prehistory), but the perception remains that the use of technology as a tool for persons with disabilities is a fairly recent phenomenon. In fact, James and Thorpe describe any number of assistive devices used as early as the 6th or 7th century B.C. (1). Their descriptions include partial dentures, artificial legs and hands, and drinking tubes or straws. The earliest documented account of optical and lens technologies, or eyeglasses, came from Venice around A.D. 1300 (2). The term assistive technology (AT) to describe devices used to facilitate the accomplishment of everyday tasks by persons with disabilities is actually the more recent development (3).

In 1988, Public law 100-407 defined AT as “Any item, piece of equipment or product system whether acquired commercially off the shelf, modified, or customized that is used to increase or improve functional capabilities of individuals with disabilities.” This definition also included a second component defining AT services as any service that directly assists an individual with a disability in the selection, acquisition, or use of an AT device. This includes

1. The evaluation of the needs of an individual with a disability, including a functional evaluation of the individual in his or her customary environment;
2. Purchasing, leasing, or otherwise providing for the acquisition of AT by persons with disabilities;
3. Selecting, designing, fitting, customizing, adapting, applying, retaining, repairing, or replacing AT devices;
4. Coordinating and using other therapies, interventions, or services with AT devices, such as those associated with existing education and rehabilitation plans and programs;
5. Training or technical assistance for the person with a disability or, if appropriate, his or her family;
6. Training or technical assistance for professionals (including individuals providing education or rehabilitation services), employers, or other individuals who provide services to, employ, or are otherwise substantially involved in the major life functions of children with disabilities (4).

This definition has also been included in other federal legislation authorizing services or supports for persons with disabilities, including the Rehabilitation Act (5) and the Individuals with Disabilities Act (6).

So what is AT? In short, AT is a tool used by someone with a disability to perform everyday tasks such as getting dressed, moving around, or controlling his or her environment, learning, working, or engaging in recreational activities. As a tool, AT is no different than using a hammer to drive a nail. Fewer than 30 years ago, there were fewer than 100 devices commercially available. Today, more than 29,000 assistive devices are listed on the AbleData Web site (www.abledata.com). AT use often begins shortly after birth and continues throughout the life span of individuals with disabilities.

Included within this textbook are chapters on wheelchairs and other mobility aids, recreational therapies, orthotics and prosthetics, and many other rehabilitation devices or tools. The devices described by the authors of these chapters also fall within the definition of AT. This chapter covers basic information on various categories of AT products, evaluation, prescription and funding for AT, use of a team approach, and outcomes measurement in the field of AT devices and services. Readers are referred to the appropriate chapter for more in-depth information on those devices covered elsewhere in the textbook.

AT tends to be divided into two major categories: low technology and high technology. Low-technology or “low-tech” devices tend to be simple, nonelectronic devices. Items such as dressing aids, pencil grips, picture-based communication boards for persons who are nonspeaking, and magnifiers for persons with visual impairments fall within the category of low-tech devices. High-technology or “high-tech” devices are typically described as sophisticated, electronic devices such as power wheelchairs, computers, or augmentative and alternative communication (AAC) devices that provide voice output for persons who are nonspeaking. These devices are usually fairly expensive and often require extensive training to ensure they are used to their fullest potential (7).

There are a number of myths surrounding the provision of AT that tend to reflect common misperceptions about both the technology and individuals with disabilities. These myths include

1. AT is the “be all and end all.”
2. AT is complicated and expensive.
3. Persons with the same disability benefit from the same devices.
4. Professionals are the best source of information for AT.
5. AT descriptions are always accurate and helpful.
6. A user’s AT requirements need to be assessed just once.
7. AT devices will always be used.
8. Individuals with disabilities want the latest, most expensive device.
9. AT is a luxury.
10. Only people with certain types of disabilities find AT useful.

Although AT does hold a great deal of promise for persons with disabilities who would like to pursue everyday tasks that most of us take for granted, myths such as these have a tendency to cause practitioners and those they serve to disregard the potential utility of assistive devices. Tasks such as navigating freely throughout the community, talking with a loved one, and writing a letter are often out of reach for persons with disabilities. Proper prescription of assistive technologies can enable persons with disabilities to learn, work, and play, just like everyone else. Dispelling these myths can do much to ensure that those who can benefit from assistive devices and services receive appropriate supports.

**AT AND ABANDONMENT**

It is important to recognize that not everyone with a disability enjoys using technology, however useful it might appear to practitioners. Depending on the type of technology, nonuse or abandonment can be as low as 8% or as high as 75%. On average, one third of more optional ATs are abandoned, most within the first 3 months after acquiring the device. To date, research has not been done to ascertain the number of individuals who must continue to use devices they are not pleased with simply because they cannot abandon the technology without severe consequences (8,9). For example, an individual who has just received a new wheelchair that does not meet expectations simply cannot stop using the chair to navigate independently within his or her community. Rather, he or she must wait until third-party funding becomes available again (often as long as 3 to 6 years) or engage in potentially difficult and unproductive discussions with the vendor who has more than likely provided the chair as it was prescribed by the assessment team.

Research does tell us the number one reason individuals with disabilities choose not to use assistive devices is because practitioners failed to consider their opinions and preferences during the device selection process. In other words, the person with a disability was not included as an active member of the team during the evaluation process (8).

**HUMAN FACTORS AND AT**

A growing body of research in the field of human factors is being applied to the design and development of AT devices for persons with disabilities. Analysis of human factors in a global sense is concerned with how humans interact with various technologies. When you sit in a new car and notice how comfortable it is—how well the seat contours with your body and how accessible the controls are for the stereo system—you have experienced the growing information derived from human factors research.

Dr. King (10), a professor at the University of Wisconsin-Eau Claire, has expanded key points found in the literature on human factors and applied them to research and development in AT. He tells us that human factors in AT must be concerned with how human beings who have special needs, limitations, or disabilities interact with devices and tools that may support, supplement, or replace some process or ability that has been lost or impaired by illness or injury. We must be concerned not only with how the user interacts with the devices but also with how the family or other close care providers react to the use of tools and devices in their settings. The interaction of the AT user alone with the technology does not tell us the whole story because persons who use assistive technologies must interact closely and frequently with, or depend on, others, for daily care and other aspects of their lives. The larger impact on those around the user must also be considered because they are key players in the implementation of any AT in the user’s life. Across all component areas of AT, human factors must be concerned with how the potential user—as well as his or her family, personal care providers, education and therapy aides, teachers, and clinicians—interacts with the assistive devices and technological systems.

Analysis of human factors in AT is concerned with finding out the special needs, capabilities, and limitations of users and then matching devices and controls to each individual user. Heterogeneity and individualization are primary considerations in dealing with persons who have special needs. However, mass-produced technologies (such as computers) must be designed for mass-market users, rather than for unique individuals. Flexibility and adaptability of technology to a wide range of user characteristics are critically important.

Human factors considerations in AT are especially focused on reducing the user’s exertion, stress, and fear of use. We all have a bit of “technophobia” when it comes to use of new tools and devices—especially complex, high-tech devices. The fear and stress, particularly when they relate to devices that may be difficult to set up or require a great deal of exertion to use, can be deleterious to the AT user. Persons relying on AT have some type of limitation or disability. Expecting an individual with a disability to become skilled in additional tools, devices, or technology can be highly stressful to users as well as caregivers because it adds more complexity to their life.

Human factors also focus on reducing the danger caused by the device to the user and persons around them. AT professionals should also strive to reduce the possibility of failure during use, which can lead to rejection or abandonment of the device for future use, even when the system has considerable merit for the person with a disability (10).
Disability itself is not precise and quantifiable. The concept of disability is not always agreed on by persons who self-identify as having a disability, persons who study disability, or the general public (11). This lack of agreement creates obstacles for studies focused on disability and to the equitable and effective administration of programs and policies intended for persons with disabilities. To facilitate agreement about the concept of disability, the World Health Organization (WHO) has developed a global common health language—one that is understood to include physical, mental, and social well-being. The WHO first published the International Classification of Impairment, Disabilities, and Handicaps (ICIDH) in 1980 as a tool for classification of the “consequences of disease.”

The newest version, International Classification of Functioning, Disability and Health, known as ICF (2001), moves away from a “consequences of diseases” classification (1980 version) to a “components of health” classification. This latest model is designed to provide a common framework and language for the description of health domains and health-related domains. Using the common language of the ICF can help health care professionals to communicate the need for health care and related services, such as the provision of AT for persons with disabilities (12).

In the context of health, the following language is used:

- **Body functions** are the physiologic and psychological functions of body systems.
- **Body structures** are anatomic parts of the body, such as organs, limbs, and their components.
- **Impairments** are problems in body function or structure, such as significant deviation or loss.
- **Activity** is the execution of a task or action by an individual.
- **Participation** is involvement in a life situation.
- **Activity limitations** are difficulties an individual may have in executing activities.
- **Participation restrictions** are problems an individual may experience in involvement in life situations.
- **Environmental factors** make up the physical, social, and attitudinal environments in which people live and conduct their lives (12).

Application of the WHO global common health language makes possible the definition of the need for health care and related services; defines health outcomes in terms of body, person, and social functioning; provides a common framework for research, clinical work, and social policy; ensures the cost-effective provision and management of health care and related services; and characterizes physical, mental, social, economic, or environmental interventions that will improve lives and levels of functioning. Provision of AT for persons with disabilities is an intervention that has the potential to diminish activity limitations and participation restrictions and in turn, improve the quality of life of individuals with disabilities. Throughout this chapter, the use of the WHO common health language is used to discuss the potential impact of AT.

**AT FOR MOBILITY IMPAIRMENTS**

Individuals with mobility impairments often present with unique needs and abilities. Some may demonstrate only lower-body impairment, such as a spinal cord injury or spina bifida, with no other complications. AT solutions might include crutches, a scooter, or a wheelchair. Simple modifications or adaptations to the environment, such as removing physical barriers to access (wide doorway or a ramp instead of stairs), may be all that is needed. For others, automobile hand controls, adapted saddles for horseback riding, sit skis for downhill skiing, or even placing bricks under a desk or table to allow the wheelchair user to work comfortably at a workstation may be necessary accommodations.

Other adaptive equipment for persons with mobility impairments might include a van with an attached lift. Many individuals who use wheelchairs drive a wide range of motor vehicles as well as bicycles using specially customized hand controls for turning and braking (Fig. 73-1). Chapter 78, Wheelchairs, and Chapter 76, Lower Extremity Orthotics, Shoes and Gait Aids, provide in-depth discussions of a wide range of assistive technologies for persons with mobility impairments that interfere with ambulation and other activities of daily living (ADLs).

For someone with upper-body mobility impairment, such as poor hand control or paralysis, assistive devices might include alternate keyboards or other input methods to access a computer. Alternate keyboards come in many shapes and sizes. There are expanded keyboards, such as the Intelliekeys (Fig. 73-2), which provides a larger surface area than a standard keyboard, larger letters with a contrasting yellow background, and options such as a delayed response of the activated

**FIGURE 73-1.** AT for mobility impairments. (Courtesy of Marlin Cohrs, Assistive Technology Partners, Department of Rehabilitation Medicine, University of Colorado Health Sciences Center.)
English alphabet, with the letters a, e, and the space bar placed in the center of the keyboard for easy access. Characters, numbers, and functions (including mouse control) keys fan out from the center of the keyboard depending on how frequently they are used by the average typist.

For individuals who are unable to use any type of keyboard or other device that requires hand mobility or finger pointing, there are a wide range of switches available that provide access not only to the computer but also to battery-operated electric toys and home-based or work appliances. These switches may be as simple as a “rocking lever” switch (Fig. 73-4) designed to be activated through a gross motor movement that involves touching or hitting the switch with the head, hand, arm, leg, or knee. Other switches can be activated by tongue touch, sipping and puffing on a straw, or through very fine movements, such as an eye blink or a single muscle twitch.

Fairly recent developments include eye gaze switches that calibrate intentional eye movement patterns and select targets such as individual keys on an on-screen keyboard, and also brain wave technology (Eye and Muscle Operated Switch or EMOS) that responds to the excitation of alpha waves to trigger the selection. Other input methods for the computer include devices such as the Head Mouse and Tracker 2000. These devices also rely on an on-screen keyboard visible on the computer monitor. The user wears a reflective dot on his or her forehead or other head-mounted signaling device to tell the computer what key or command he or she wishes to activate.

Individuals with disabilities, even those with the most severe motor impairments, can monitor and control an unlimited array of home, work, and school activities.

There are thousands of lower-tech assistive devices available for persons with motor impairments. Commonly referred to as aids for daily living, these devices include such things as weighted spoons, scoop plates, and other devices used to

FIGURE 73-2. Intellikeys keyboard. (Courtesy of Jim Sandstrum, Assistive Technology Partners, Department of Rehabilitation Medicine, University of Colorado Health Sciences Center.)

FIGURE 73-3. Tash Mini keyboard. (Courtesy of Marlin Cohrs, Assistive Technology Partners, Department of Rehabilitation Medicine, University of Colorado Health Sciences Center.)

FIGURE 73-4. Switches. (Courtesy of Marlin Cohrs, Assistive Technology Partners, Department of Rehabilitation Medicine, University of Colorado Health Sciences Center.)
communication impairments can benefit academically, vocationally, emotionally, and socially from the provision of a device that allows them to communicate their thoughts, learn and share information and ideas, and otherwise participate in life activities.

It is important to note these AAC devices, although extremely useful and important to the lifestyles of many individuals who are nonspeaking, do not replace natural speech. Instead they augment or provide an alternate form of communication. Other communicative modalities, such as vocalizations, gestures, sign language systems, eye gaze, etc., remain valid and acceptable forms of communication and should be encouraged to develop along with the facility to communicate out loud using an AAC device.

Congenital conditions such as autism, cerebral palsy, mental retardation, developmental verbal apraxia, and developmental language disorders can result in severe expressive communication impairments, necessitating the need for AAC interventions (13). Acquired disorders for which AAC is often used include traumatic brain injury, stroke, amyotrophic lateral sclerosis (ALS), tetraplegia, multiple sclerosis (MS), and laryngectomy that is due to cancer (7,14). Augmentative communication systems (Fig. 73-7) have also been used successfully for individuals who are temporarily nonspeaking as a result of ventilator dependence (15).

There are two common myths surrounding the use of AAC. The first is that individuals who are nonspeaking must demonstrate certain developmental prerequisites before the...
There are a number of low-tech voice output AAC devices available that are simple to program and use. The Talking Symbol Notepad, GoTalk, and TechTalk8 communication devices all use digital speech output (Fig. 73-9). Digital speech devices work much like a tape recorder. The person prescription of an AAC device. On the contrary, there are no strict sets of cognitive or physical prerequisites to the use of AAC devices. Specific AAC techniques that match the individual’s needs and abilities are chosen on the basis of a comprehensive evaluation by a qualified team of clinicians, family members, teachers, and others. The second myth is that the use of an AAC device will stifle or preclude the development or return of natural speech. However, the research tells us that use of an AAC device can actually support improved speech production and it in no way creates barriers to the development or return of natural speech (16).

Low-tech picture and alphabet boards are often used as either a preliminary step before purchasing an electronic voice output system, or as backup systems, should an electronic device break or its use not be convenient (e.g., during swim lessons). These low-tech systems (Fig. 73-8) can be made from picture library software available through commercial vendors (BoardMaker, PCS Symbols) or can be handmade using digital photographs, pictures from catalogs or books, or by simply using a marker to write or draw letters, words, phrases, or pictures. Adults with progressive diseases such as ALS or MS frequently use low-tech picture or alphabet boards to clarify their meaning as communication abilities decrease or as they fatigue during the day.

**FIGURE 73-7.** Using an AAC device. (Courtesy of Jim Sandstrum, Assistive Technology Partners, Department of Rehabilitation Medicine, University of Colorado Health Sciences Center.)

**FIGURE 73-8.** Low-tech AAC device. (Courtesy of Marlin Cohrs, Assistive Technology Partners, Department of Rehabilitation Medicine, University of Colorado Health Sciences Center.)

**FIGURE 73-9.** Using the Big Mac. (Courtesy of Diane Brians, Assistive Technology Partners, Department of Rehabilitation Medicine, University of Colorado Health Sciences Center.)
setting up the device simply holds down a button and records live voice into the microphone. The speech is digitized or recorded within the device. When the end user chooses to speak, he or she simply presses to activate the device. It then speaks the prerecorded message. The good news is these devices are simple and fairly inexpensive to purchase. It is important to bear in mind, however, that they are intentionally designed to communicate quick, simple messages, such as "hi," "let’s play," or "leave me alone." These devices are not appropriate for individuals who have more than a few things to say or who have the ability to generate complex thoughts and feelings.

There are a number of more sophisticated digital devices available as well. These include the DynavoxM3, Chatbox40 (Fig. 73-10), Springboard, and others. These devices have the capacity to encode several minutes of live voice and are often used by individuals who are not yet literate, have developmental disabilities, or simply wish to have a simpler device to use when going to the store or out to eat.

At the high end of AAC systems are the synthesized speech devices (Fig. 73-11). These voice output systems use built-in speech synthesis to speak messages that have been typed and previously stored into the device. They are capable of encoding several thousand words, phrases, and sentences. AAC devices such as these are frequently designed to provide an alternate access to the computer and control appliances such as TVs, door openers, and other electrical items found in many environments. Four popular text-to-speech devices are the ECO-14, LightWriter (Fig. 73-12), Tango, and the VMax. It is important to note that these AAC systems can cost thousands of dollars. But more importantly, they also form a critical link to the world for individuals with severe expressive communication impairments.

All of the AAC devices on the market can be activated through direct selection using either a finger or other pointing device, such as a head or mouth stick. A large number of the digital and text-to-speech devices can be activated using either direct selection or an alternate input mode. Alternate input includes all of the switches described above, as well as other additional infrared and wireless switches currently on the market. When an individual with a severe motor impairment and severe expressive communication impairment chooses to
use a switch to access the AAC device, the device can be preset to scan.

The most commonly used switch access method is called row column scanning. In this method, the end user starts by activating a switch to begin the scan. When the correct row is highlighted, he/she hits the switch again to start the lights scanning across the row (hence column). When the correct button is highlighted, the user hits the switch again to activate the chosen button. As can be surmised, using a scanning system to activate an AAC device can be a slow and tedious process. For many individuals, however, this method provides the only available access to spoken and written communication, and for those select persons it is highly prized as a window to the world.

One of the more vigorous debates within the field of AAC centers around the encoding or mapping strategies used to represent language on a communication device. Currently available systems vary greatly in vocabulary storage and retrieval methods, but all systems are based on communication symbols, whether orthographic or pictographic. Symbols vary in their transparency (guessability) and translucency (learnability). In order to select a set of symbols for an individual’s communication system, it is important to match these factors with the individual’s cognitive and perceptual abilities.

Additionally, the world of technology moves at a very rapid pace. AAC devices currently on the market have been designed using today’s technologies. As more and more innovations become available, such as cell phones, text messaging, and wireless global positioning systems (GPS), the capabilities and flexibility of these devices will change. As with all technologies, it is important that clinicians and practitioners remain current on these and future technology developments.

AT FOR VISUAL IMPAIRMENTS, INCLUDING BLINDNESS

The term visual impairment technically encompasses all degrees of permanent vision loss, including total blindness, which affects a person’s ability to perform the usual tasks of daily life. Low vision refers to a vision loss that is severe enough to impede performance of everyday tasks but still allows some useful visual discrimination. Low vision cannot be corrected to normal by regular eyeglasses or contact lenses.

For individuals with visual impairments, there are a variety of AT devices and strategies available to assist them to perform daily activities such as reading (Fig. 73-13), writing, daily care, mobility, and recreational activities. Among the low-tech solutions are simple handheld magnifiers, the use of large print, or mobility devices (e.g., long cane) for safe and efficient travel. High-contrast tape or markers can be used to indicate what an item is or where it is located within a physical plant.

Other low-tech devices include items such as wind chimes to facilitate direction finding, using easily legible type fonts such as Verdana (16 point or larger), and using vanilla- or beige-colored paper rather than white to improve visibility of text. For children and adults interested in recreational activities, solutions include beeper balls, three-dimensional puzzles, and “Braille Trails” specifically designed to improve access to wilderness and other outdoor activities (Fig. 73-14).

Many restaurants now provide large-print, Braille, and picture-based menus for customers with a variety of abilities. Books on tape are another readily available resource for

FIGURE 73-13. Low-tech magnification. (Courtesy of Marlin Cohrs, Assistive Technology Partners, Department of Rehabilitation Medicine, University of Colorado Health Sciences Center.)

FIGURE 73-14. Low-tech recreation aids. (Courtesy of Marlin Cohrs, Assistive Technology Partners, Department of Rehabilitation Medicine, University of Colorado Health Sciences Center.)
enable the end user to choose the amount and type of magnification he or she prefers for optimal computer access. A recent addition to the screen magnification software list is one called Bigshot. This software is less expensive ($199) and provides fewer features. However, it appears to be a nice alternative for computer users who do not need access to the more sophisticated computer functions, and it is affordable for public facilities such as libraries that wish to provide reasonable accommodations for customers accessing information databases.

**AT FOR LEARNING AND COGNITION**

Children and adults can present with a variety of learning and cognitive impairments resulting from either acquired or developmental disabilities. Not only can AT provide important accommodations for those with disabilities of all types, it can also become a critical tool to be used during recovery or acquisition of functional skill sets. For those with learning disabilities, there are a wide range of abilities that may benefit from an AT solution.

For example, many children with learning and cognitive impairments struggle with developing literacy skills. Fortunately, there are a number of both low- and high-technology solutions available to assist them. Many benefit from the use of specially designed software programs that predict the word or phrase they are trying to spell as they type the first letter(s) of a word (Co:Writer). Other software programs provide highlighted text and voice output so they are able to hear the words that they are generating (Write Outloud, Kurzweil 3000) on the computer. Individuals who are unable to read print materials often use some of the software solutions mentioned in the previous section, such as Jaws for voice output or books on tape.

Voice recognition has become a popular request both for persons with mobility impairments who are unable to type using their hands and for individuals whose learning disabilities are so significant they are unable to develop literacy skills. Voice recognition software enables an individual to speak into a microphone in order to input words, phrases, and sentences

---

**TABLE 73.1 Resources: American Federation for the Blind (AFB)**

<table>
<thead>
<tr>
<th>Resource</th>
<th>Contact Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>American Federation of the Blind Information Center</td>
<td>(800) AFB-LINE, (800) 232-5463</td>
</tr>
<tr>
<td></td>
<td>E-mail: <a href="mailto:afbinfo@afb.net">afbinfo@afb.net</a></td>
</tr>
<tr>
<td></td>
<td>(800) 824-2184</td>
</tr>
<tr>
<td>AccessWorld Solutions</td>
<td>E-mail: <a href="mailto:awsolutions@afb.net">awsolutions@afb.net</a></td>
</tr>
<tr>
<td></td>
<td>(800) 232-3044</td>
</tr>
<tr>
<td></td>
<td>or (412) 741-1398</td>
</tr>
<tr>
<td>Journal of Visual Impairment and Blindness and Access World</td>
<td>E-mail: <a href="mailto:afbsub@abdintl.com">afbsub@abdintl.com</a></td>
</tr>
<tr>
<td></td>
<td>(212) 502-7642</td>
</tr>
<tr>
<td>National Technology Center</td>
<td>E-mail: <a href="mailto:techctrn@afb.net">techctrn@afb.net</a></td>
</tr>
</tbody>
</table>

**FIGURE 73-15.** Zoomtext Xtra. (Courtesy of Marlin Cohrs, Assistive Technology Partners, Department of Rehabilitation Medicine, University of Colorado Health Sciences Center.)
into standard computer word processing programs such as Microsoft Word. Although a bit trickier, voice recognition can be used for input and control functions for other software, such as database programs and Windows.

Although voice recognition software is a rapidly developing technology, it is important to remember that it takes a fifth- to sixth-grade reading ability to train the standard software, simply because the individual’s voice file cannot be developed unless he or she uses the standard training package that comes with the software. Dragon Systems has developed a version for children, but its success rate for children with learning and other cognitive disabilities has not yet been published. In addition, ambient noise, such as in a typical classroom, as well as fluctuating vocal abilities (e.g., fatigue) found in many individuals with disabilities will have an impact on the accuracy of the voice recognition. In general, it takes more than 20 hours to train the software to an acceptable level of accuracy (>90%). Although caution is in order when prescribing this type of software, the rapid pace of development bodes well for future use of this solution for persons with disabilities.

Low-tech adaptations for persons with learning and/or cognitive impairments can include colored highlighter tape, pencil grips, enlarged text, and other easy modifications, such as using a copy holder to hold print materials for easy viewing. Reminder lists with important times, places, and activities highlighted with a marker are often useful for individuals who need subtle memory prompt to be in the right place at the right time.

Recent technology developments include the handheld personal digital assistants or PDAs. AT software developers (AbleLink Technologies, Inc.) have taken this technology a step further by developing a software application named PocketCoach (Fig. 73-16) that provides auditory prompts for individuals with cognitive disabilities. This software can be set up to remind someone what step they need to take next. It can be used for something as simple as a vocational task, such as mopping a floor, all the way to the complexity of prompting someone through the steps of a math problem. The latest edition of this software combines voice prompts with visual prompts (Visual Assistant). The individual setting up the system can simply take a picture with the accompanying digital camera and combine the pictures with digitally recorded voice prompts to further facilitate memory and cognition.

A number of software packages are available for these populations that focus on a range of topics, including academics, money management, personal skills development, behavior training, development of cognitive skills, memory improvement, problem solving, time concepts, safety awareness, speech and language therapy, telephone usage, recreation, and games. For individuals with cognitive and learning impairments, a number of simple techniques can be kept in mind when evaluating or designing activities and materials for use by persons with cognitive and learning impairments. Table 73-2 includes some components of accessible interfaces to keep in mind.

**AT FOR HEARING**

For an individual who is deaf or hard of hearing, there are two major effects of hearing loss: lack of auditory input and compromised ability to monitor speech output. AT devices, such as hearing aids and FM systems can often be used to facilitate both auditory input and speech output. Other types of AT devices provide a visual representation of the auditory signal. These include flashing lights to indicate an emergency alarm (fire, tornado), the phone ringing, or someone at the door.

Environmental adaptations can frequently support individuals who are deaf or hard of hearing. For example, when speaking to someone who has difficulty hearing, do not stand...
in front of a light source (windows, lamps, etc.), and do not over exaggerate lip movements, but do include gestures, which can be helpful. For individuals who wear hearing aids, there are a number of additional technologies that can facilitate hearing in large rooms or in crowded environments, such as a restaurant. The Conference Mate and Whisper Voice are two products especially designed for this issue. The person with the hearing loss wears a neck loop (it looks much like a bolo tie). In the case of the Conference Mate, a small octagonal device is placed on a convenient table. This device picks up voices within the room that are transmitted to the neck loop and then to the hearing aid for better reception. Although it may sound cumbersome, it can be an excellent solution for office- and school-based environments. The Whisper Voice works much the same way, except the device contains a small microphone and is portable. It can be passed from one speaker to another, with sound transmitted to the neck loop and then to the hearing aid for amplification. Chapter 15, Speech, Languages, Swallowing and Auditory Rehabilitation contain additional information on hearing loss and auditory rehabilitation.

### TABLE 73.2 Components of Accessible Interfaces

<table>
<thead>
<tr>
<th>Simplicity of layout, operation, and appearance</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Is the interface crowded, complex, or otherwise overwhelming?</td>
</tr>
<tr>
<td>• Does the interface require complex mouse actions or keystroke combinations?</td>
</tr>
<tr>
<td>• Is the language level of the interface too complex for the user?</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Consistency of critical elements in the interface</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Are interface elements and controls located consistently throughout the application?</td>
</tr>
<tr>
<td>• Are interface elements and controls activated in the same manner throughout the application?</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Saliency of active elements</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Is the user directed visually or otherwise toward the central content of the interface?</td>
</tr>
<tr>
<td>• Is key information highlighted?</td>
</tr>
<tr>
<td>• Is there a minimum of competing information?</td>
</tr>
<tr>
<td>• Is it clear to the user when actions or changes occur in the interface?</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Intuitive operation</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Is the operation of the software obvious to the user?</td>
</tr>
<tr>
<td>• Is the selection of user interface components simple and direct?</td>
</tr>
<tr>
<td>• Is it clear to the user that the interface is responding to his or her input?</td>
</tr>
<tr>
<td>• Is there clear feedback when the application is busy completing a task?</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Organization</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Are similar functions grouped logically together?</td>
</tr>
<tr>
<td>• Does the organization of the interface make sense visually?</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Adaptability</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Does the interface offer a choice of modalities for the user (e.g., can on-screen text be read aloud)?</td>
</tr>
<tr>
<td>• Can interface elements be added or removed easily to adjust to a user’s abilities?</td>
</tr>
<tr>
<td>• Does the interface offer context-sensitive help, such as tool tips?</td>
</tr>
<tr>
<td>• Are on-screen instructions provided?</td>
</tr>
<tr>
<td>• Are cues provided automatically to the user if he or she waits for assistance?</td>
</tr>
<tr>
<td>• Is there a timed response?</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Recoverability</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Can the user easily recover from an error?</td>
</tr>
<tr>
<td>• Is clear warning provided if an action cannot be undone?</td>
</tr>
<tr>
<td>• Can the user explore the interface safely without causing instability?</td>
</tr>
</tbody>
</table>

### AT FOR ERGONOMICS AND PREVENTION OF SECONDARY INJURIES

A rapidly growing area of concern for AT practitioners is the development of repetitive motion disorders. For many persons with disabilities, the use of computer keyboards and other technology presents an opportunity for secondary injuries to occur. Computer desks, tables, and chairs used in computer labs, classrooms, and at the office do not always match the physical needs of the end user. When those with disabilities (and those without) are not positioned properly and then spend hours repetitively performing the same motor movement, they can and do develop (incur) injuries.

An entire industry of AT has developed over the past few years dealing with repetitive motion disorders. Potential solutions for someone demonstrating this type of impairment include raising or lowering a chair or desk for the appropriate fit, implementing routine breaks within activities for the individual to move about or do something different, lumbar and other supports, specially designed ergonomic keyboards, and other ATs.
ELECTRONIC AIDS TO DAILY LIVING

Electronic Aids to Daily Living (EADLs), also described as environmental control units, provide alternative control of electrical and battery operated devices within the environment. These devices may include the TV, VCR, stereo, lights, appliances, telephone, door, electric bed, and more. EADLs are designed to improve independence of ADLs. EADLs are primarily used in the home, but can also be used at work and school.

EADLs provide alternative control and are designed for persons who are unable to use standard controls such as light switches or other electronic or battery-operated appliances and fixtures. EADLs can be helpful for persons with physical and cognitive disabilities. For example, a person who has aphasia and motor impairments that are due to stroke may not be able to easily select a new TV channel. This person may benefit from a solution as simple as a standard TV remote control.

A person with cerebral palsy may have difficulty with small buttons on a remote control. This person may benefit from an EADL that is accessed by a switch to scan choices. Persons of all ages can benefit from this technology as well. For example, entry level EADLs provide alternative control of toys for very young children.

EADLs are controlled by three different access methods: direct, switch, and voice. Direct access is generally finger-to-button, as on a standard remote control. Some EADLs have enlarged buttons or keyguards to assist direct access. Enlarged buttons can also make the buttons easier to see. Typically, individuals who use direct EADLs have fair to good fine motor control and vision.

In switch access, any type of switch can be placed at the best location for activation by the person. The first switch activation begins a scan of choices, usually of general categories (i.e., TV, lights, phone). The second switch activation chooses one of these categories. Choices within that category are now scanned (i.e., channel up, channel down, mute). A third switch activation selects the desired function, and the signal is sent to the TV. Most of these systems have visual displays with small text in English (although some EADLs are available in other languages) and no speech feedback. The person generally must have good sequencing skills and vision and be able to read.

Voice-operated EADLs respond to verbal commands. For example, if the user says “TV on,” a signal is immediately sent to the TV to turn it on. The individual using the device needs to have a consistent, understandable voice to operate these EADLs. They must also remember the available commands or be able to read a list to remind themselves. A person with a high-level spinal cord injury is a typical patient who could benefit from this type of device.

Many augmentative communication devices are capable of sending signals to control devices within the home environment. This can be advantageous for several reasons. First, because the EADL capability is built in, no additional funding is needed. Second, these AAC devices allow the use of larger text, graphics, and auditory scanning. Auditory scanning verbally announces each choice as it is scanned. Depending on the AAC device, these auditory prompts can even be recorded in another language. This can be very helpful for individuals with low or no vision, who need prompts to assist with memory or sequencing challenges, or who do not read (or do not read English). If they are already using a communication device, these features can be easily programmed to increase independence in the home. Some individuals who do not require an AAC device still use one solely for the EADL features because

<table>
<thead>
<tr>
<th>Table 73.3 Internet-Based Resources</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Resource</strong></td>
</tr>
<tr>
<td>Occupational Safety and Health Administration (OSHA)</td>
</tr>
<tr>
<td>MacWorld—Parent’s Guide to Ergonomics</td>
</tr>
<tr>
<td>UCLA Ergonomics</td>
</tr>
</tbody>
</table>
of the visual and cognitive advantages. AAC devices can be accessed directly by switch scanning, mouse control, or joystick.

Finally, software and hardware are available for computers to provide control of devices within the environment. Typically, this technology is designed for computer users who are interested in automating their home. These EADLs are not designed for persons with disabilities and can be challenging to use both cognitively and via motor control. Whatever method a person is using to access his or her computer (i.e., keyboard) is what is used to access the EADL features.

People must both feel and be safe in their home, especially if they are left alone for any period of time. Primary safety concerns are control of the telephone and door opener. The person with a disability must have independent access to both, if at all possible. The telephone is crucial for calling emergency personnel (i.e., medical emergency, possible burglary), calling when an attendant does not show up, and calling when a caregiver becomes incapacitated. Of course, phone access is also dependent on his or her ability to judge an emergency appropriately and call emergency services when warranted. The phone must have battery backup, as some emergencies are linked to power outages.

If the person is mobile and the home entrance is accessible, a power door opener may be necessary. However, if the individual has cognitive limitations affecting judgment and should not leave the home without supervision, independent control of the door may not be appropriate. Independent control of a door opener is crucial for leaving the home in an emergency such as a fire and letting in caregivers and visitors so safety and security are not compromised.

Considerations in EADL Selection
Features that are essential to consider in EADL selection are (a) portability, (b) whether the client needs to use the EADL from their bed, and (c) safety. A portable device is very important for an individual who is mobile within the home (i.e., driving a power wheelchair) and for people who need to use the EADL from more than one position, such as from a wheelchair and also from bed. A portable EADL can be moved from one location to another.

Accessing the EADL from bed is very important for many persons with disabilities. Alerting a caregiver of a need in the night, controlling an electric bed, turning on a light, and turning off some quiet music to help someone go to sleep are just a few examples. Motor control can change dramatically from sitting to lying down, which may require a different access method in general or two different access methods (i.e., hand switch in the wheelchair and head switch in bed).

There are a number of important questions to be discussed during the EADL selection process. These include
1. What environments will the person be in, and how much time is spent in each setting?
2. What appliances or devices need to be accessed?
3. Will needs and abilities change throughout the day or in different environments?
4. What are the individual’s cognitive and sensory abilities? (17)

AT specialists should assess for the most appropriate EADL based on what devices the person would like to control, as well as motor, cognitive, and sensory skills. The AT specialist or the AT supplier or vendor can also assist with procuring funding, installing equipment in the home, and training the end user and caregivers in its use.

 Appropriately prescribed EADLs result in a more productive and satisfying lifestyle. In the hospital setting, EADLs allow patients to experience greater independence and can decrease the level of required nursing supervision and support. In the home, an EADL allows an individual more independence and flexibility while demonstrating the cost benefit of decreased home health care and assistance.

TEAMING AND AT
"Teaming means you work together, no matter what. You do it because you'll come up with better ideas. And, if (or when) you disagree, you just figure it out—without fighting" (18). Equal participation in the collaborative teamwork process by the person with a disability, family members, and service providers is critical for individuals to achieve their goals. As it is in so many areas of physical medicine and rehabilitation, AT services are delivered in a wide variety of settings, including comprehensive medical rehabilitation centers, university-affiliated clinics, state agency-based AT programs, private rehabilitation engineering and technology firms, and nonprofit disability organizations (19). Because AT is a relatively new field and preservice and in-service preparation is just beginning to register an impact, persons with disabilities may encounter difficulty locating experienced and credentialed professionals to deliver AT services.

A transdisciplinary model of service delivery is preferred as it provides a larger pool of resources and expertise. The team may include occupational and physical therapists, rehabilitation engineers, speech language pathologists, physiatrists, case managers, and other professionals identified as important to meeting the individualized goals of the person with a disability. It is critical that the team includes as recognized members the person with the disability(s), and his or her family and significant others when appropriate. It is also critical that at least one member of the team has some background knowledge and training in the field of AT. There are a number of university and online courses available for professionals to build expertise as well as resources in the field of AT.

The field of AT, like many other growing professions, is working to develop standards of practice and credentialing opportunities. The Rehabilitation Engineering and Assistive Technology Society of North America (RESNA), an interdisciplin ary association for the advancement of rehabilitation and
to achieve as a result of the evaluation (i.e., equipment ideas, potential success with vocational or educational objectives).

When selecting team members to conduct an AT evaluation, appropriate disciplines should be chosen based on the identified needs of the person with the disability. For example, if the individual presents with both severe motor and communication impairments, then an occupational or physical therapist with expertise in AT, as well as a speech-language pathologist with a background in working with persons with severe communication impairments and alternative forms of communication, should be included as members of the team. If a cognitive impairment has been identified during the intake process, someone versed in learning processes such as a psychologist, neurolinguist, teacher, or special educator might be appropriate members of the team. If there is an ergonomic issue (i.e., carpal tunnel syndrome), an evaluator with training in ergonomic assessment or a background in physical or occupational therapy is a necessary component for a successful experience.

It is not appropriate for an AT vendor to be called in to perform the AT evaluation. Although vendors can and should be considered as identified members of the team, it must be recognized that they have an inherent conflict of interest. They are there to sell products. When requested by the team, vendors should demonstrate their products, discuss pertinent features, and assist in setting up the equipment for evaluation and trial usage. However, other team members, including the end user and their family, should carry out the actual evaluation and make the final recommendation(s).

### Phase I of the Assessment Process

Knowledge within the field of AT continues to grow and change, sometimes on a daily basis. As this evolution continues, a number of important variables are being identified that directly impact whether the AT recommended by the assessment team will be used or abandoned by the consumer (9,20). As a result of this information, the evaluation process continues to be refined. Many researchers are working to develop standardized AT measurement tools (9,21,22), but the fact remains there are few available resources to guide practitioners who have not received formalized training in AT.

### ASSESSMENT USING A TEAM APPROACH

The goal of any AT evaluation is to determine whether the individual receiving this service has the potential, and the desire, to benefit from AT devices and services at home, school, work, or play. Other outcomes of an AT evaluation include providing a safe and supportive environment for the person with a disability and his or her family to learn about and review available assistive devices; identifying necessary AT services such as training, modification, etc., that may be necessary for the equipment to be effective; and developing a potential list of recommended devices for trial usage before a final determination is made. Also, the individual and family, as well as the involved professionals, should specify exactly what they hope to achieve as a result of the evaluation (i.e., equipment ideas, potential success with vocational or educational objectives).

AT, has developed guidelines and credentialing examinations for three categories of specialists in AT (Table 73.4).

For more information about the credentialing process and criteria for credentialing, contact RESNA at http://www.resna.org. A number of universities throughout the United States and Canada offer training in AT for a wide range of audiences. Table 73-5 lists only a few of the available options. For more information, visit the Web and type assistive technology training into the search engine.
As mentioned earlier in this chapter, the number one reason AT is abandoned is because the needs and preferences of the consumer were not taken into account during the evaluation process. Other reasons cited for abandonment of devices include:

- Changes in consumer functional abilities or activities
- Lack of consumer motivation to use the device or do the task
- Lack of meaningful training on how to use the device
- Ineffective device performance or frequent breakdown
- Environmental obstacles to use, such as narrow doorways
- Lack of access to and information about repair and maintenance
- Lack of sufficient need for the device functions
- Device aesthetics, devices, weight, size, and appearance (8)

Careful review of these factors suggests that many of these issues can be considered during the evaluation process. At the University of Colorado Denver, School of Medicine, the assessment protocol has evolved from a group of practitioners trying any number of devices with the individual to a team process that starts by leaving the technology out of sight. The process about to be described may sound laborious and cumbersome. With practice, we have reduced the time necessary for the evaluation process and have increased the likelihood that the individual will be using it selects the appropriate technology. In addition, this process has decreased both installation and follow-up training time and has resulted in improved outcomes for end users.

Phase I of the assessment process is initiated once a referral is received. Standard intake information is collected, usually over the phone that provides the name, primary diagnosis, age, reason for referral, etc. In the majority of cases, cognitive, motor, vision, and other standard clinical assessments have already been performed, and a release of information is requested from the individual or his or her caregivers for this information to be forwarded to the team. If it has not previously occurred, these evaluations are scheduled as a component of Phase II of the assessment process.

Based on the preliminary information, an appropriate team of professionals is assembled and a date chosen for the evaluation. The team leader takes responsibility for ensuring that the individual with the disability, his or her family, and any other significant individuals are invited to the evaluation. It is not unusual to vary the schedule to meet the needs of the family rather than the professionals.

When the time arrives for the evaluation, the team members are invited to gather and spend some time getting to know the individual. Using methods described by Cook and Hussey (22) and Galvin and Scherer (8), the team first identifies the life roles of the consumer (e.g., student, brother, musician, etc.). Then the specific activities engaged in by the individual to fulfill that role are identified. For example, if he is a brother, then that means he may play hide and seek with a sibling, squabble over toys, or otherwise engage in brotherly activities. If he is a musician, then he may want or need to have access to musical instruments, sheet music, or simply a radio.

Next, the team identifies any problems that may occur during these activities. For example, the musician may not have enough hand control to play the piano or may experience visual or cognitive difficulties with sheet music. Specific questions are asked regarding where and when these difficulties occur (activity limitations). Perhaps problems occur when the individual is tired, or not properly positioned, or when he or she tries to communicate with others. The individual is also asked to describe instances of success with these activities and to discuss what made them successful (prior history with and without technology). Interestingly enough, the team by now is usually able to recognize patterns of success and failure from the individual’s perspective, as common themes across environments emerge.

Finally, we ask the team to prioritize the order in which we can address identified barriers to participation, and a specific plan of action is developed. Within the specific plan of action, “must statements” are also developed. For example, the device must have a visible display in sunlight, or the technology chosen must weigh less than 2 lb. In one instance, the must statement read, “It must be purple.”

It is at this point that the team may be reconfigured. For example, if the individual is not properly seated and positioned, they are referred first to the occupational or physical therapist for a seating and positioning evaluation before any other technology issues being addressed. At all times, the configuration of the team includes the individual being assessed and the caregivers as the primary members to be consulted.

In many instances, various members of the team in collaboration with others determine that further assessment from their perspective is not warranted for the technology component to proceed. In other situations, it is determined that additional team members who were not previously considered should be invited to participate (e.g., vision specialist).

**Phase II of the AT Assessment**

Once the team has agreed on the specific plan of action and those things that must occur, phase II of the assessment process begins. The person with a disability and his/her caregivers are asked to preview any number of assistive devices that may serve to reduce activity limitations and increase participation in their chosen environments. These ATs are tried with the individual, and various adaptations, modifications, and setups are explored to ensure an appropriate match of the technology to the individual is made.

It is at this point that the AT skill sets of the clinician become critical. If trial devices are not properly configured or if the wrong information is given to the consumer, then they will be unable to make an appropriate selection. Because so many devices require extensive training and follow-up, it is also critical that realistic information regarding training issues (including learning time) is provided and appropriate resources within the local community be identified.
In a number of instances, the technology that appears to be optimal for an individual does not carry with it the appropriate community supports. In those cases, it is often advantageous to work first to identify local resources or local AT professionals willing to receive additional training before sending the device home with the end user. At all times, the end user and his or her families should be informed and updated so that they can make the final decision regarding when and where they wish the equipment to be delivered.

With very few exceptions, the wise course of action involves borrowing or renting the equipment before a final purchase decision. For many individuals with disabilities, the actual use of various technologies on a day-to-day basis elicits new problems that must be resolved. Unexpected benefits, including changes in role and status, also occur as a result of improved functioning. In some cases, these unexpected benefits create an entirely new set of problems that must be addressed. For most, these disruptions can be resolved with time and energy. Others decide that they either prefer the old way of doing things, or they are interested in adding or changing the technology once they have had a chance to experiment with it in different settings.

**Writing the Evaluation Report**

When writing the evaluation report for an AT assessment, it is important to ensure that a number of items are included. First and foremost, case managers, educators, and others unfamiliar with assistive technologies appreciate layman’s terms when discussing the need for AT and what it will accomplish.

In cases in which medical insurance is being used to purchase the technology, it is critical to document the actual medical necessity for the device(s). For example, “Mrs. Smith will use this device to communicate her health care needs and to meet the functional goals outlined in the attached report.”

In instances when the evaluation was requested to determine educational or vocational benefit with assistive devices, it is important to document how these specific needs will be met with the prescribed equipment.

It is extremely important that all components of the assistive devices be included (e.g., cables, ancillary peripherals, or consumable supplies) in the list of recommended equipment. In many instances, devices are recommended for purchase as a “system.” When this occurs, acquisition and implementation can be delayed for months because an item was not included in the initial list. It is also important to include contact information for the various vendors who sell the equipment. Many purchasers are unfamiliar with these companies, and acquisition can be delayed for months if this information is not included in the report.

**FUNDING AT**

The funding sources for AT fall into several basic categories (Table 73-6). One source to be investigated is private or government medical insurance. Medical insurance defines AT as medical equipment necessary for treatment of a specific illness or injury. A physician’s prescription is usually required. When writing a prescription for an AT device, it is important that the physician is aware of the costs and benefits of the devices he or she is prescribing and is prepared to justify his or her prescriptions to third-party payers. Funding includes not only the initial cost of the device, but the expense involved in equipment maintenance and patient education, as well as the potential economic benefits it provides to the patient (e.g., return to work).

According to a publication sponsored by the American Medical Association (23), the following items (reprinted with

---

**TABLE 73.6  Potential Funding Sources for AT**

<table>
<thead>
<tr>
<th>Public Programs</th>
<th>Alternative Financing</th>
<th>U.S. Tax Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medicare</td>
<td>Private insurance</td>
<td>Medical care expense deduction</td>
</tr>
<tr>
<td>Medicaid—Early and Periodic Screening, Diagnosis, and Treatment (EPSDT)</td>
<td>Private foundations</td>
<td>Business deductions</td>
</tr>
<tr>
<td>State grants</td>
<td>State loan programs</td>
<td>Americans with Disabilities Act credit for business</td>
</tr>
<tr>
<td>Individuals with Disabilities Education Act (Parts B and C)</td>
<td>Employee accommodation programs</td>
<td>Charitable contributions deduction</td>
</tr>
<tr>
<td>Vocational rehabilitation state grants, including Title VH, Chapter 2</td>
<td>Corporate-sponsored loans</td>
<td>Targeted jobs tax credit</td>
</tr>
<tr>
<td>The Developmental Disabilities state grants</td>
<td>Community reinvestment programs</td>
<td></td>
</tr>
<tr>
<td>Workers Compensation Programs</td>
<td>Family and friends</td>
<td></td>
</tr>
<tr>
<td>Social Security Supplemental Security Income PASS Program</td>
<td>Religious organizations</td>
<td></td>
</tr>
<tr>
<td>CHAMPUS/TRI-CARE</td>
<td>Community groups</td>
<td></td>
</tr>
<tr>
<td>Department of Veterans Affairs</td>
<td>Service clubs and advocacy organizations</td>
<td></td>
</tr>
</tbody>
</table>
permission) should be taken into account when prescribing AT and certifying medical necessity:

1. The physician must provide evidence of individual medical necessity.
2. An “appropriate” prescription is one that takes into consideration the comprehensive assessment process, including motivation and availability of training, the potential patient functional outcome, and the cost/benefit of available products.
3. Physicians should be prepared to provide sufficient information to insurance companies to ensure approval. Dialogue is often necessary to show medical necessity of complex assistive technologies (power wheelchairs, computer-based environmental control systems).
4. Basic knowledge of AT reimbursement for patient and physician includes familiarity with established medical necessity forms and prior authorization procedures.
5. Avoid making static decisions on a dynamic problem; anticipate future need.
6. Base decision on both expected performance and durability of the device.

**Documentation in the Medical Record**

In addition to prescribing and certifying medical necessity on various forms, physicians must be sure to maintain complete patient records that include the following information:

- Patient diagnosis or diagnoses
- Duration of the patient’s condition
- Expected clinical course
- Prognosis
- Nature and extent of functional limitations
- Therapeutic interventions and results
- Past experience with related items
- Consultations and reports from other physicians, interdisciplinary team, home health agencies, etc.
- Complete listing of all assistive devices the patient is using, including copies of prescriptions and certification forms or letters
- Tracking system for device performance, including follow-up assessment schedules and lists of professional and vendor names to contact if problems occur

**Letters of Medical Necessity**

These letters should include the following areas:

1. Diagnoses 1CD-10-CM codes
2. Functional limitation(s) (a partial list of disabilities as examples follows):
   - Balance disorder
   - Developmental delay
   - Hypotonia
   - Joint deformity or instability
   - Hemiparesis
   - Side affected or bilateral
   - Diaparesis
   - Paraparesis

3. “Because of the patient’s functional limitation, he/she is unable to…”
   a. Perform
      i. ADLs
      ii. Instrumental ADLs
      iii. ADLs and functional mobility
      iv. Functional mobility
      v. Work activities
   b. Communicate
      i. Verbally
      ii. In writing
      iii. Independently over the phone
   c. Use of equipment “The use of the equipment will/allow the patient to…”
      - Function independently
      - Function independently with the device/equipment
      - Perform independent wheelchair mobility in the home
      - Perform independent wheelchair mobility in the home and community
      - Return home
      - Be required as a lifetime medical need (if shorter duration, explain need)
      - Improve the patient’s functional ability

5. Description of equipment (a partial list as examples follows):
   a. Wheelchair
   b. Wheelchair frame
      i. Electric
      ii. Manual
      iii. Manual backup
      iv. One-arm drive
      v. Lightweight
      vi. Nonstandard
      vii. Reclining
      viii. Miscellaneous
   c. Other
      i. Bathing aids
      ii. Toileting aids
      iii. High-technology vision enhancers
      iv. Other hearing assistive devices
      v. Hospital bed
      vi. Stander

6. Rationale (a partial list as examples follows):
   - Safety or safe positioning for an activity
   - Cost-effectiveness in prevention of secondary complications (e.g., pressure sores)
   - Mobility restrictions preventing independent activity
   - Access to areas in home, such as bathroom and kitchen
   - Access to workplace, school
   - Duration of expected use
   - Past experience, interventions, and results (failure of less expensive solutions)
   - Duration of expected use
   - Goals and benefits to patient

AT is usually covered under policy provisions for durable medical equipment, orthotics and prosthetics, or daily living...
PART VI | MANAGEMENT METHODS

and mobility aids. With private insurance, AT providers request funding under the specific provisions of the individual policy, appealing any denials (an inevitable) and offering medical justification for coverage. With government insurance policies, such as Medicaid and Medicare, coverage is based on existing law and regulations. In 2002, regulations were promulgated by Medicare to include coverage of AAC devices (24).

Information on covered services and how to request funding is available from the Medicaid programs in individual states and from the regional offices for Medicare. AT professionals and other health care providers should continually advocate for adequate coverage of AT in all health care plans.

Funding of AT is also available from other federal and state government entities, such as the Veterans Administration, State Vocational Rehabilitation, Rehabilitation Services Administration, State Independent Living Rehabilitation Centers, and State Education Services. Local school districts may fund education-related AT for children. Each agency or program sets criteria for the funding of AT based on the mission of the agency and the purpose of the technology. For example, vocational rehabilitation agencies generally pay for devices to facilitate gainful employment, and education program funding is directed toward enhancing the client’s performance in school.

Private funding is often available through subsidized loan programs, churches, charitable organizations, and disability-related nonprofit groups. The AT provider must keep abreast of the requirements of various funding sources in order to direct the client to appropriate organizations. Often a combination of funding from several sources is needed to reduce personal out-of-pocket costs. Because funding for replacement of AT devices is also difficult to obtain, careful selection of the initial device is required. Providers can also assist clients by considering funding when making equipment recommendations by including both low- and high-cost alternatives with their relative advantages. Funding is generally available for AT, but persistence and advocacy by the AT provider are required for success (25,26).

MEASURING OUTCOMES IN AT

The study of the impact of AT devices for individuals with disabilities poses a challenge in outcomes research. The field itself is a multidisciplinary area of study encompassing medicine, rehabilitation, psychology, education, engineering, and biotechnology specialties, and involves physical, cognitive, psychosocial, sensory, and physiological effects. Consequently, there is a lack of consistency in what has been studied, how the outcomes have been measured, and where the results have been recorded. In the field of AT, there is also a paucity of outcomes measurement research in general (27).

More than 10 years ago, Frank DeRuyter of the Rancho Los Amigos Medical Center in Downey, California, stated that “evaluating AT services to demonstrate quality or to measure outcomes is the ethical obligation of the entire AT community” (28, p. 3). He went on to remind us that the systematic application of outcomes management research within AT has been fragmented and limited at best, and suggested it is incumbent upon all stakeholders in the AT community to evaluate the value and outcomes in all aspects of services delivery. In 1996, RESNA dedicated an entire issue to AT outcomes. In his editorial, Larry Trachtman again reminded AT professionals to develop a methodology for measuring and reporting outcomes.

He argued that no accepted way exists to collect data in order to verify trends or to support or refute practices. As a field, there is little or no agreement on the measures, data collection points, or even the desired outcomes (27, pp. 67–68).

Persons with functional limitations and the AT devices provided by professionals do not operate in a vacuum. They exist on a broad continuum and are impacted by such things as environmental and psychosocial issues, family finances, cultural differences, and other contextual factors. Services are often fragmented, with many consumers receiving interventions from any number of teams and facilities. It is not unusual to hear families talk about their school team, hospital team, and any number of private therapists as independent service providers who do not interact. Rarely are discussions held regarding appropriateness of devices across environments, cost-effectiveness, or prevention of secondary conditions. General agreement within the field suggests that outcome measurement is a critical, unmet need. But a conceptual framework for developing measurement tools and measurement research has remained elusive (28–32). Studies of the treatment efficacy of AT devices and services have typically been relegated to single case study reports and occasional multiple case reports showing changes from baseline (33,34).

AT traditionally includes the prescription and implementation of devices for sensory augmentation (speech, hearing, vision, etc.), but in recent years, the concept of AT has been broadened to encompass any technology that can improve a person’s function (35). This is an important distinction, because it places nonoperative rehabilitation interventions, such as orthotics, prosthetics, electrical stimulation, and functional neuromuscular stimulation, in the realm of AT.

The application of technology to improve human function has long been the goal of the AT professional. In many cases, clinicians working in AT have been the most successful at crossing traditional clinical boundaries to reach out to their health care partners who may be less familiar with AT, producing collaborations that are both innovative and productive. The AT specialist has the hands-on clinical experience to see what works and understands those factors leading to technology abandonment. Typical clinical practice, however, does not lend itself to the development of experimental methodologies to objectively evaluate patient performance with AT devices and services. Moreover, most AT clinicians do not have the resources to actively participate in a sustained program of research, nor are these behaviors emphasized as a component of clinical intervention in most training programs.

Despite this limitation, AT professionals and the AT service delivery model have been effective in getting technology into
the hands of the people who need it, creating a foundation for rehabilitation intervention service delivery in general. Because the AT specialist functions across disciplines, he or she is often the first to notice the impact of other treatment modalities. For example, it is typical for a child with an acquired disability to enter rehabilitation services with a variety of needs and assignments to various disciplines for treatment. It is often the AT specialist who notices incompatibilities between systems, such as a seating system with a lap tray that interferes with a child in the development of an alternate access method to a computer used to complete educational tasks.

In recent years, the National Institute on Disability and Rehabilitation Research and the National Institutes of Health’s National Center for Medical Rehabilitation Research have begun to fund various research activities devoted to developing standardized outcomes measurement systems in order to determine the efficacy of various AT devices and services. The plan for these activities calls for the dissemination of information to individuals with disabilities, their families, caregivers, funding sources, and manufacturers. The field of AT is one of growth and excitement. Results from studies such as these are a welcome and necessary component for the continued development of this discipline.

THE FUTURE OF ASSISTIVE TECHNOLOGIES

The future of assistive technologies and universal design is exciting to contemplate. Intensive research and development activities are focusing on human-robot interaction (HRI), GPS, and a wide range of other context aware sensor technologies. Advances in the commercial manufacturing industry is leading the way by designing cheaper, less intrusive support systems for persons without disabilities.

Researchers and engineers focused on these technologies are quickly taking advantage of new discoveries as well as adding to the body of knowledge and the array of products available for persons with disabilities. HRI is often frequent and intense in manufacturing and service environment(s). Utilizing a soft computing toolbox approach, especially with fuzzy set-based learning techniques, can be effectively adopted for modeling human behavior patterns as well as for processing human biosignals including facial expressions, hand/body gestures, Electromyogram (EMG), and so forth (36).

With over 600,000 people each year surviving a stroke, it has become the leading cause of serious long-term disability in the United States. Studies have proven through repetitive task training, neural circuits can be remapped, thus increasing the mobility of the patient. This fuels the emerging field of rehabilitation robotics. As technology advances, new therapy robots are developing that are increasingly compliant and captivating to use. One example is the robotic gait trainer (RGT) developed in the human machine integration laboratory at Arizona State University. The RGT is a tripod mechanism, where the patient’s leg is the fixed link, controlled on a Matlab and Simulink platform. This device facilitates retraining of gait and movement patterns impeded as a result of stroke or injury (37).

Context aware sensors and GPS systems are emerging technologies that can be useful in identifying where someone is; the status of his or her environment (temperature, location); and, changes in positioning, weight, etc. Utilization of these technologies can alert caregivers when a patient with dementia is attempting to open a door, taking a medication, lost, or otherwise impaired within their current environment. These same technologies are also proving to be an effective tool for persons with intellectual/developmental disabilities, MS, and other debilitating injuries or illnesses.

The field of AT is one of rapid change and growth. New technologies, new research possibilities, and a growing universal acceptance of AT devices and services support the recognition that persons with disabilities are persons with abilities who have much to contribute.

ACKNOWLEDGMENTS

Funding for this chapter is due in large part to NIDDRR grant no. H224A940014-01: Colorado Assistive Technology Project, State Grants for Technology Related Assistance. Special thanks are also due to staff of Assistive Technology Partners, Department of Physical Medicine and Rehabilitation, University of Colorado Health Sciences Center, including Maureen Melonis, Jim Sandstrom, Pat McAleese, Marlin Cohrs, Julia Beems, Diane Brians, and Brian Simms. The authors also wish to thank Dr. Thomas W. King, Department of Communication Disorders, University of Wisconsin-Eau Claire; Department of Geriatric Health, American Medical Association; Michelle Lange for her work on the EADL section; and the previous AT chapter authors.

REFERENCES

Major limb amputation impacts multiple aspects of an individual’s life: body image, self-care activities, mobility, psychosocial health, vocational, and avocational opportunities. Successful rehabilitation allows the individual with an amputation to return to their highest level of activity and function. Tempering expectations with reality, balancing the use of prosthetic technology with the individual’s perceived needs, physical and cognitive capabilities, social support network, and financial resources are essential features of rehabilitative care.

Advances in the care and prosthetic restoration of the individual with an amputation have always come from multiple arenas: new surgical techniques, improvements in preoperative and postoperative management, advances in prosthetic technology, and better understanding of the psychosocial implications of limb loss. In the past decade, the greatest advances have taken place in the areas of prosthetic technologies, fabrication of prosthetic sockets, and improved components that more effectively replace the lost function of the extremity. Teams of health care providers in regional centers, who treat a large number of individuals with an amputation, are able to provide optimal prosthetic rehabilitation because of their combined experience (1). While such integrated teams are ideal, a less formal coordination of efforts between surgeons, physiatrists, prosthetists, and therapists in the community can provide effective care to most individuals with an amputation who lack access to specialized centers.

In this chapter, the causes of amputation, basic surgical issues, and the overall approach to the medical and physiatric care of the individual with an amputation are reviewed. A discussion on the prosthetic management of amputation at different levels in both the lower and upper limbs follows to aid the practitioner in organizing the myriad options available for restoring the function of the lost limb. Finally, common problems, medical complications, and special issues in pediatric amputation are discussed.

**INCIDENCE AND ETIOLOGY**

### Acquired Amputation

The etiology of limb loss influences the clinical treatment, management, and functional expectations of the individual with an amputation. Data from the Agency for Healthcare Research and Quality (AHRQ) and the Veterans Health Administration (VHA) from the late 1980s to the late 1990s estimates that a total of 140,000 amputations are performed yearly in the United States (2,3). Acquired amputation accounts for 96% to 99% of all limb losses, with the remaining 1% to 4% being related to congenital causes.

The distributions of upper and lower extremity (LE) amputation by level are shown in Figure 74-1. In the LE, 75% to 93% of acquired amputations are the result of vascular disease (diabetic vascular disease, atherosclerosis, immunologic, and idiopathic). Diabetes alone contributes to two thirds of all LE amputations (2,3). Approximately 6% to 10% of acquired LE amputations result from traumatic injuries with the remainder due to benign or malignant tumors. While accounting for a smaller overall percentage of LE amputation, trauma is the most common cause for LE amputation in the second and third decades of life. Among those between the ages of 10 and 20 years, tumor is the most frequent cause of all amputations (4–6).

In the upper extremity (UE), trauma is the leading cause of limb loss, accounting for 80% of amputations with the vast majority of injuries limited to digital amputations. There are an estimated 10,000 to 15,000 upper limb amputations at the transradial level and above yearly in the United States (Table 74-1) (7). Most of these occur in individuals between the ages of 20 and 40 years. As in the LE, tumor is the most common cause of UE amputation in children.

From 1988 to 1996, the rates of trauma- and cancer-associated amputations declined by more than 40% (2,8). This decline likely reflects improved surgical reconstruction techniques for trauma, limb-sparing management of musculoskeletal tumors, and greater prevention through improved safety awareness. In contrast to the reduction in tumor- and trauma-related limb loss, dysvascular amputation rates have increased 10% to 19% over the past several decades. The increase in dysvascular-related amputation has occurred despite considerable evidence that comprehensive management of the diabetic foot at risk can substantially reduce or delay amputation. Factors that account for this adverse trend include the difficulty of implementing systematic comprehensive management strategies for the diabetic foot coupled with an increasing prevalence of diabetes, hypertension, and hypercholesterolemia placing more individuals at risk (2,9,10).

### Congenital Amputation

The absence of part or all of an extremity at birth is more appropriately referred to as a congenital skeletal deficiency rather than a congenital amputation. The Birth Defects
Monitoring Program, a national program monitoring congenital malformations in the United States, reports the incidence of congenital upper and lower limb skeletal deficiency as 2.41/10,000 births (upper limb, 1.58/10,000 births; lower limb, 0.83/10,000 births) (11). Although a few genetically determined syndromes such as heart-hand malformation (Holt-Oram), renal dysfunction (Fanconi), thrombocytopenia (multiple named syndromes), absent radius and drugs (e.g., thalidomide) have been associated with skeletal deficiencies (12–14), no etiology can be determined for most congenital limb loss.

Multiple systems for classifying congenital limb deficiencies have been proposed but none has been universally accepted. A commonly used and preferred system is based on the International Society for Prosthetics and Orthotics recommendations that classify congenital limb loss as either a transverse or longitudinal skeletal deficiency. Transverse limb deficiency, also referred to as terminal deficiency, is defined as the loss of all skeletal components distal to a particular transverse axis (e.g., transverse forearm or transverse radial limb deficiency). Longitudinal limb deficiency, also referred to as intercalary limb loss, is defined as the loss (complete or partial) of one or more skeletal elements within the longitudinal axis of the limb, with preservation of some or all of the distal skeletal elements. Classification of congenital limb deficiencies is often confused further when congenital deficiencies are described in the same terms used for acquired amputations, such as “transradial” or “transtibial.” Similar terminologies are used either because the congenital deficiency appears similar to an acquired amputation or because a congenital skeletal deficiency has undergone a surgical conversion to accommodate appropriate prosthetic restoration. Surgical conversion has been estimated to be necessary in the management of 50% of LE congenital deficiencies and 8% of UE deficiencies (15).

**Amputation Surgery**

The underlying principle in choosing the amputation level is to preserve as much limb length as possible that will be consistent with the body's functionality and overall quality of life. The selection of the appropriate level of amputation is guided by several factors, including the extent of the injury, the patient's age and physical condition, and the potential for functional restoration. This approach allows for a better fit of the prosthesis, improved mobility, and enhanced quality of life for the individual. Additionally, rehabilitation and training are critical components in facilitating the return to normal or near-normal activities.

### TABLE 74.1

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Lower limb (total)</td>
<td>953,367 (97)</td>
<td>61,605 (31)</td>
<td>8,351 (76.1)</td>
</tr>
<tr>
<td>Toe</td>
<td>309,589 (31.5)</td>
<td>27,233 (13.9)</td>
<td>1,466 (13.4)</td>
</tr>
<tr>
<td>Foot</td>
<td>102,872 (10.5)</td>
<td>4,483 (2.3)</td>
<td>482 (4.4)</td>
</tr>
<tr>
<td>Ankle</td>
<td>7,478 (0.8)</td>
<td>823 (0.4)</td>
<td>164 (1.5)</td>
</tr>
<tr>
<td>Transtibial</td>
<td>271,550 (27.6)</td>
<td>14,244 (7.3)</td>
<td>1,501 (13.7)</td>
</tr>
<tr>
<td>Through-knee</td>
<td>4,237 (0.4)</td>
<td>921 (0.5)</td>
<td>133 (1.2)</td>
</tr>
<tr>
<td>Transfemoral</td>
<td>253,145 (25.8)</td>
<td>10,821 (5.5)</td>
<td>2,499 (22.8)</td>
</tr>
<tr>
<td>Hip disarticulation</td>
<td>3,554 (0.4)</td>
<td>418 (0.2)</td>
<td>726 (6.6)</td>
</tr>
<tr>
<td>Pelvic</td>
<td>469 (0.1)</td>
<td>52 (0.03)</td>
<td>1,369 (12.5)</td>
</tr>
<tr>
<td>Bilateral</td>
<td>0 (0)</td>
<td>1,504 (0.8)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Upper limb (total)</td>
<td>29,426 (3)</td>
<td>134,421 (68.6)</td>
<td>2,617 (23.9)</td>
</tr>
<tr>
<td>Thumb</td>
<td>2,344 (0.2)</td>
<td>24,325 (12.4)</td>
<td>352 (3.2)</td>
</tr>
<tr>
<td>Finger(s)</td>
<td>21,427 (2.2)</td>
<td>100,316 (51.2)</td>
<td>529 (4.8)</td>
</tr>
<tr>
<td>Hand</td>
<td>1,255 (0.1)</td>
<td>983 (0.5)</td>
<td>92 (0.8)</td>
</tr>
<tr>
<td>Wrist</td>
<td>514 (0.1)</td>
<td>415 (0.2)</td>
<td>21 (0.2)</td>
</tr>
<tr>
<td>Transradial</td>
<td>1,626 (0.2)</td>
<td>4,001 (2.0)</td>
<td>212 (1.9)</td>
</tr>
<tr>
<td>Through-elbow</td>
<td>385 (0.04)</td>
<td>346 (0.2)</td>
<td>123 (1.1)</td>
</tr>
<tr>
<td>Transhumeral</td>
<td>1,511 (0.2)</td>
<td>3,008 (1.5)</td>
<td>468 (4.4)</td>
</tr>
<tr>
<td>Shoulder</td>
<td>236 (0.02)</td>
<td>154 (0.1)</td>
<td>365 (3.3)</td>
</tr>
<tr>
<td>Bilateral</td>
<td>0 (0)</td>
<td>462 (0.2)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Forequarter</td>
<td>132 (0.01)</td>
<td>15 (0.01)</td>
<td>439 (4)</td>
</tr>
</tbody>
</table>
with wound healing, an acceptable soft-tissue envelope, and functional prosthetic fitting. While in principle, decision making is straightforward, in practice many issues must be weighed including the severity of the underlying process, tissue viability, overall medical condition of the patient, the morbidity associated with limb salvage, expected functional level after amputation, and ultimately patient–physician preference. In patients with vascular disease, noninvasive vascular studies can assist in predicting wound healing; an absolute ankle Doppler blood pressure of 70 mm Hg or greater, an ankle-brachial index of 0.5 or greater, and transcutaneous oxygen pressure of greater than 20 to 30 mm Hg are all suggestive of a greater likelihood of healing at the transtibial level. Other nonvascular factors associated with compromised wound healing include poor nutritional status (albumin below 3.5 g/dL) or immuno-compromised status (total lymphocyte count <1,500). Despite these identified predictive factors, the final choice of amputation level in the vascular patient often cannot be made until the time of surgery, when the amount of blood flow in the relevant tissues can be observed (8,16–18). In most medical centers, the physiatrist plays a limited role in presurgical planning. In centers where multidisciplinary teams are involved, the physiatrist can offer useful insights into the likelihood of achieving various functional goals at differing levels of amputation. This information allows the health care team and the patient to weigh surgical options when uncertainty exists.

The decision between amputation and limb salvage following trauma is a complex one. Early or immediate amputation may be required in life-threatening multiple trauma situations when the physiologic demands associated with repeated surgery to salvage a limb would not be tolerated. The presence of preexisting medical conditions such as peripheral vascular disease (PVD) or neurologic injury that adversely affect function of the residual limb may also favor immediate amputation. However, when the injuries are not immediately life-threatening, the decision to amputate versus attempting limb salvage must be based on an assessment of the approach that will most effectively restore function and return the individual to his or her preinjury activities. Extensive soft-tissue loss, proximal arterial injuries, multiple arterial injuries, and sciatic or tibial nerve damage are poor prognostic indicators for successful limb salvage (17,19). In general, limb salvage of the LE requires more surgeries, involves longer hospital stays, delays weight bearing, and slows the return to preinjury activities compared to amputation (19). Because the functional demands of the UE are different from the LE, a bias towards limb salvage has developed in UE trauma surgery. The lack of weight-bearing forces, the ability to function with partial sensation, and the limited functionality of UE prostheses are reasons cited for greater effort directed toward limb salvage and reimplantation (20,21).

Amputation surgery must be approached as a reconstructive procedure. Bones are beveled to minimize the sharp edges that can cause tissue trauma and pain with weight bearing. The nerves are sharply transected and allowed to retract into proximal soft tissues so that they do not become adherent in scar or remain in a location subjected to high loading forces from a prosthesis. Appropriate myofascial closure of the muscle or myodesis provides for good control of the remaining bone in the residual limb, and appropriate placement of the skin incision line avoids bony prominences and adherence to underlying bone. Such attention to detail will result in a well-shaped residual limb that can be effectively fitted with a prosthesis. Ilizarov techniques or free fibular grafts (22–24) have been used to lengthen short residual limbs, though the indications for their use and their success are highly individualized. Skin grafts and myocutaneous free flaps (25,26) have been used successfully to preserve length in nonvascular individuals with an amputation, but in general the presence of skin grafts or insensate skin in the residual limb often results in recurring skin breakdown. When recurring skin breakdown occurs, stump revision, use of tissue expanders, and creative approaches to prosthetic design may be needed. The preferred levels of amputation are as follows.

**Lower Extremity**
- Toe amputations
- Ray resections
- Transmetatarsal amputations
- Syme amputation (i.e., ankle disarticulation)
- Transtibial amputation (between the junction of the middle and distal thirds of the leg)
- Knee disarticulation (KD)
- Transfemoral amputation (8 cm or more proximal to the level of the knee joint)
- Hip disarticulation (short transfemoral amputation at/or proximal to the greater trochanter is functionally a hip disarticulation)
- Hemipelvectomy

**Upper Extremity**
- Digital amputation
- Ray resection
- Transmetacarpal resection
- Wrist disarticulation (WD)
- Transradial amputation
- Elbow disarticulation
- Transhumeral amputation (THA) (i.e., 6.5 cm or more proximal to the elbow joint)
- Shoulder disarticulation
- Forequarter amputation (interscapulothoracic disarticulation)

**LE AMPUTATION**

**General Principles of LE Amputation Management**

The interaction between the health care team and the patient to achieve the goal of prosthetic restoration and rehabilitation is referred to as prosthetic management (27). The process of prosthetic rehabilitation can be organized into a four-phase
process: preprosthetic management, postoperative care, prosthetic fitting and training, and long-term follow-up care. This staging permits the rehabilitation physician to assess the individual with an amputation and organize the rehabilitation program.

**Preprosthetic Patient Evaluation and Management**

Preprosthetic management begins when the decision to perform an amputation is made, when a patient is initially evaluated after a traumatic amputation, or when a child is born with a congenital skeletal deficiency. It ends with the fitting of a prosthesis. Optimal care is ensured when members of the prosthetic team can evaluate the patient before amputation, but often the events surrounding an amputation delay the rehabilitation assessment until the postoperative period.

The preprosthetic evaluation, whether performed preoperatively or postoperatively, should focus on identifying factors that will affect the ultimate functional status of the patient and optimize prosthetic fitting. Issues that need evaluation include assessing the premorbid functional status, identifying coexisting musculoskeletal, neurologic, and cardiopulmonary disease that will influence the rehabilitation potential, determining the available social support network, and understanding the patient’s goals and expectations. Education of the patient and family about the functional consequences of amputation and the steps involved in prosthetic rehabilitation will help allay some of the fears the patient may have about his or her future. Therapy programs for range of motion, conditioning exercises, correct positioning of the residual limb, ambulation with gait aids, relaxation techniques, and activities of daily living (ADLs) should be started as soon as medically appropriate. The patient is often better able to absorb and comply with a therapy program during the preoperative period than during the early postoperative period, when incisional pain, medication, or apprehension may interfere with the ability to participate.

**Postoperative Care**

The goals that direct the postoperative, preprosthetic management of the individual with an amputation are outlined in Table 74-2. During the immediate postoperative period, general medical care focuses on optimizing control of underlying disorders that can interfere with rehabilitation: diabetes, coronary artery disease, congestive heart failure, renal disease. Maintaining nutritional status is frequently neglected, nevertheless it plays a critical role in ensuring wound healing (8) and in facilitating the muscular strength adaptations needed for prosthetic mobility. The principles guiding residual limb care are based on ensuring primary wound healing, controlling pain, minimizing edema, and preventing contractures.

Options for wound management include soft dressings, semi-rigid dressings (Unna casts), rigid dressings (plaster or fiberglass casts), and air splints. Each option has advantages and disadvantages. Soft dressings are typically used with an elastic bandage wrap (i.e., Ace bandage) or a compressive stockinette. Soft dressings have the advantage of being readily available, quickly applied, and allowing frequent wound inspection. However, they do not provide protection from external trauma and only have a limited ability to control edema. If poorly applied, elastic wraps can lead to tourniquet effect. Elastic bandages require considerable cooperation, skill, and attention on the part of the patient, family, and medical staff because the wraps need to be reapplied frequently and carefully to be successful. In practice their use is problematic enough and alternatives such as compressive stockinettes, elastic stump shrinkers, or roll-on gel liners are often a better choice of edema management. Despite a number of limitations, soft dressings remain the most commonly used wound care approach following amputation (28).

Rigid dressings have been reported to reduce wound-healing time and lead to more rapid and improved rehabilitation (29,30). The primary concerns surrounding rigid dressings are the inability to inspect the wound and the potential increase in wound breakdown from incorrect application or early weight bearing in dysvascular individuals with an amputation. In spite of these limitations, postoperative rigid dressings may be the preferred method of wound care, especially for the transtibial individual with an amputation, but their clinical use and acceptance has been limited by the lack of expertise in their application. Rigid dressings can be fabricated in a removable form that resembles a transtibial prosthetic socket or as a nonremovable cast that extends to the mid-thigh level. The rigid dressing is most commonly made using standard orthopedic cast materials but commercially available prefabricated devices are also available. A nonremovable rigid dressing is typically applied during or shortly after surgery and replaced every 7 to 14 days. The mid-thigh length of the dressing prevents knee flexion contractures and is continued until adequate wound healing has occurred so that concerns over contracture development are minimized. Subsequent dressings are fabricated as removable rigid dressings (RRDs) that can be taken off whenever the wound needs to be inspected. Rigid dressings have been predominantly used in the individuals with a traumatic amputation because of lessened concern over wound healing or limb injury from the dressing. While rigid dressings can be used simply as a wound care strategy, they can also be used with a pylon attachment to which components

---

**TABLE 74.2 Goals of Postoperative Management of the Individual with an Amputation**

- Successful healing of the amputation
- Pain control
- Maintaining range of motion in the remaining proximal joints of the amputated extremity
- Strengthening of residual muscle groups needed for biomechanical compensation
- Preparation of the residual limb for prosthetic fitting
- Achieving independence in ADLs and mobility without a prosthetic limb
- Education about the process of prosthetic limb fitting and expected functional outcome
- Psychosocial support for the adaptations resulting from the amputation
can be attached, creating a preparatory prosthesis that enables immediate or early weight bearing.

Because little objective data exist that clearly identify a superior wound dressing strategy, the choice of wound management appears to be driven largely by practice conventions, availability of skilled staff, and the personal experience of the surgeon. Greater attention is needed to facilitate rapid wound healing, especially in the individual with a dysvascular amputation in whom the effects of prolonged immobilization may substantially complicate rehabilitation effects. The preprosthetic phase of management, before preparatory prosthetic fitting, can typically last 6 to 10 weeks for the individual with a dysvascular LE amputation, a considerably shorter period of time for the individual with a traumatic LE amputation, and 3 to 6 weeks for the individual with a UE amputation (12).

Muscle imbalance and postoperative positioning to facilitate comfort leads to the development of knee flexion contracture in the transtibial residual limb and to hip flexion and abduction contractures in the individual with a transfemoral amputation. Contractures are preventable through a postoperative therapy program that emphasizes range of motion exercises and early mobilization. Strengthening of muscle groups that biomechanically substitute for the lost function of the limb is needed. Exercise programs are required to accomplish this task. In the individual with a LE amputation, the hip extensors (gluteus maximus and hamstrings), gluteus medius, hip flexors, and the contralateral ankle plantar flexors all contribute to restoring ambulation ability (31,32). In the individual with a UE amputation, proximal shoulder girdle muscle strengthening should be taught, emphasizing the trapezius, serratus anterior, pectoralis major, as well as any residual deltoid and biceps functions.

An individual’s psychological response to amputation may be compared to the grieving process that varies greatly and includes identifiable stages of denial, anger, depression, coping, and acceptance. Not every person ultimately adapts to limb loss. The individual’s ultimate response to the psychosocial impact of limb loss is determined by many factors, including the cause of amputation, personal life experience and inner strengths, the available social support system, the care provided by the prosthetic team, and the functional outcome that is achieved through rehabilitation.

Prosthetic Fitting and Training
An understanding of the functional needs of the individual with an amputation, his/her interest and motivation in pursuing prosthetic fitting, and an assessment of his/her ambulatory potential are required to set realistic goals for prosthetic fitting and training. Not all the individuals with an amputation are candidates for prostheses. Although the factors that predict success in prosthetic use are partially understood, a number of factors have been associated with a poor outcome in returning the individual with an amputation to functional ambulation at household or community levels. Negative prognostic factors include a delay in wound healing, the presence of joint contractures, dementia or cognitive disorders, medical comorbidities, and higher levels of limb amputation (transfemoral) (33–35). Age has inconsistently been identified as a predictor of prosthetic success, implying that except in advanced age (>80 to 85 years) other factors play a more important role in determining the rehabilitation potential of the individual with an amputation.

As a result of the uncertainty in identifying prosthetic candidates, considerable clinical judgment is required. Some general guidelines can be followed. An individual with an amputation should have reasonable cardiovascular reserve, adequate wound healing, and good soft-tissue coverage, range of motion, muscle strength, motor control, and learning ability to achieve useful prosthetic function. Individuals with an LE amputation who can walk with a walker or crutches without a prosthesis usually possess the necessary balance, strength, and cardiovascular reserve to walk with a prosthesis. Examples of poor candidates for functional prosthetic fitting would be an individual with a dysvascular LE amputation with an open or poorly healed incision, an individual with a transfemoral amputation with a 30-degree flexion contracture at the hip, or an individual with a transradial amputation with a flail elbow and shoulder. Generally, individuals with a bilateral, short, transfemoral amputation over the age of 45 years are considered unlikely candidates for full-length prosthetic fitting. Additional medical problems such as severe coronary artery disease, pulmonary disease, severe polyneuropathy, or multiple-joint arthritis may result in an individual with an amputation who could be fitted with a prosthesis but who may not be a functional prosthetic user. Patients in whom prognosis is poor, life expectancy is short, or with a disease that results in significant fluctuations in body weight are not good candidates. In borderline cases, it may be necessary to proceed with actual prosthetic fitting to determine the eventual prosthetic function. The use of a less costly, RRD with pylon and foot or a preparatory prosthesis is appropriate before a decision is made about fitting such a person with a more costly definitive prosthesis. The overall success rate in restoring functional ambulation in the individual with a lower limb amputation varies approximately from 36% to 70%. Amputation resulting from vascular disease is a manifestation of a severe systemic vasculopathy. The early mortality following major LE amputation is 15% to 20%, largely related to myocardial infarction. Overall, the individuals with dysvascular amputation have a 3- to 5-year 50% mortality, which underlies the importance of successful early rehabilitation to allow for an improved quality of life in their remaining years.

The timing of prosthetic fitting for the individual with an LE amputation remains controversial, reflecting the clinical uncertainty over early versus delayed weight bearing. Because the majority of LE amputations occur as a result of PVD, primary wound healing at the amputation site is of paramount importance. When the rigid dressing was introduced on a wide scale in the 1970s, it was used to implement immediate postoperative prosthesis (IPOP) (a rigid dressing with a pylon and foot) as a means to speed rehabilitation for individuals with LE amputation (36). Problems with wound healing and
residual limb trauma from poorly fabricated devices and a lack of experienced teams to manage this approach to early postoperative care led to abandoning their use in the individual with a dysvascular amputation. Despite these problems, in selected centers with adequate experience and a process to monitor closely the residual limb, an immediate or early postoperative prosthesis fabricated several weeks after surgery has been used safely in individuals with a dysvascular amputation (37). Immediate fitting in the younger patient with traumatic amputation has been more successful and is a reasonable method of treatment. Immediate and early postoperative prostheses are, in effect, an RRD with a pylon and foot attached. This device is used to achieve limited partial to full weight bearing, reduce edema, and accomplish initial gait training. Because the fit of these devices is always suboptimal compared to a custom-molded socket, they are not recommended for extended use.

When concern over wound healing dominates clinical care in the postoperative period, prosthetic fitting is delayed until the residual limb has healed adequately to allow unrestricted weight bearing. Providing a prosthesis is typically performed in two stages: a preparatory prosthetic limb phase is followed by the provision of a definitive prosthesis. The preparatory prosthesis is often of simple design, lower performance, and is more accommodating to changes in residual limb volume than is the definitive limb. It allows the individual with an amputation to gain skill and confidence in walking with prostheses, facilitates residual limb maturation, and affords the rehabilitation team the opportunity to better define the ultimate functional level of the individual. When stump maturation has occurred, a definitive prosthesis is prescribed to meet the anticipated needs of the individual with an amputation.

Stump maturation is an imprecisely defined concept that occurs when the volume of the residual limb has stabilized, soft-tissue atrophy has occurred, and the residual limb has been molded into a cylindrical shape that optimizes prosthetic fitting. This can usually be determined when the individual with an amputation reports a plateau in the number of sock plies worn from day to day and by clinical exam that shows edema resolution. Residual limb maturation, typically, takes about 4 months (38) but may extend substantially longer depending on the activity level, amount of prosthetic limb use, and coexisting medical disease. After stump maturation occurs, a definitive prosthesis is prescribed to specifically meet the ADLs and vocational and avocational needs of the individual with an amputation. In the case of young children, the prosthesis prescription must also meet any needs related to the development of age-appropriate motor milestones. Although a two-stage approach (preparatory followed by definitive limb) is commonly used, financial considerations are becoming increasingly important with many health insurance programs allowing for only a single limb. Under these situations, the prosthetic team may recommend as the initial prostheses a limb that is projected to meet all the long-term needs of the individual with an amputation. Patients who are not candidates for functional prosthetic use may choose to have a cosmetic prosthesis that has an appearance similar to that of the opposite limb.

**Gait Training**

After completing the final prosthetic evaluation, the individual with a new amputation will require a period of gait training under the supervision of the physical therapist. The individual with an amputation is instructed on how to put on and take off the prosthesis, how to determine the appropriate number of limb socks to be worn, when and how to check the skin for evidence of irritation, and how to clean and care for the prosthesis. For the individual with a new amputation it is best if the initial gait training occurs while the prosthesis is still capable of being adjusted to permit alignment or length changes that may become apparent during gait training. Gait training often occurs on an outpatient basis and may last from weeks to months. The more proximal levels of amputation require lengthier gait training.

Gait training begins with weight shifting and balance activities while still in the parallel bars. Once weight shifting and balance activities have been mastered, a program of progressive ambulation begins in the parallel bars and progresses to the most independent level of ambulation possible with or without gait aids. Specific training should focus on transfers, knee stability, equal step lengths, and avoiding lateral trunk bending. Following mastery of ambulation on flat, level surfaces, techniques for managing uneven terrain, stairs, ramps, curbs, and falling and getting up off the ground are learned. Moving from a walker to less cumbersome gait aids can be achieved for most individuals with an LE amputation. For higher functioning individuals with an amputation, prosthetic training should include instruction and practice in driving, recreation, and vocational pursuits. Developing the optimal benefit from a prosthesis must take into account the specific mechanical attributes of the components used. For example, using a dynamic response (i.e., energy-storing) prosthetic foot requires loading the prosthetic toe during mid-stance and late stance to capture energy for push off assistance or to activate a prosthetic knee to initiate the swing phase.

Wearing tolerance for the prosthesis must be gradually increased. Initially, the individual with an amputation will wear the prosthesis only for 15 to 20 minutes, removing it to check the condition of the skin. As tolerance to weight bearing increases, the length of wearing time is gradually increased. Several weeks may be required before the individual with an amputation is able to wear the prosthesis full-time. The individual with an amputation may take the prosthesis home when safe and independent ambulation has been demonstrated and residual limb skin checks are assured. Common gait deviations and their causes are highlighted in Table 74-3.

**LE Prosthetic Follow-up**

During the initial 6 to 18 months, most individuals with an amputation will experience continued loss of residual limb volume, resulting in a prosthetic socket that will be too large. During this period, return visits should occur frequently enough to ensure that this loss of residual limb volume is being compensated for by the use of additional limb socks or by appropriate modifications of the prosthetic socket. It is usual for an individual with a new amputation to require replacement of the
<table>
<thead>
<tr>
<th>Gait Cycle</th>
<th>Observed Gait Abnormality</th>
<th>Possible Cause</th>
<th>Suggested Modifications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial contact to loading response</td>
<td>Abrupt heel contact, rapid knee flexion</td>
<td>Excessive heel lever&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Realign prosthetic foot, change heel stiffness</td>
</tr>
<tr>
<td></td>
<td>Prolonged heel contact, knee remains fully extended</td>
<td>Inadequate heel lever&lt;sup&gt;a&lt;/sup&gt; or heel worn out, improper socket flexion, learned gait pattern, quadriceps weakness</td>
<td>Increase heel stiffness, realign prosthesis, gait training and strengthening</td>
</tr>
<tr>
<td></td>
<td>Jerky knee motion</td>
<td>Socket loose, poor alignment, inadequate suspension</td>
<td></td>
</tr>
<tr>
<td>Mid-stance</td>
<td>Medial or lateral socket thrust, lateral trunk shift over prosthesis</td>
<td>Foot too far outset or inset, socket loose</td>
<td>Realign prosthesis, replace socket or adjust socks</td>
</tr>
<tr>
<td>Mid-stance to terminal stance</td>
<td>Pelvis drops or elevates Early knee flexion or “drop off”</td>
<td>Prosthesis too short/too long, inadequate toe lever&lt;sup&gt;c&lt;/sup&gt;</td>
<td>Adjust prosthetic length, realign prosthesis, replace foot</td>
</tr>
<tr>
<td>Terminal stance</td>
<td>Heel off too early</td>
<td>Excessive toe lever&lt;sup&gt;c&lt;/sup&gt;, too much socket extension</td>
<td>Realign prosthesis</td>
</tr>
<tr>
<td></td>
<td>Heel off excessively delayed</td>
<td>Inadequate toe lever&lt;sup&gt;c&lt;/sup&gt;, too much socket flexion</td>
<td>Realign prosthesis</td>
</tr>
<tr>
<td>Swing phase</td>
<td>Prosthetic foot drags</td>
<td>Prosthesis too long, inadequate suspension</td>
<td>Shorten limb, modify suspension</td>
</tr>
<tr>
<td>Successive double support</td>
<td>Uneven step length</td>
<td>Hip flexion contracture, gait insecurity Uncomfortable socket</td>
<td>Physical therapy, adjust socket fit</td>
</tr>
</tbody>
</table>

**Transtibial Amputee Gait**

<table>
<thead>
<tr>
<th>Gait Cycle</th>
<th>Observed Gait Abnormality</th>
<th>Possible Cause</th>
<th>Suggested Modifications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial contact to loading response</td>
<td>Foot rotation at heel strike</td>
<td>Poor socket fit/rotation</td>
<td>Adjust socket fit, add belt for rotation control, reduce heel stiffness</td>
</tr>
<tr>
<td></td>
<td>Knee buckling</td>
<td>Heel too firm Excessive heel lever&lt;sup&gt;a&lt;/sup&gt; Incorrect prosthetic knee alignment Weak hip extensors</td>
<td>Realign limb, reduce heel stiffness, realign TKA relationship, gait training and strengthening</td>
</tr>
<tr>
<td>Mid-stance</td>
<td>Lateral trunk bend or shift over prosthesis</td>
<td>Prosthetic limb abducted Too much socket abduction, foot too far outset Prosthesis too long Medial groin pain Poor medial-lateral prosthetic control Poor socket fit Weak hip adductors Short residual limb Prosthesis too short</td>
<td>Realign prosthesis</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Shorten prosthesis</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Adjust socket fit</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Adjust prosthesis length, gait training and strengthening, accept, possibly add hip joint, adjust knee friction or damping</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Adjust knee friction or damping</td>
</tr>
<tr>
<td>Initial swing</td>
<td>Uneven heel rise</td>
<td>Knee friction too tight or loose Knee extension</td>
<td>Realign prosthesis</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Prosthetic foot drags</td>
<td></td>
</tr>
<tr>
<td>Swing phase</td>
<td>Circumduction or prosthetic limb</td>
<td>Inadequate knee flexion, knee too stiff Prosthesis too long, inadequate suspension Poor gait pattern Improper knee rotational alignment Excessive socket rotation</td>
<td>Adjust knee flexion or damping, adjust prosthesis length, physical therapy, realign prosthesis, adjust socket fit</td>
</tr>
<tr>
<td></td>
<td>Whips</td>
<td>Hip flexion contracture Insufficient socket flexion</td>
<td>Physical therapy, realign prosthesis</td>
</tr>
</tbody>
</table>

<sup>a</sup>Causes of excessive heel lever—foot dorsiflexed too much, foot too far posterior, heel cushion too hard, shoe heel too hard.  
<sup>b</sup>Causes of inadequate heel lever—foot plantarflexed too much, foot too far anterior, heel cushion too soft.  
<sup>c</sup>Causes of inadequate toe lever—foot dorsiflexed too much, foot too far posterior, foot keel too soft/ flexible.  
<sup>d</sup>Causes of excessive toe lever—foot plantarflexed too much, foot too far anterior, foot keel too stiff.
prosthetic socket during this time because of the significant loss of soft-tissue volume. During follow-up visits, the condition of the residual limb, the prosthesis, the individual’s gait, and the level of function are reviewed (39). Appropriate medical treatment, prosthetic modifications, or additional therapies are prescribed as needed. When the residual limb volume has stabilized sufficiently and the patient is doing well with the prosthesis, yearly visits to the amputee clinic are appropriate. Once the residual limb has stabilized, the average life expectancy for an LE prosthesis before replacement should be 3 to 5 years.

**LE Prostheses**

The LE prosthetic prescription must balance the individual’s need for stability, mobility, durability, and cosmesis with available resources and cost. Understanding the role and importance of prosthetic ambulation in achieving the mobility goals of the individual with an amputation is essential for correctly prescribing a prosthetic device. Prosthetic ambulation is usually the primary mode of mobility for the younger individual with an amputation as well as for other patients across a wider age range when the amputation is at the transtibial and more distal levels. For the elderly, with dysvascular amputation above the knee and a more proximal level of amputation, prosthetic ambulation is often limited to transfers, indoors, or short community distances. The prescription of the LE prosthesis is based on several principles: maximizing comfort, matching specific components to the mobility needs of the individual with an amputation, and providing acceptable cosmesis. Comfort is the most critical aspect and depends on achieving an appropriate distribution of forces between the residual limb and the socket. A poorly fitting or uncomfortable socket will limit the mobility and often lead to rejection of the prosthesis. Once comfort has been established, the appropriate choice of components facilitates achieving maximal independence and function during sitting, standing, transferring, walking, and running.

Lastly, cosmetic concerns are considered. Cosmesis is influenced by personal preferences and psychosocial dynamics but is usually satisfactorily achieved using contoured foam and a nylon or rubber skin tone cover. Some individuals with an amputation prefer not to have their prosthesis covered because of the possible interference with prosthetic component function.

Medicare, a major funding source for prosthetic limbs in the United States, requires that the functional level of the individual with an amputation be taken into account when prescribing a prosthesis. The functional index is referred to as the Medicare “K” code and limits the components that can be used when fabricating the prosthesis. Although only required for Medicare, the “K” code classification is a simple but useful hierarchical framework for classifying the mobility potential of all individuals with an LE amputation (Table 74-4).

**LE Prosthetic Components**

The continual introduction of new component designs and the overlap of functional features of components from various manufacturers make it difficult to stay abreast of available prosthetic options. Collaboration between health care providers (physician, prosthetist, and therapist) is essential in developing an appropriate, individualized limb prescription. Seldom is there a single correct choice of components for a prosthesis, rather most individuals with an amputation can be successfully fit using components that span a reasonable range of mechanical and functional characteristics. Because objective data, linking prosthetic component characteristics to the demographics of individuals with an amputation, are limited, empiric approaches and experience play a major role in limb prescription. The prescription for an LE prosthesis should include the Medicare “K” code, diagnosis, type of prosthesis (with modifiers), socket type, liner, suspension method, foot, knee and hip systems (as required by amputation level), diagnostic or check socket, and supplies.

### Table 74.4: Medicare Guidelines for Functional Classification of Patients with Prosthesis

<table>
<thead>
<tr>
<th>K Code Level</th>
<th>Functional Level</th>
<th>Activity Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>K0</td>
<td>Not a potential user for ambulation or transfer</td>
<td>Does not have the ability or potential to ambulate or transfer safely with or without assistance, and a prosthesis does not enhance their quality of life or mobility.</td>
</tr>
<tr>
<td>K1</td>
<td>A potential household ambulator including transfers</td>
<td>Has the ability or potential to use a prosthesis for transfer or ambulation on level surfaces at fixed cadence. Typical of the limited and unlimited household ambulator.</td>
</tr>
<tr>
<td>K2</td>
<td>A potential limited community ambulatory</td>
<td>Has the ability or potential for ambulation with the ability to traverse low level environmental barriers such as curbs, stairs or uneven surfaces. Typical of the limited community ambulator</td>
</tr>
<tr>
<td>K3</td>
<td>Community ambulator using variable cadence, including therapeutic exercise or vocation</td>
<td>Has the ability or potential for ambulation with variable cadence. Typical of the community ambulator that has the ability to traverse most environmental barriers and may have vocation, therapeutic, or exercise activity that demands prosthetic utilization beyond simple locomotion.</td>
</tr>
<tr>
<td>K4</td>
<td>High activity user which exceeds normal ambulation skills</td>
<td>Has the ability or potential for ambulation that exceeds basic ambulation skills, exhibiting high impact, stress, or energy levels. Typical of the prosthetic demands of the child, active adult, or athlete.</td>
</tr>
</tbody>
</table>

Prosthetic Feet

Prostheses for amputations at or proximal to the ankle require the use of a prosthetic foot. The selection of an appropriate prosthetic foot is complicated by the wide range of foot designs, marketing-driven claims of performance, and the limited availability of objective data comparing the relative biomechanical and functional advantages of different feet. In the clinical setting, the selection of a prosthetic foot is largely empirically based on the conceptual goal of matching the functional characteristics of the foot to the expected activity needs of the individual (40–45). Within this approach, it is useful to group feet by their major functional feature(s) as belonging to rigid keel, flexible keel, single/multiaxial, or dynamic response (or energy-storing) categories. It is acceptable for the prescribing physician to define the functional features desired in the foot and to rely on the prosthetist who typically has a better working understanding of the commercially available feet to select the specific manufacturer and foot within the desired functional class. This multidisciplinary approach is increasingly important as foot designs become more sophisticated, more costly, and combine different functional characteristics into a single foot. Occasionally, another characteristic of a foot such as an adjustable heel height, cosmesis, or being waterproof is the primary determinate in its selection.

The solid ankle cushion heel (SACH) foot (Fig. 74-2) is the least expensive and most commonly prescribed prosthetic foot. It is durable and lightweight, which accounts in part for its usefulness. The SACH foot has no moving parts and consists of a wooden or composite keel with a compressible foam heel and toes that flex under load, allowing limited simulation of the effects of the heel and forefoot rocker mechanisms of the normal foot. A SACH foot is appropriate for individuals with an amputation who have a lower activity level (K1 to K2), with ambulation primarily limited to level surfaces. It can be used in a wide range of individuals with an amputation for the preparatory prosthesis and upgraded as the individual with an amputation progresses to a higher activity level. For a juvenile with an amputation, the SACH foot is often the most cost-effective foot due to the need for frequent foot changes because of rapid growth.

The flexible keel foot (Fig. 74-2) is designed to mimic the motion of the forefoot rocker mechanism by replacing the rigid keel of the SACH foot with a flexible keel. The keel bends with controlled stiffness as the foot moves from midstance through preswing. Several versions of flexible keel feet are commercially available, each with different construction but sharing similar function. The stationary-ankle-flexible endoskeletal (SAFE) II foot is a commonly used flexible keel foot. The flexible keel foot allows some inversion and eversion, and gives a smoother rollover than a SACH foot, making it appropriate for general mobility needs in the individual with an amputation with a low to moderate activity level. However, the more active individual with an amputation may perceive the flexible keel foot as being too soft, especially for fast walking or running activities.

Articulating prosthetic feet include both single axis and multiaxial designs. The single axis foot allows controlled movement in the sagittal plane (plantar-flexion and dorsiflexion), adjusted by using different durometer bumpers. The primary advantage of the single axis foot is its ability to reduce knee-bending movements during limb loading, thus improving knee stability. Disadvantages include a greater weight than many other feet and more maintenance to ensure correct function. This foot is primarily used in the individual with a proximal amputation that requires better knee stabilization, such as the elderly individual with a transfemoral amputation or the individual with a transfemoral amputation and a short residual limb.

Multiaxial foot designs allow for varying degrees of controlled movement in the sagittal, coronal, and transverse planes (plantar/dorsiflexion, inversion/eversion, some degree of transverse rotation). Multiaxial feet can use mechanical joints to supply motion such as the Greissinger foot or the College Park foot (see Fig. 74-2), but increasingly rely on the inherent flexibility of rubber and polymer materials to provide multiaxial motion. Using material flexibility improves durability and reduces both weight and maintenance compared to mechanical jointed feet. Multiaxial “ankle” motion can be integrated into the foot (e.g., Endolite foot, Luxon) or added to other feet through the use of separate multiaxial ankle components (e.g., impulse ankle—Ohio Willow Wood, Mt. Sterling, Ohio). Multiaxial capabilities are appropriate for the individual with an amputation who needs improved ankle motion to accommodate to uneven terrain and for the active individual with an amputation who requires greater ankle movement to adjust to different speeds or for cutting and pivoting quickly.
Dynamic response (i.e., energy-storing) prosthetic feet incorporate elastic (spring-like) elements that store energy in the foot during limb loading and mid-stance as the elastic material compresses or flexes. Energy is returned at the time of push-off as the spring components of the foot returns to its normal shape or configuration. Examples include the Flex-foot, the Springlite feet, Seattle foot, and the impulse feet (see Fig. 74-2). The dynamic energy characteristics of these feet make them particularly suitable for individuals with an amputation involved in activities requiring running and jumping. Many individuals with an amputation believe that they are more functional with a dynamic response foot. Dynamic elastic response (DER) feet were expected to make ambulation more efficient by reducing the oxygen consumption of individuals with an amputation but the results of objective studies have been mixed (44,46). The metabolic benefits of DER designs are limited and primarily seen at faster walking speeds.

**Prosthesis by Level of Amputation**

**Partial Foot Amputation**

Toe amputations, ray resections, and transmetatarsal amputations require minimal prosthetic/orthotic intervention. At the more distal foot amputation levels and for the less active individual with a transmetatarsal amputation, accommodative shoes with custom insoles, arch supports, and toe fillers are usually adequate. More active individuals with a transmetatarsal amputation may benefit from orthotic modifications that better substitute for the lost anterior foot lever arm. Options include the addition of carbon fiber or spring steel sole shanks, rocker soles, or short ankle foot orthosis. Partial foot amputations at the tarsal-metatarsal and transtarsal levels (e.g., Lisfranc, Chopart) are relatively uncommon and have historically been associated with equinovarus contracture of the hind foot, increasing the likelihood of skin breakdown over the plantar surface of the foot. However, improved surgical techniques that include Achilles tendon lengthening/resection and anterior tibialis and peroneus tendon transfers have reduced equinovarus deformities and result in a functional and useful amputation level (47,48). Prosthetic/orthotic devices for the individual with a proximal partial foot amputation need to supply medial-lateral stabilization of the hind foot and substitute for the lost forefoot lever. Options include: (a) an extra-depth shoe with toe filler, steel shank, and rocker bottom modifications; (b) custom posterior leaf-spring ankle-foot orthosis with toe filler; or (c) a custom prosthetic foot with a self-suspending rear-opening split socket (47,48). A major advantage of all partial foot amputations is the ability to be fully end-bearing, allowing ambulation without any devices.

**Syme Amputation**

Similar to the hind foot amputation, the Syme (tibiotalar disarticulation) amputation is capable of full weight end bearing. The heel flap is anchored to the distal end of the tibia and fibula, and following healing, allows short distance ambulation without a prosthesis. The substantial leg length discrepancy makes long distance ambulation impractical. Over time, posterior migration of the distal heel pad occurs in some individuals with a Syme amputation leading to problems with skin breakdown and difficulty in prosthetic fitting (49–51). The relatively bulbous distal end of the residual limb has the advantage of enabling the use of self-suspending prosthetic designs; however, it also contributes to the major disadvantage of the Syme amputation—poor cosmesis due to the bulkiness of the prosthesis around the ankle joint. There are several different types of prostheses available for individuals with a Syme amputation.

The most common prosthetic style uses a total contact socket with a removable medial window (Fig. 74-3). The distal removable window allows the bulbous portion of the residual to slip easily into the socket, which is then held in place by closing and securing the window with Velcro straps. The major disadvantages to this prosthesis style are the poor cosmesis and the reduced strength of the socket due to the window. A second option uses a fixed posterior opening socket. This type of prosthesis is used for a very bulbous residual limb. This prosthesis is prone to breakage at the ankle joint and is not recommended for heavy-duty users.

Alternative designs to window sockets use a flexible socket wall or liner to allow donning the prosthesis around the distal bulbous end of the residual limb. In the “stovepipe” design, a pelite or similar liner is used that is built proximal to the ankle area, creating a cylindrical stovepipe-shaped inner liner. To don the prosthesis, the patient slips the distal bulbous residual past the narrow center of the liner that can then be easily inserted into the socket. This style of prosthesis is somewhat bulky but can be easily modified and is durable, making it suitable for use as a preparatory prosthesis or in active or obese patients.

The expandable wall prosthesis uses a double-wall socket that has a flexible, expandable inside liner and a rigid outer frame. The inner wall is made of silicone or other elastomers that are flexible enough to allow the bulbous residual limb to slide into the prosthesis. This prosthesis is typically difficult to modify but has the advantage of being very strong and useful for active users or obese patients, and is generally easy to don and doff, making it suitable for individuals with upper limb impairment or cognitive impairment.
determined. Next, the suspension system is chosen and finally pylon and foot/ankle components are selected.

The patella tendon-bearing (PTB) total contact socket has been the internationally accepted standard transtibial socket since the 1960s (52). The PTB total contact socket is fabricated from a cast of the residual limb which is modified to specific weight bearing (SWB) regions that are pressure tolerant and correspondingly modified to decrease pressure over bony prominences such as the tibia crest, fibula, and distal portion of the tibia (Fig. 74-4).

The standard PTB design has several variations (53–55). The PTB-Supracondylar (PTBSC) socket has high medial and lateral sidewalls that extend above and over the femoral condyles, providing enhanced mediolateral stability and self-suspension of the prosthesis. The PTB-Supracondylar/suprapatellar (PTBSCSP) socket further extends the PTBSC socket concept by also extending the anterior aspect of the socket so that the patellar is encompassed within the socket. The PTBSCSP gives additional stiffness to the mediolateral walls and applies force proximal to the patella during stance to provide sensory feedback to limit genu recurvatum. Both the PTBSC and PTBSCSP are primarily used in individuals with an amputation and with short residual limbs to improve varus/valgus control and to provide greater surface area for weight distribution (Fig. 74-5).

An alternative socket design for individuals with a transtibial amputation is the total surface-bearing (TSB) socket made practical by the development of gel and elastomeric liner systems (discussed below). The TSB socket is made from a cast of the residual limb with minimal modifications. When used with gel liners, the TSB socket is believed to distribute pressure more uniformly within the socket. The relative advantages and disadvantages of TSB versus PTB total contact sockets remain poorly understood. When a comfortable fit cannot be achieved with one type of socket, empirically switching to the other can be successful.

Low-profile feet are needed for Syme prosthesis due to the limited space available beneath the socket. Acceptable foot options range from the rigid keel SACH feet through multiaxial and dynamic response feet.

**Transtibial Amputation**

Transtibial amputations are the most common amputation level seen in general practice. Considerable effort over the past several decades has gone into designing components to address the needs of individuals with transtibial amputations. The large number of options available increases the uncertainty when choosing components but if approached systematically, straightforward reasoned decisions can be used to generate a prosthetic prescription. The following discussion of transtibial prosthetic components parallels the recommended approach for prescribing prostheses. Initially, the socket and liner system that is anticipated to optimize comfort and skin protection is determined.
Socket fit coupled with the choice of a liner are key considerations in assuring comfort and acceptance of the prosthetic limb. The liner functions as the primary interface between the residual limb and the remainder of the prosthesis. In this role, it must complement socket fit to ensure optimal pressure distribution while also eliminating harmful shear forces and providing a favorable moisture, heat, and chemical environment that prevents skin breakdown. PTB total contact sockets can be fit as hard sockets that do not use a liner or more commonly use a liner made from closed cell foam such as Pe-Lite for improved comfort. Using PTB total contact sockets and pelite liners are often an advantage in the preparatory prosthesis because of the relative ease with which the liner can be modified to accommodate changes in residual limb volume (52,55,56). Roll-on silicone or elastomeric gel liners are another option that can be used with a PTB total contact socket but are generally recommended for use with TSB socket designs. Gel liners are thought to enhance comfort and reduce shear, making them the initial choice for residual limbs with scarring or skin grafts that compromise skin integrity (52,56). Gel liners result in more sweating and are generally less tolerated in warm climates than other liners. Contraindications to the use of gel liners are residual limbs with open wounds, poor hygiene, or a history of contact dermatitis.

The suspension system for the transtibial prosthesis must securely attach the limb during activities, minimize pistoning, and be comfortable when sitting. When working with individuals with an amputation who run, play sports, or are involved in climbing activities, ensuring effective suspension is especially important. Suspension systems can be grouped into categories that include straps, sleeves, gel liners with locking mechanisms, and suction (Fig. 74-6). A commonly used suspension system is the supracondylar cuff strap. Several variants exist, all of which consist of a multipart strap that attaches to the sidewalls of the socket and encircles the distal thigh using the normal anatomic flare of the supracondylar portion of the femur to maintain suspension. The supracondylar strap is inexpensive, easily applied, and is comfortable during sitting. It supplies adequate suspension for the low-to-moderate activity-level individual with an amputation and is often the best option when impaired hand function limits grip strength and coordination. Waist belts with an anterior fork strap that attaches to the socket are a rarely used suspension option. This type of suspension is most commonly used in conjunction with a PTB total contact socket with side joint and a high corset. The weight and bulk of the resulting limb makes it a poor initial choice for a contemporary prosthetic limb, but it does remain a useful option for long-term users of this design of prosthesis. Other individuals with an amputation who may benefit from this type of limb are those with a short transtibial amputation who require maximal medial-lateral stability for outdoor or work activities. This type of limb may also be preferred when coexisting ligamentous instability of the knee is present, or to partially off load a painful or weight-intolerant residual limb.

Sleeve suspension systems consist of rubber, neoprene, or elastic sleeves that are pulled up onto the distal thigh after donning the prosthesis (56). Sleeves are a general purpose suspension system that is inexpensive and effective for individuals with an amputation across a wide spectrum of activity levels. The primary disadvantages are related to excessive heat or sweating, the need for good grip strength to pull the sleeve up, and the occasional occurrence of contact dermatitis, especially with the use of neoprene-based sleeves. Silicone and elastomeric gel liners are prosthetic sock-shaped sleeves made from a variety of silicone and urethane elastomeric compounds that are rolled onto the residual limb. They function as both an interface and suspension method. The suspension function requires either a metal pin attached to the distal end of the liner which inserts into a locking mechanism in the bottom of the socket or by a Velcro lanyard strap that passes through a slot in the socket and connects with its counterpart attached to the outside of the socket. The suspension pin or lanyard securely anchors the liner to the socket, and the subsequent friction and suction that develops between the liner and the residual limb.

**FIGURE 74-6.** Suspension systems used for the transtibial prosthesis.
Most contemporary prostheses are endoskeletal in design. Using an endoskeletal pylon allows alignment changes after prosthetic fabrication and enables the use of additional components that can absorb forces or allow motion between the socket and the remainder of residual limb. Commonly used components include transverse rotators that reduce axial torques and vertical shock absorbers that cushion impact loading and may reduce oxygen consumption (46,58). The selection of a prosthetic foot completes the transtibial prosthesis prescription.

Knee Disarticulation

Knee disarticulations (KD) amputations share some of the same advantages and disadvantages as the Syme amputation (ankle disarticulation) (59). Similar to the Syme amputation, full weight bearing on the distal end of the KD residual limb is usually possible and the anatomic flare of the femoral condyles can be used for self-suspension of the prosthesis. Because of the improved distal weight bearing, the KD amputation does not require an ischial weight-bearing socket leading to enhanced comfort and sitting tolerance as does a transfemoral amputation. The KD has a bulbous distal end, which compromises prosthetic cosmesis. Compared to the individual with a transfemoral amputation, the long length of the KD residual improves the prosthetic control and allows a greater degree of dynamic muscular stability. However, the long residual limb limits the choice of prosthetic knee units that can be used to maintain symmetric knee centers between the amputated and nonamputated side. Advances in the design and development of the four-bar linkage or polycentric knee units (Fig. 74-9) offer good biomechanical function and acceptable limb cosmesis. KD, while an uncommon amputation level, is reemerging as an alternative to the transfemoral amputation when wound healing concerns are acceptable because of the improved sitting balance, reduced energy cost of walking, and better acceptance rate than that for a transfemoral amputation (60).

FIGURE 74-7. TSB socket with gel liner and pin system, and custom-fabricated liners to accommodate any irregularities of the residual limb. If pin or lanyard system is not being used then the socket will be held by a suspension sleeve. (Courtesy of Otto Bock and Evolution Liners, see Web site listing.)

FIGURE 74-8. VASS works by use of a vertical shock pylon that acts like a vacuum pump and continually withdraws air from the sealed socket while ambulating. (Courtesy of Otto Bock, see Web site listing.)
Transfemoral Amputation
The development of new technology, materials, and prosthetic components over the past decade has arguably had the greatest impact on the care of the individual with an amputation at the transfemoral level. The following discussion of transfemoral prosthetic components is organized to parallel a reasonable approach to prescribing a prosthesis. Initially, the socket style is specified. Next the residual limb interface/liner and suspension system are determined. Because liner and suspension options are closely linked at the transfemoral level, they are discussed together. Following selection of socket and suspension, the knee unit is selected and finally pylon and foot/ankle components are chosen.

The quadrilateral socket introduced following World War II had been the standard socket design for transfemoral prostheses until the emergence of a new socket design, the ischial containment socket (ICS) during the past 15 years. The quadrilateral socket, named for its quadrilateral shape as viewed in the transverse plane, is designed for ischial weight bearing on the posterior brim. Quadrilateral sockets are still used for long-term wearers who have become accustomed to its weight-bearing and control characteristics but also remain an option for the individual with a transfemoral amputation with a long residual limb and for individuals who require UE aids for trunk stability (61) (Fig. 74-10). However, for the majority of individuals with new transfemoral amputations, the ICS is believed to provide a more normal anatomic alignment of the femur inside the prosthesis. This is accomplished by extending the socket trimlines proximally and contouring the medial aspect of the socket to capture the ischial tuberosity inside the socket rather than allowing the tuberosity to sit on the posterior brim as in the quadrilateral socket. The ICS allows more effective hip and pelvic stabilizing forces to be developed during stance phase (61–63) which improves the medial-lateral control of the trunk. While the ICS is useful for all individuals with transfemoral amputations, the improved stability is especially beneficial for those with a short residual limb.

Several variants of the ICS have been introduced that include the normal shape-normal alignment, contoured adducted trochanter-controlled alignment method, the Sabolich ICS, and the Northwestern ICS and the M.A.S. (Marlo Anatomical Socket) ICS (61–65). These options vary somewhat in contour details but all retain the basic functional characteristics of ICS. No clear consensus has emerged favoring one-design variant over the others but all require considerable skill to fabricate correctly. Both the ISC and quadrilateral sockets can be fabricated as a rigid socket or by using a flexible thermoplastic inner socket supported by graphite-reinforced, laminated open framework (66,67). The advantages of the rigid frame-flexible liner type of socket designs are flexible walls that increase comfort, especially at the proximal brim, improved proprioception, accommodation of minor volume changes, less heat build-up, and enhanced suspension as the inner socket warms during use increasing its flexibility and improving the intimacy of the fit (63).

A number of interface liner systems and suspension options exist for the transfemoral prosthesis (Fig. 74-11)
The simplest liner system uses a hard socket with wool or cotton socks adjusted in plies to achieve a comfortable fit. Suspension for this type of prosthesis most commonly uses a Silesian belt or total elastic suspension (TES) belt. The Silesian belt attaches to the anterior and lateral portions of the proximal prosthetic socket and passes over the opposite iliac crest. The TES belt is made of the same neoprene material used for transtibial suspension sleeves. It slips over the outside of the prosthetic socket and surrounds the waist above the iliac crests to provide suspension. Both Silesian and TES suspension systems are simple to don, can be adapted for use by individuals with impaired hand function, and usually provide acceptable prosthetic limb suspension for low activity-level patients. Disadvantages include some inevitable pistoning of the prosthesis, reduced comfort due to bandage pressure, and heat or occasional dermatitis, especially with the TES belt. The TES and Silesian suspension systems can be used as a primary suspension system or coupled with other suspension options (discussed below) as an axillary suspension system when additional suspension security or rotation control is needed. Other suspension belt options include the pelvic band and hip joint. This option uses a single-axis hip joint integrated into the lateral socket wall, which is attached to a pelvic band and belt closely contoured about the iliac crest. The side joint and band discourage rotation of the prosthesis and extend the lateral lever arm stabilizing the prosthesis, making it especially useful for the short residual transfemoral prosthesis.

Suction socket suspension is the second major type of suspension for the individual with a transfemoral amputation and, whenever feasible, is generally the preferred option (56). Suction sockets are total contact sockets worn directly against the skin of the residual limb that incorporate a one-way air valve in the distal socket. To don the prosthesis, the individual with an amputation has to pull the residual limb into the socket. This can be accomplished using an elastic stockinette, an Ace bandage, or an “EZ pull” sock made from a thin slippery nylon fabric. An alternative method to pull socks uses a “wet fit” process in which a liquid powder is applied to the residual limb, temporarily lubricating the skin, which can then be slipped into the socket. Once the socket is donned, the suction valve is functional and the remaining air is expelled through the valve, creating a small vacuum that holds the socket on the limb. For most users, suction suspension provides a very secure and comfortable suspension effect free from external belts or straps. Suction suspension requires a stable residual limb volume, and as a result is not a good option for preparatory prostheses when rapid limb shrinkage is expected. The presence of scar tissue can compromise the ability to maintain suction and may be poorly tolerated by the inherent high skin friction present in the socket. Prostheses that use suction sockets must be donned while standing and require both good balance and adequate hand strength and coordination to manipulate the pull sock and to install the valve. Difficulty with consistently pulling all of the proximal soft tissues of the thigh into the socket can lead to an adductor roll (a compression of the soft tissues of the medial groin between the proximal socket and the pelvis) that is painful and can lead to skin breakdown. Correcting this problem may require socket modifications, changes to the donning technique, or the use of external vacuum pumps to help elongate the tissue of the residual limb to promote a better fit (68).

A variant of the hard socket uses hypobaric socks, conventional limb socks with an impregnated silicone band near the proximal end of the sock. The silicone band provides a seal to prevent air leak, allowing for a partial suction suspension effect that, when coupled with an auxiliary Silesian suspension or TES, is an alternative option to the standard suction socket that can accommodate changing residual limb volume secondary to limb shrinkage or weight change.

Similar to a transtibial prosthesis, silicone and elastomeric gel suspension liners with a lanyard strap or pin attachment can be used to provide a beltless pseudosuction type of suspension. This suspension system provides good suspension, minimal pistoning, effective control of adductor rolls, and good comfort across a wide range of activity levels. Socks can be added over the suspension liner to accommodate for volume change of the residual limb. Effective use of gel liners can be problematic in the residual limb with excess or loose soft tissue because of difficulties in consistently donning the liner, excessive traction on soft tissue during swing, and poor rotational...
control of the prosthesis. The use of auxiliary TES or Silesian suspension may be needed in this situation.

During the last decade, a wide variety of prosthetic knee joints have been designed, fabricated, and made commercially available. A recent survey cataloged over 200 knee units, ranging from the simple single-axis knee joints to the completely computerized knee units (Fig. 74-12). The primary purposes of the prosthetic knee are to provide stability during stance (to prevent knee buckling), either through alignment or mechanical means, knee motion during swing to permit clearance of the toe, and adequate flexion to allow the knee to bend when sitting. The knee unit should control the heel rise of the shank, assisting or resisting the acceleration and deceleration of the shank during the swing phase. The selection of an appropriate knee unit is primarily based on matching the stance phase stability features and the swing control aspects of the knee to the anticipated activity level and usage of the prosthesis (69–71). Hydraulic or pneumatic knee units are used for the K3 and K4 individual with an amputation and provide either swing phase control or swing and stance phase control for individuals who change their cadence frequently. For the K1 or K2 individual with an amputation who has difficulty maintaining knee stability during stance, locking knees and weight-activated–stance control knees are available. The four-bar, five-bar, and six-bar knee units are becoming increasingly popular due to the moveable centroid, which gives better stability of the knee. Table 74-5 outlines the major types of knee joints and their advantages and disadvantages.

The same pylon components (vertical shock pylons [VSPs] and rotators) that are available for the individual with an amputation at the transtibial level are also available at the transfemoral level. In addition, a thigh rotator can be added for the individual with a transfemoral amputation who has a need to cross the prosthetic leg for ADLs (i.e., riding in a car, donning and doffing clothing or shoes, etc.) (Fig. 74-13).

**Hip Disarticulation/Hemipelvectomy**

Individuals with a transfemoral amputation, with less than 5 cm of residual femur, usually are fitted at hip disarticulation level. The standard prosthesis for a hip disarticulation is the Canadian hip disarticulation prosthesis (Fig. 74-14). The socket of this prosthesis encloses the hemipelvis on the side of the amputation and extends around the hemipelvis of the nonamputated side, leaving an opening for the nonamputated LE. There is a flexible anterior wall with an opening that allows the prosthesis to be donned. Weight is borne on the ischial tuberosity of the amputated side. Endoskeletal prosthetic components are preferred for this level of amputation to reduce the overall weight. The endoskeletal hip joint has an extension assist, as does the knee unit, which usually is a constant-friction knee. Endoskeletal components may be made from aluminum, titanium, or carbon graphite composite materials. Traditionally, a single-axis or SACH foot with a soft heel has been the most common choice for the prosthetic foot. The newer lightweight foot/ankle combination such as the Endolite foot/ankle complex or the Endolite ankle with a Seattle Lite foot may be a better option for this level. A cosmetic cover completes the prosthetic prescription. If necessary, locking hip or knee joints can be used.

The prosthesis for a hemipelvectomy resembles that for the hip disarticulation except in the interior configuration of the socket. In the hemipelvectomy, most of the weight is borne by the soft tissues on the amputated side, with some of the weight being borne by the sacrum, the rib cage, and the opposite ischial tuberosity.

**Hemicorporectomy**

Although rare, the hemicorporectomy (HCP), or translumbar amputation, has the most significant physiologic and psychologic implications. During this procedure the bony pelvis, pelvic contents, lower extremities, and external genitalia are removed following disarticulation of the lumbar spine and transection of the spinal cord. HCP is usually a last resort for patients with life-threatening conditions such as advanced pelvic tumors, pelvic osteomyelitis, crushing pelvic trauma, or intractable decubiti in
### TABLE 74.5 General Information on Prosthetic Knees and Usage

<table>
<thead>
<tr>
<th>Type of Knee</th>
<th>Advantage</th>
<th>Disadvantage</th>
<th>Possible Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single-axis constant friction</td>
<td>Simple</td>
<td>Only constant swing phase control</td>
<td>Excellent for pediatrics</td>
</tr>
<tr>
<td></td>
<td>Durable</td>
<td>No stance control</td>
<td>Having good voluntary control of swing and stance phase but</td>
</tr>
<tr>
<td></td>
<td>Low-maintenance</td>
<td>Single cadence</td>
<td>single cadence</td>
</tr>
<tr>
<td>Polycentric (not fluid control)</td>
<td>Varying stability through stance.</td>
<td>Increased weight and bulk</td>
<td>Knee disartics</td>
</tr>
<tr>
<td></td>
<td>Shortens shank during swing for better toe</td>
<td>Complex mechanism</td>
<td>Long transfemoral for appearance and short transfemoral</td>
</tr>
<tr>
<td></td>
<td>clearance.</td>
<td>Single cadence</td>
<td>for knee stability</td>
</tr>
<tr>
<td></td>
<td>While sitting, some give natural and better</td>
<td></td>
<td>Weak hip extensors</td>
</tr>
<tr>
<td></td>
<td>cosmetic appearance.</td>
<td></td>
<td>Geriatrics</td>
</tr>
<tr>
<td>Weight-activated—stance control</td>
<td>Do not have adequate control to</td>
<td>Requires regular maintenance.</td>
<td>Short residual limb</td>
</tr>
<tr>
<td></td>
<td>manage a bending knee or good enough hip</td>
<td>Not very responsive for active walker.</td>
<td>General debility</td>
</tr>
<tr>
<td></td>
<td>control to stabilize braking mechanism if</td>
<td></td>
<td>Uneven surfaces</td>
</tr>
<tr>
<td></td>
<td>weight applied with knee flexed 0–20 degrees</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Helpful for slower clients</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Manual lock</td>
<td>Total stability in stance phase.</td>
<td>No swing phase flexion, resulting in stiff knee</td>
<td>Patient requires mechanical stance</td>
</tr>
<tr>
<td></td>
<td></td>
<td>gait</td>
<td>stability in stance</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Awkward in sitting</td>
<td>Last resort</td>
</tr>
</tbody>
</table>

**Fluid Control**

<table>
<thead>
<tr>
<th>Type of Knee</th>
<th>Advantage</th>
<th>Disadvantage</th>
<th>Possible Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single-axis pneumatic control</td>
<td>Responds to changing gait speeds.</td>
<td>Higher cost</td>
<td>From pediatrics to adults, with go control</td>
</tr>
<tr>
<td>Single-axis hydraulic control</td>
<td>Swing respond to changing gait speeds.</td>
<td>May need more maintenance</td>
<td>From pediatrics to adults, with go control</td>
</tr>
<tr>
<td></td>
<td>If stance phase has hydraulic resistance to</td>
<td>Heavier</td>
<td>Excellent reliability</td>
</tr>
<tr>
<td></td>
<td>knee flexion during weight bearing</td>
<td></td>
<td>For the more active amputee</td>
</tr>
<tr>
<td>Polycentric and multiaxis fluid</td>
<td>Varying stability through stance.</td>
<td>Higher cost</td>
<td>Knee disartics</td>
</tr>
<tr>
<td>control</td>
<td>Shortens shank during swing for better toe</td>
<td>May need more maintenance</td>
<td>Long transfemoral for appearance and short transfemoral</td>
</tr>
<tr>
<td></td>
<td>clearance.</td>
<td>Heavier</td>
<td>for knee stability</td>
</tr>
<tr>
<td></td>
<td>While sitting some give natural and</td>
<td></td>
<td>Range from homebound to highly active amputee</td>
</tr>
<tr>
<td></td>
<td>better cosmetic appearance.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Variable cadence</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Microprocessor Control**

<table>
<thead>
<tr>
<th>Type of Knee</th>
<th>Advantage</th>
<th>Disadvantage</th>
<th>Possible Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single axis or multiaxis fluid</td>
<td>On board computer adjusting knee for variable</td>
<td>Highest cost</td>
<td>For the more active patients. Some computerized knees use a computer-</td>
</tr>
<tr>
<td>control</td>
<td>gait cycles.</td>
<td>Heavy</td>
<td>regulated valve to adjust the swing phase resistance of a pneumatic</td>
</tr>
<tr>
<td></td>
<td>Energy saving</td>
<td>Unproven track record for dependability</td>
<td>cylinder. Some use the computer to control swing phase function</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>and stance phase stability. Some systems using multiple sensors to</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>send messages about changes walking back to the microchip 50 times a</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>second.</td>
</tr>
</tbody>
</table>
the pelvic region. Once the amputation site is adequately healed then a prosthetist should begin fabrication of the prosthesis.

The prosthesis is bucket-shaped with two cut apertures to accommodate the patient’s colostomy and ileostomy sites. This type of prosthesis is similar to those used in other cases of HCP (72). When using the prosthesis for the first time, the patient must limit the use and allow frequent checks for skin breakdown for the next several weeks. At first, the base of the prosthesis is attached to a flat board to provide a level seating surface and possibly later replaced with a rocker type board for greater mobility when transferring to and from a wheelchair and promoting easier wheelchair use.

**UE AMPUTATION**

**General Principles of UE Amputation Management**

The general approach to the care of the individual with a UE amputation parallels that used in treating a lower limb amputation outlined earlier in this chapter. The main differences will be highlighted and areas of unique management discussed.

**Preprosthetic Patient Evaluation and Management**

The preprosthetic evaluation, whether performed preoperatively or postoperatively, should focus on identifying factors that will affect rehabilitation and prosthetic fitting. Of particular importance is the presence of coexisting musculoskeletal or neurologic compromise such as significant shoulder range of motion loss or brachial plexus injury that would compromise or prevent the successful use of prosthesis. Initial education about the rehabilitation process, prosthetic options, and a discussion of realistic expectations should occur during this stage.

**Postoperative Care**

The options for postoperative care of the residual limb are similar to those used in the LE. Soft dressings along with stump shrinkers or elastic compressive stockinettes are the most commonly used wound care approach. In centers with experience in the application of RRDs, their use in the immediate and early postoperative period is advocated because of presumed benefits of better residual limb protection, edema control, and pain reduction. RRDs work best at the transradial amputation level because of the ease in suspending the cast. An additional advantage of RRDs is the ability to implement immediate or early postprosthetic fitting by attaching a hook, harness, and control system to the cast. Early prosthetic fitting is particularly important for the individual with a bilateral UE amputation to avoid the potentially profound psychology consequences resulting from the complete dependence that accompanies the loss of both UEs.

**Prosthetic Fitting**

Even more so than in the individual with an LE amputation, an understanding of the individual’s vocational and avocational needs, motivation in pursuing prosthetic fitting, availability of a supportive social network, and cosmetic concerns are critical in goal setting and decision making regarding prosthetic prescription. There are no fixed guidelines for deciding which individual with an upper limb amputation will benefit from a prosthetic limb. Factors that often prevent successful fitting and training include significant cognitive deficits, upper motor neuron syndromes that impair volitional coordination of the residual limb, plexopathies or peripheral nerve injuries, significant sensory loss, hyperpathia of the residual limb, or range of motion loss in the shoulder. The typical sequence of prosthetic rehabilitation involves initial fitting and training with a body-powered preparatory prostheses followed by a definitive prosthesis when residual limb maturation has occurred 3 to 6 months postsurgery. Advantages of using a body-powered preparatory prosthesis include the greater ease of fitting, greater adaptability to changes in residual limb volume, early training in ADLs, and a lower cost compared to external powered devices. Since it is difficult to predict either the likelihood of long-term acceptance of an upper limb prosthesis or an individual’s preferred type of prosthetic system (body or external powered), the preparatory prosthetic phase should be used to explore various prosthetic options. Different types of terminal devices

**FIGURE 74-14.** Canadian type hip disarticulation with free motion hip joint and four bar knee with outline of cover. (Courtesy of Otto Bock, see Web site listing.)
should be tried and the individual with a UE amputation given enough time to explore the advantages and disadvantages of using the prosthesis in a variety of home, work, and social settings. This type of clinical trial allows more informed decision making and participation of the patient at the time of definitive prosthetic prescription.

The overall acceptance and continued use of a prosthetic limb is influenced by a number of factors (73–76). Similar to the LE, as the amputation level moves more proximal, the prosthetic rejection rate increases. The majority of individuals with an amputation at the shoulder disarticulation or forequarter level ultimately reject prostheses. Approximately 40% of individuals with a THA become long-term users of a prosthetic limb. At the transradial amputation level, approximately 90% of individuals continue to use a prosthesis occasionally. More distal amputations in the hand have a lower rate of prosthetic limb acceptance, typically because of the greater preserved function of the residual limb. Individuals with bilateral amputations at all levels tend to have a higher acceptance rate for prosthetic limb use. Early work by Malone et al. (73) indicated that early fitting within 30 days postoperatively was associated with higher long-term acceptance. More recent studies (74,75) have not shown as strong a correlation, suggesting that even delayed fitting can be successful. Common reasons for rejecting a prosthetic limb include limited usefulness, excessive weight, residual limb pain, and poor durability (myoelectric prosthesis).

**Final Prosthetic Evaluation and Control Training**

Prosthetic training is an important component of the rehabilitation process and affects the successful use and acceptance of the prosthesis (77). General strengthening and range of motion exercises for the proximal residual limb started in the preprosthetic phase should continue. After the prosthesis has been fabricated, it should be checked by the members of the prosthetic team to ensure that the fit is comfortable and that the control system is properly adjusted for maximum functional operation. During initial use and training, the prosthetic socket should be removed every 15 to 30 minutes to check for signs of excessive pressure or irritation that may occur with poor socket fit or overuse. As the skin tolerance increases, wearing time is gradually increased during the first few days of wear and thereafter more quickly. With a well-fitting prosthesis, an individual with a UE amputation can wear the limb for an entire day within a week or two of receiving it.

Specific training involves instruction in how to put on and remove the prosthesis, adjusting the number of sock plies, and how to clean and care for the residual limb and prosthesis. Initial skills are acquired to control the terminal device and to activate position and lock the elbow in the individual with a THA. Grasping and releasing objects, transferring objects, and positioning of the terminal device for functional activities should be practiced under the guidance and feedback of an experienced therapist. ADLs, homemaking, and occupational and recreational activities should be undertaken and simulated in the training sessions. Successful prosthetic users rely on basic skills learned in therapy, which are supplemented and polished by practice at home during everyday tasks. The individual who is a long-term user of the upper limb often has very specific use patterns and preferences for components, harness type, and control cable adjustments. It is prudent to incorporate the desires of the individual into the decision-making and prescription process.

Control training for externally powered prostheses is more complex than for body-powered UE prostheses. The same goals as outlined for conventional prostheses are appropriate for myoelectric controlled prostheses. In addition, the individual must learn to separate, modulate, and sustain voluntary muscle contractions in the muscles selected to control the powered functions of the prosthesis. Training with an externally powered (electric) prosthesis often requires more time than with a body-powered prosthesis because more exacting individual muscle motions and quantity of muscle contraction are required for control of the prosthesis.

Adults and older children can be expected to practice specific tasks and routines, both in therapy and at home, as outlined by the therapist to achieve the necessary skills for independent function. The training time in a young child with a UE prosthesis will be significantly longer than that for an adult or older child.

**UE Prosthetic Follow-Up**

The routine follow-up visits for an individual with a new amputation should occur initially 4 to 6 weeks after delivery of the prosthesis, then every 2 to 6 months until a definite prosthesis is prescribed. Once clinically stable in a definitive prosthesis, yearly clinic visits or whenever a problem arises, are usually adequate. At these follow-up visits, the individual’s use and function with the prosthesis should be reviewed, difficulties or problems resolved, the fit and condition of the prosthesis evaluated, and the condition of the residual limb noted. If necessary, additional therapy may be suggested, repairs to the prosthesis made, medical problems with the residual limb addressed, and a new prosthesis prescribed if indicated. With average use, a UE prosthesis can be expected to be worn for 3 to 5 years before total replacement is necessary. The socket itself may need to be replaced more frequently than the other components.

Although our emphasis has been on prosthetic restoration, the focus of rehabilitation should remain on the individual with the UE amputation and his or her desired lifestyle following limb loss. Many individuals with a UE amputation do well without the aid of a prosthesis, and should not be viewed as having failed if they choose not to wear a prosthesis.

**Bilateral UE Amputations**

The individual with a bilateral UE amputation is immediately faced with the loss of ability to perform almost every ADL. Early restoration of any ADL is important. Providing a utensil cuff, which can be attached to a residual arm, can assist the patient with feeding and tooth brushing. In the individual with a bilateral
UE amputation, early prosthetic fitting should be accomplished, even with temporary or preparatory prostheses. In the individual with a bilateral amputation, the longer residual limb usually assumes dominance. Special component considerations apply in this case. Wrist flexion units, at least on the dominant side or perhaps bilaterally, will permit mid-line activities such as shirt buttoning, belt buckling, and toileting. Also wrist rotator units, which provide terminal device positioning, provide for easier bilateral prosthetic use. Special toileting techniques must be taught for patient independence. In addition, foot skills should be reviewed, and LE mobilizing exercises should be performed.

**UE Prostheses**

A UE prosthesis attempts to replace very complex functions. The hand is able to perform a wide range of functional activities ranging from fine dexterity tasks requiring light prehensile forces to gross grasping movements with great prehensile forces, all under the guidance of sensory feedback. To accomplish these tasks in the dynamic world around us, the coordinated movements of the proximal muscles and joints of the UE must position the hand for functional activities. The prosthetic replacement for the UE is a limited substitute for the lost body part. With practice, it can replace several of the simple grasping and manipulating functions of the hand through a mildly to moderately restricted sphere of functional reach. A critical limitation of all UE prosthetic limbs is the severe restriction in sensory feedback from the terminal device relegating its use to that of simply assisting bimanual activities.

The limited ability of the prosthetic limb to replace normal hand function typically results in a shift in hand dominance with amputation of the dominant hand.

UE prostheses can be divided into three groups: conventional or body-powered, external powered or electric, and passive or cosmetic. The advantages and disadvantages for each of the three general types of upper limb prostheses are presented in Table 74-6.

**Body-Powered Prostheses**

Since the majority of individuals with UE amputation first learn to use a body-powered limb, an initial discussion of the function of various components (78,79) used in this type of prosthetic limb follows. Understanding the role, use, and limitations of body-powered prosthetic limbs forms the framework for subsequent discussions of externally powered devices. All conventional body-powered UE prostheses have these component parts:

- **Socket**
- **Suspension**
- **Control cable system**
- **Terminal device**
- **Interposing joints (wrist, elbow, shoulder) as needed by the level of amputation**

**Socket**

Traditionally UE prostheses have used a dual-wall socket design fabricated from lightweight plastic or graphite composite.
materials. In dual wall designs, a rigid inner socket is fabricated from a custom mold of the residual limb and is the primary interface between the user and the prosthesis. Comfort and function are directly tied to the quality of the fit of the inner socket. The outer socket wall is fabricated to have the general shape, length, and contour of the normal arm or forearm, and serves a cosmetic function and also supplies the foundation for the mounting of required components. This type of socket is durable and easily accommodates variation in residual limb volume using socks to adjust the fit.

An alternative approach to the design of the socket parallels the rigid frame, flexible liner approach used in lower limb prostheses. An inner socket is fabricated from flexible plastic materials to provide a total contact fit and is optimized for the use of suction suspension. Surrounding the inner socket is a rigid frame that provides the structural integrity of the socket. Windows in the outer socket allow for the movement of the muscles of residual limb to enhance comfort.

**Suspension**

The suspension system must hold the prosthesis securely to the residual limb and accommodate and distribute the forces associated with the weight of the prosthesis and any superimposed lifting loads. Suspension systems can be classified as harness-based, self-suspending sockets, or suction. Table 74-7 outlines the types of harnesses and their suggested uses, advantages, and disadvantages.

The most commonly used suspension for body-powered prostheses are variants of the harness-based system. The most commonly used harness is the figure-8 strap (Fig. 74-15). On the intact side, the harness loops around the axilla anchoring the limb and providing the counterforce for suspension and control cable forces. On the prosthetic side, the anterior strap carries the major suspending forces to the prosthesis by attaching directly to the socket (transhumeral) or indirectly through an intermediate Y-strap and triceps pad (transradial). The posterior strap on the prosthetic side attaches to the control cable. For heavier lifting activities or when axillary pressure from a figure-8 harness is unacceptable, a shoulder saddle with chest strap system is a useful alternative suspension system. A leather or flexible plastic saddle is positioned over the prosthetic side shoulder and is secured using a chest strap that wraps around the intact chest wall in an infra-axillary location. The suspension of the prosthesis occurs through an inverted U-shaped cable or strap that is anchored to the top of the saddle and drapes downward to attach to the front and back of the prosthetic socket (transhumeral) or triceps pad (transradial). The control cable attaches to the posterior aspect of the chest strap.

Self-suspending socket designs are used when the bony configuration of the residual limb allows suspension, similar to the effect used in Syme amputations. Self-suspending sockets are largely limited to wrist and elbow disarticulations. In both harness and self-suspending systems, socks can be worn to optimize fit and improve comfort.

Suction suspension sockets are similar to those available to the individual with an LE amputation and use either a total contact socket with a one-way air valve or a roll-on silicone gel sleeve with a locking pin. Suction sockets are appropriate primarily for the transhumeral residual limb with good soft-tissue

**TABLE 74.7 General Information on UE Prosthetic Suspension Options**

<table>
<thead>
<tr>
<th>Suspension</th>
<th>Indications</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Harness</td>
<td>Figure-8</td>
<td>Transradial Transhumeral Light to normal duty activities Transradial Transhumeral Heavy lifting</td>
<td>Simple, durable, adjustable. Axillary pressure reduces discomfort.</td>
</tr>
<tr>
<td>Shoulder saddle and chest strap</td>
<td>Transhumeral</td>
<td>Greater lifting ability, more comfortable than figure-8</td>
<td>Difficult to adjust in women because straps cross breasts. Reduced control compared to figure-8 harness</td>
</tr>
<tr>
<td>Self-suspending</td>
<td>Muenster Northwestern supracondylar</td>
<td>Wrist disarticulation Elbow disarticulation Short transradial Myoelectric transradial</td>
<td>Ease of use</td>
</tr>
<tr>
<td>Suction</td>
<td>Suction socket with air valve</td>
<td>Transhumeral with good soft-tissue cover</td>
<td>Secure suspension, elimination of suspension straps</td>
</tr>
<tr>
<td>Gel sleeve with locking pin</td>
<td>Transradial Transhumeral Compromised limbs with scarring or impaired skin integrity</td>
<td>Accommodate limb volume change with socks, Reduced skin shear.</td>
<td>Requires stable residual volume, harder to don than other suspension systems Greater cleaning and hygiene requirements. Can be uncomfortable in hot climates</td>
</tr>
</tbody>
</table>

---

**Figure 74-15**

Transradial Transhumeral Transradial Transhumeral Transradial Transhumeral Heavy lifting Greater lifting ability, more comfortable than figure-8

---

**Table 74-7 outlines the types of harnesses and their suggested uses, advantages, and disadvantages.**

**Suspension**

The suspension system must hold the prosthesis securely to the residual limb and accommodate and distribute the forces associated with the weight of the prosthesis and any superimposed lifting loads. Suspension systems can be classified as harness-based, self-suspending sockets, or suction. Table 74-7 outlines the types of harnesses and their suggested uses, advantages, and disadvantages.

The most commonly used suspension for body-powered prostheses are variants of the harness-based system. The most commonly used harness is the figure-8 strap (Fig. 74-15). On the intact side, the harness loops around the axilla anchoring the limb and providing the counterforce for suspension and control cable forces. On the prosthetic side, the anterior strap carries the major suspending forces to the prosthesis by attaching directly to the socket (transhumeral) or indirectly through an intermediate Y-strap and triceps pad (transradial). The posterior strap on the prosthetic side attaches to the control cable. For heavier lifting activities or when axillary pressure from a figure-8 harness is unacceptable, a shoulder saddle with chest strap system is a useful alternative suspension system. A leather or flexible plastic saddle is positioned over the prosthetic side shoulder and is secured using a chest strap that wraps around the intact chest wall in an infra-axillary location. The suspension of the prosthesis occurs through an inverted U-shaped cable or strap that is anchored to the top of the saddle and drapes downward to attach to the front and back of the prosthetic socket (transhumeral) or triceps pad (transradial). The control cable attaches to the posterior aspect of the chest strap.

Self-suspending socket designs are used when the bony configuration of the residual limb allows suspension, similar to the effect used in Syme amputations. Self-suspending sockets are largely limited to wrist and elbow disarticulations. In both harness and self-suspending systems, socks can be worn to optimize fit and improve comfort.

Suction suspension sockets are similar to those available to the individual with an LE amputation and use either a total contact socket with a one-way air valve or a roll-on silicone gel sleeve with a locking pin. Suction sockets are appropriate primarily for the transhumeral residual limb with good soft-tissue
envelope, an absence of invaginated scarring, and a stable volume. Donning the suction socket with air valve involves the use of water-based lubricants or a pull sock to seat the residual limb into the socket. The one-way air valve creates suction inside of the socket, maintaining suspension. When using a silicone gel (or similar polymer materials) sleeve with a distal attachment pin, the sleeve is rolled onto the residual limb and inserted into the socket where the pin mates with a locking mechanism. Suspension is accomplished through a combination of suction and skin friction.

Control Cable Mechanism

Body-powered prosthetic limbs use cables to link movements of the shoulder and humerus to activation of the terminal device and elbow. The movements that are captured for control include scapular abduction, shoulder depression and abduction, and humeral flexion. Control cables used to activate a single prosthetic component such as the terminal device are known as single-control cables or Bowden cables. Dual-control cable systems use the same cable to control two prosthetic functions, typically elbow flexion and terminal device opening in the transhumeral prosthesis. The control cable is attached to the figure-8 or chest strap harness used for suspension of the prosthesis. When a body-powered prosthesis incorporates a self-suspending socket design or suction suspension, a simple figure-8 harness strap system is required to allow control functions to be performed.

Terminal Device

Terminal devices for body-powered prostheses can be hooks, functional or active hands, cosmetic or passive hands, or special terminal devices designed for specific activities (e.g., bowling ball terminal device, golf club holder). Hook-style terminal devices provide the equivalent of a lateral pinch grip while active hand terminal devices provide a three-point chuck action (80). The most commonly used active terminal device is the voluntary-opening hook. While different voluntary-opening hook designs are available for various applications, the most commonly prescribed design for general use is the Dorrance 5X, 5XA, and 7 (Hosmer Dorrance, Campbell, CA) (Fig. 74-16). With voluntary-opening devices, the individual with a UE amputation provides power through a cabling system to open the terminal device and relies on springs or rubber bands to provide the closing prehensile force. Typical closing forces range from 5 to 10 lb. Voluntary-opening active hands are available and in general are more cosmetic, but are heavier, interfere with visualizing the object being grasped, and provide lower prehensile forces. Voluntary-closing terminal devices allow the individual to provide a variable prehensile force transmitted through the control cable to the terminal device. Voluntary-closing devices are capable of providing larger prehensile forces up to 20 to 25 lb (9 to 11 kg) and provide indirect sensory feedback through the force exerted on the control cable. A significant disadvantage of voluntary-closing devices is the need for a constant pull on the control cable during prolonged grasping, a skill that is difficult to accomplish during dynamic tasks. Passive (cosmetic) hands are lighter than active
important for the individual with a bilateral UE amputation. A quick-disconnect option permits the easy interchange of different terminal devices, such as a hook for a hand.

Additional details regarding components, sockets, suspension, and control systems are discussed where appropriate in the following sections relating to specific levels of amputation.

**Prostheses by Level of Amputation**

**Partial Hand**

For partial hand amputations (e.g., phalanges, ray resections, transmetacarpal), a prosthesis may not be necessary. To be functionally useful, the residual hand needs to be able to provide a rudimentary grasp. This requires two opposing posts that can be moved into contact with each other to provide a prehensile force. When possible, surgical reconstruction of the remaining hand is often the preferred approach to preserving or enhancing function while maintaining sensation. When only one movable digit remains, as in a transmetacarpal amputation of digits 2 through 5, either an open or mitt-shaped prosthetic opposition post can be used to provide a stable surface for opposition with the thumb. In general, the substantial variability in the anatomy of the remaining partial hand requires creative custom solutions to optimize function (Fig. 74-17).

At times, an individual with a UE amputation may require devices customized for specific activities. A cosmetic prosthesis is frequently provided for this level of amputation. Long-term usage varies but the majority of individuals continue to wear them at least occasionally for social situations (80).

**Transradial Amputation/Wrist Disarticulation**

The length of the bony forearm, measured from the medial epicondyle, classifies transradial amputations as: very short (<35%), short (35% to 55%), and long (55% to 90%). Longer residual limb length enhances the forearm lever, making lifting easier, and allows capturing residual pronation-supination motion of the forearm. Long transradial residual limbs retain from 60 to 120 degrees of supination-pronation, which decreases to less than 60 degrees in short transradial residual limbs. For short and long transradial amputations, a dual-wall socket is attached to a triceps pad with flexible elbow hinges (straps) to allow pronation-supination. The triceps pad helps to distribute suspension forces and is needed to anchor the control cable. Most commonly, a figure-8 harness system is used for suspension and control. For very short transradial amputation levels, rigid hinges are generally used to provide greater stability of the socket on the residual limb. With transradial amputations, in which range of motion is limited at the elbow, polycentric elbow joints or a split socket with step-up hinges...
can be used to provide additional flexion. The additional flexion gained with the use of these elbow hinges is offset by a loss of elbow flexion power and lifting ability.

The wrist disarticulation (WD) prosthesis is a variant of the transradial prosthesis. Because WD spares the distal radial-ulnar joint, full forearm supination-pronation is preserved. Socket designs for the WD level are flattened distally to form an oval to capture supination-pronation, allowing active rotational positioning of the terminal device during activities. The distal flare of the residual limb can be taken advantage of to fabricate a self-suspending socket but this usually leads to a bulbous, cosmetically compromised appearance similar to that occurring in the LE Syme amputation. The long residual limb necessitates the use of a special thin wrist unit to minimize the overall length of the prosthesis. If cosmesis is of primary importance to the patient, a long transradial amputation may be a more appropriate amputation level.

An example for the prescription of a transradial/WD limb could read: dual-wall total contact socket, flexible elbow hinges, triceps pad, figure-8 harness, single control cable, constant friction wrist unit with quick disconnect, SXA hook, and cosmetic hand.

Transhumeral Amputation/Elbow Disarticulation
The length of the residual humerus measured from the acromion classifies the transhumeral amputation (THA) as: humeral neck (<30%), short transhumeral (30% to 50%), standard transhumeral (50% to 90%), and elbow disarticulation (90% to 100%). For short and standard transhumeral residual limb lengths, the traditional dual-wall socket extends to just below the acromion and attaches to either a figure-8 or a shoulder saddle and chest strap harness for suspension. With shorter residual limbs, securing the socket to the residual, especially under load, is more difficult. To accommodate this problem, the socket extends proximal and medial to the acromion, creating a partial shoulder cap. This socket design often can be suspended with only a chest strap, but other harness systems can be used for additional security or to improve control functions. Suction socket suspension systems are becoming the preferred system for individuals with a THA because of the improved suspension and greater ability to position the limb for activities. Even when suction suspension systems are used, an axillary harness is needed for control and can augment suspension, especially when lifting larger loads. Suction suspension, when used with externally powered myoelectric components, can result in a self-suspending prosthesis free of any harness.

The standard elbow component for the transhumeral prosthesis is the internal elbow joint. Internal elbow units allow for 135 degrees of flexion and can be manually locked into a number of preset flexed positions. The standard internal elbow unit incorporates a turntable that allows passive internal or external rotation of the forearm. Elbow spring-lift assist units are available and are generally recommended for internal elbow units to help counterbalance the weight of the forearm, making elbow flexion easier for the individual with an amputation. This unit requires approximately 5 cm of length. If the level of amputation is less than 2 in (5 cm) proximal to the epicondyles, then an internal elbow unit cannot be used unless an asymmetric elbow unit compared to the intact limb is cosmetically acceptable. When the internal elbow unit cannot be used, locking external elbow joints are available but these are less durable and less cosmetic.

The control system for the transhumeral prosthesis uses two cables: a dual control cable that controls the elbow and terminal device and a secondary elbow locking cable. To control the prosthetic limb, the individual with an amputation uses scapular protraction and humeral flexion to flex the elbow into the desired position. The elbow is locked using the secondary control cable. Once locked, the same shoulder movements that powered the elbow are now available to activate the terminal device. Locking the elbow is typically accomplished by using a control cable that is routed along the anterior aspect of the socket and attaches to the front of the harness. Shoulder depression and humeral extension movements are used to lock-unlock the elbow.

An elbow disarticulation prosthesis is a variant of the transhumeral prosthesis. The socket is flat and broad distally to conform to the anatomic configuration of the epicondyles of the distal humerus. This design provides some self-suspension and allows the individual with an amputation active rotation of the prosthesis (internal and external rotation of the humerus). The length of the residual limb requires the use of external elbow joints, with a cable-operated locking mechanism. The harness is either a figure-8 or a shoulder saddle and chest strap. The control system for this level is the same as for the individual with a more proximal THA.

Example prescription for a short transhumeral prosthetic limb read as: flexible wall/rigid frame suction socket, figure-8 axillary suspension and control harness, dual-control cable, internal locking elbow with turntable and flexion assist, lightweight forearm shell, constant friction wrist unit with quick disconnect, SXA hook, and cosmetic hand.

Shoulder Disarticulation/Forequarter Amputation
For shoulder disarticulation and forequarter amputation, the socket extends onto the thorax to suspend and stabilize the prosthesis. The portion of the thorax covered by the socket is more extensive for the forequarter amputation. In some cases, an open-frame socket rather than a plastic laminated socket is chosen for these levels to reduce prosthetic weight and to minimize heat build up by reducing the amount of skin coverage. Prosthetic components are similar to those for the transhumeral prosthesis with the addition of a shoulder unit, which allows passive positioning of the shoulder joint in flexion-extension and abduction-adduction. Chest straps are attached to the anterior and posterior socket for suspension. The loss of ipsilateral shoulder motion for control purposes severely compromises the use of the prosthesis. A harness and control cable system that uses three individual cables can be used. Individual cables use intact side humeral flexion for prosthetic elbow control, chest expansion for terminal device control, and a manual nudge or pull cable for elbow locking.
The body-powered prosthesis is cumbersome to don, has limited functionality, and is often used mainly for cosmesis. The difficulty in providing suitable body-powered prostheses at these proximal amputation levels argues against their routine use. For many individuals, a cosmetic limb is sufficient. For the highly motivated individual, externally powered prostheses may be more functional and hence can be considered.

### Externally Powered Prostheses

Externally powered prosthetic limbs use small electric motors incorporated into the prosthesis component to control its function. Reliable external power units are available for terminal device operation, wrist rotation, and elbow flexion-extension. Myoelectric signals or switches control these electric motors. It is often difficult to predict the preferred type of power for the prosthetic limb for a particular individual until trial use of both has occurred. Compared to body-powered prostheses, externally powered prostheses are typically heavier, more costly, and less durable, especially if manual labor activities are frequently performed. Important advantages of external-powered prostheses compared to body-powered devices include improved comfort due to the reduced harness needs, better control, and lifting capacity for short transhumeral and shoulder disarticulation amputation levels, and the greater terminal device grip force of electric hooks and hands. In the case of the transradial level amputation, the prosthesis can use a self-suspending socket that eliminates the need for a harness (Fig. 74-18).

The preferred control system for external-powered prostheses is myoelectrical control. Myoelectrical control uses surface electrodes embedded in the prosthetic socket that make contact with the skin and detect muscle action potentials from voluntarily contracting muscles in the residual limb. The myoelectric signal controls an electric motor to provide a function (e.g., terminal device operation, wrist rotation, elbow flexion). Prior to the prescription and fitting of a myoelectric prosthesis, the individual’s ability to reliably contract and relax at least one muscle group in the residual limb should be ascertained. This can be accomplished using electromyographic biofeedback equipment or a myoelectric tester to identify the most appropriate electrode control location(s). Several different types of myoelectric controllers are available depending on the number of prosthetic functions that need to be controlled and the number of useful electrode sites identified. The dual-electrode system uses two sets of electrodes positioned over antagonist muscles allowing natural and intuitive myoelectric control to occur. For example, at the transradial level, activation of the forearm flexor muscles closes the terminal device while activation of the forearm extensors opens the terminal device. Most contemporary myoelectric control systems use proportional control so that the speed or strength of terminal device activation varies with the intensity of muscle contraction. When there are an inadequate number of usable electrode sites available to control all desired functions of the prosthesis, alternative control strategies can be used. Single-site controllers use the strength of voluntary contraction from a single electrode (i.e., amplitude of the myoelectric signal) to control the motion (i.e., a weak contraction will close the terminal device and a strong contraction will open the terminal device) (81). Sequential or multistate controllers use the same electrode pair to control several functions (e.g., terminal device control and elbow activation). This type of controller uses a brief co-contraction to switch between control modes. Any of the alternative control schemes take longer to learn and are not consistently mastered by all individuals with an amputation.

When myoelectric control is not available or there are not enough electrode sites available, switches incorporated into the prosthetic socket or the harness can be used to implement or augment the control of various components. Simple on-off switches can be used to implement basic control of a powered component. More sophisticated servo control techniques based on position or force switches in the harness systems can be used to provide proportional control (82).

Myoelectric prosthetic components are available and sized to fit a wide range of people, from infants to adults. Most myoelectric terminal devices are hands, but some electric hook designs are also available. Myoelectrically controlled prostheses can and have been fit immediately after surgery (83), but it is generally recommended that fitting be delayed until the residual limb is healed and the limb volume has stabilized. A stable limb volume is needed to ensure a consistent socket fit and reproducible electrode position and skin contact. Myoelectric components have been combined with body-powered components to result in a hybrid prosthesis, which may provide better function for some individuals with an amputation than either myoelectric or body-powered control used alone.

### Cosmetic Prosthesis

A conventional body-powered or myoelectric prosthesis with a hand terminal device supplies adequate cosmesis for many individuals with a UE amputation. Using custom covers that are color and texture matched to the skin of the intact side can enhance cosmesis. These covers are expensive and have a limited life expectancy if used during functional activities. When an individual with an amputation is not a candidate for an active prosthetic limb, a cosmetic prosthesis can be fabricated using lightweight components that can be passively positioned to create a symmetric body image while wearing clothing.
Phantom Limb Sensation/Pain
Phantom sensation, defined as the awareness of a nonpainful sensation in the amputated part of the limb, occurs in nearly all acquired amputations (84–86). Phantom sensation is most prominent immediately after amputation and generally diminishes over time, often in a telescoping fashion. The most vivid sensations are typically in the distal portion of the limb. Generally, phantom limb sensations that persist do not require treatment. An occasional problem with phantom sensation occurs when individuals are confused by the phantom sensation and attempt to walk without using a prosthesis.

Phantom pain is pain perceived in the amputated portion of the extremity. The incidence of phantom pain has been difficult to determine with reported frequency ranging from 4% to 80% (70,85,86). Like phantom sensation, phantom pain is most common immediately after amputation and usually becomes less frequent, less intense, and shorter in duration over time. Persistent phantom pain requiring treatment occurs in approximately 5% of individuals with an amputation. Phantom pain is often described as burning, cramping, stabbing, or squeezing but is sometimes reported as bizarre contortions of the limb. The pain may be generalized but more commonly it is experienced mainly in the distal parts of the missing limb. The cause and underlying physiology of phantom pain remain poorly understood but the incidence of phantom pain has been associated with the presence of preamputation pain, phantom sensation, and residual limb pain. The correlation between preamputation pain and the subsequent development of phantom pain has led to attempts at preventing phantom pain by controlling periamputation limb pain with continuous epidural or peripheral nerve anesthesia. These attempts have shown mixed effectiveness (86). Once established, the successful treatment of phantom pain can be difficult and many therapeutic modalities have been tried (70,87–90). Although convincing evidence-based algorithms for the treatment of phantom pain do not exist, empiric clinical guidelines are frequently used. Medication is usually the first line of treatment employing tricyclic antidepressants and anticonvulsants (e.g., carbamazepine, gabapentin), either alone or in combination. Other drugs that have been used with some success include mexiletine, calcitonin, N-methyl-D-aspartate (NMDA) receptor antagonists, and opioids. Evaluation and correction of any coexisting residual limb pain or prosthetic fit problems is also an important component of the initial management of phantom pain. Range-of-motion exercises, relaxation exercises, residual limb massage, transcutaneous electrical nerve stimulation (TENS), compressive stocking, and encouraging prosthesis use have little risk and may be useful adjuncts to medical management. Since problematic phantom pain often occurs only intermittently and is typically short-lived, the patient’s participation in decision making is essential to weigh possible benefits of drug trials with the inconvenience and possible side effects of medications that need to be taken on a regular, continuous basis. Surgical treatments have not been shown to provide lasting pain relief and are rarely used.

Residual Limb Pain
In contrast to phantom limb pain, residual limb pain is the pain perceived as originating in and affecting the residual portion of the limb. Persistent residual limb pain occurs in up to 70% of individuals with lower limb amputation with about half reporting the pain as moderately to severely bothersome (85–88). Residual limb pain is commonly described as aching, sharp, throbbing, and burning in character. The underlying causes of residual limb pain can be classified as intrinsic or extrinsic.

Intrinsic residual limb pain is caused by changes or complications in the underlying neurovascular, bony, or soft tissues of the residual limb. Neuromas develop in all residual limbs after amputation but may only become problematic when entrapped in scar or positioned such that they are exposed to external mechanical loading. Diagnosing an underlying neuroma as the cause of pain can be difficult. When present, neuropathic symptoms including typical dysesthetic pain descriptors, radiation of pain in a specific nerve distribution, and the presence of pain when not using a prosthesis are helpful diagnostic aids but at times the pain location and description is nonspecific. Neuroma-related pain may be precipitated by tapping (Tinel’s sign), direct compression from manual palpation or socket pressure, or traction on an adherent scarred nerve. Larger neuromas can be imaged with magnetic resonance imaging (91). When prosthetic use exacerbates neuroma pain, initial treatment interventions should include prosthetic modifications that attempt to reduce loading of pressure-sensitive areas. Useful approaches include the use of gel socks or liners to better distribute loads and to reduce shearing of adherent tissues, flexible sockets, or socket modification to relieve sensitive areas. Infiltration of the perineuronal region with local anesthetics combined with steroids can be a useful diagnostic and therapeutic procedure. If the injection relieves the pain, a series of similar injections over several months can be attempted to try to achieve lasting relief. When neuroma pain persists and limits function, surgery to resect and move the neuroma to a more protected location can often be helpful. However, following neuroma resection, the neuroma will reform and, on occasion, again become symptomatic.

Bony overgrowth from the distal end of the residual limb skeletal elements occasionally occurs in adults but is primarily a problem that occurs in children with an amputation. Heterotopic bone can form in the soft tissues of the residual limb and may follow trauma, hematoma, or fracture of the residual limb. Poorly contoured bone edges following surgery
can lead to regions of high-pressure concentration. In any of these situations, the abnormal bone leads to localized tissue compression, pain, and tenderness that can progress to the development of adventitial bursa or soft-tissue ulceration. Diagnosis is made with plain radiographs of the residual limb. Management is focused on prosthetic socket modifications to offload painful areas; but achieving a lasting comfortable fit can be difficult. When prosthetic approaches fail, surgical revision is typically needed.

Osteomyelitis, tumor recurrence, stress fractures, and persistent limb ischemia also may occur in patients after amputation. These may cause more generalized residual limb pain and require medical and surgical management.

**Extrinsic residual limb pain** is caused by a mismatch between residual limb tissue tolerance and prosthetic loads imposed on the soft tissues. Poor socket fit or limb malalignment are the main causes. The coexistence of a compromised residual limb from intrinsic pathology reduces the safety margin between tissue tolerance and prosthetic socket loads, making it more difficult to achieve and maintain a comfortable fit. The ability to attain a comfortable socket fit is one of the most important aspects of prosthetic acceptance and function. Most contemporary prosthetic sockets are designed for total contact with modifications to the socket shape to preferentially load weight-tolerance tissues. Socket fit is inevitably compromised with body weight changes or residual limb soft-tissue atrophy over time. Clinical manifestations of poor fit or excessive local tissue loading include a gradual onset of pain while using a previously comfortable prosthesis, erythema persisting more than 15 to 20 minutes after wearing the limb, or the development of blisters, bursas, calluses, or skin ulceration. Changing the number of socks, adding pads to the socket, relieving high-pressure areas, or substituting a gel liner for socks can often restore a comfortable fit and prolong the useful life of a socket. When minor modifications fail, a replacement socket will need to be fabricated.

Malalignment of lower limb prostheses can create abnormally high or prolonged loading forces in the residual limb leading to pain even when the socket fit is acceptable. In the individual with a transfemoral amputation, sagittal plane alignment problems most commonly affect the distal tibia region while frontal plane malalignment primarily affects loading forces along the fibula. In the individual with a transfemoral amputation, distal femur pain is often seen with alignment problems.

**Choke Syndrome**

A specific socket fit problem, almost exclusively seen in the individual with an LE amputation, is the choke syndrome. A choke syndrome develops when there is a simultaneous impairment of venous return from a prosthetic socket that is too tight proximally and a lack of total contact between the residual limb and the socket. Edema develops in the residual limb where total contact is lost. Most commonly, this occurs in the distal aspect of the transfibial residual limb as atrophy takes place over time and additional socks are used to fill in for the loss of soft-tissue volume. Because there is no corresponding volume change in the proximal bony aspect of the limb, the fit becomes too tight and constricts the venous return. Initially, a circumscribed indurated region develops while wearing the prosthesis. If significant edema develops acutely, there can be associated weeping or blistering of the skin. The area of choke is tender to palpation and is prone to developing cellulitis. As the choke syndrome becomes chronic, the tissues become increasingly thickened and indurated, verrucous hyperplasia develops, and skin becomes hyperpigmented because of hemosiderin deposition. Choke syndromes are treated by relieving the proximal constriction and restoring total contact between the residual limb and the socket (90,92). When the choke syndrome is mild, reducing the number of socks to decrease proximal residual limb constriction and adding or modifying the distal end pad may be adequate but typically a new total contact socket is needed.

**Dermatologic Disorders**

Dermatologic problems are common, particularly in the individual with an LE amputation, with surveys estimating that 30% to 50% of individuals with amputations experience one or more skin problems because of a prosthesis (91–94). Dermatologic complaints can be classified as related to hyperhidrosis, physical effects of prosthetic use, contact dermatitis, and infection.

Hyperhidrosis, while not indicative of underlying disease, is one of the most common skin-related complaints and has become more common with the introduction of silicone liners in the 1990s. An increase in sweating is reported by about half of the patients with the use of silicone liners. Over several weeks, adaptation typically occurs and excessive sweating often resolves. Persistent hyperhidrosis makes it more difficult to maintain hygiene, increases the likelihood of skin maceration, and may contribute to the development of contact dermatitis (92–95). When using suction sockets, suspension effectiveness can be compromised. Problematic sweating can be controlled by using concentrated antiperspirants, such as Drysol (Person & Covey, Glendale, CA), on the residual limb or by changing to a liner system that allows the individual with an amputation to wear socks directly against the skin.

The skin of the residual limb is subjected to repeated shear, frictional, and loading forces. The physical effects of these repetitive loads can lead to keratin plugging of sebaceous glands and follicular hyperkeratosis, leading to the development of epidermoid cysts, folliculitis, and dermal granulomas. Cysts are frequently very tender and can spontaneously break open or become secondarily infected. Commonly affected areas are those that are subjected to high loading and shearing forces such as the groin region in the individual with a transfemoral amputation and the medial tibial flare and popliteal fossa in the individual with a transfibial amputation. Treatment is directed at local skin care as well as prosthetic modifications to reduce mechanical skin forces. Meticulous hygiene needs to be encouraged to keep skin, socks, and liner clean. Cosmetic or acne scrub pads can be used to help keep skin pores open. Warm compresses to promote drainage and oral antibiotics are useful in managing folliculitis and infected cysts. Larger and more persistent sebaceous cysts may require surgical drainage.
or excision. Concurrent with medical management, a review of prosthetic fit and alignment to ensure that loading of the affected region is optimized should be undertaken. Recurring problems can be helped with the use of prosthetic components that reduce residual limb shear and loading forces. These include gel liners, rotators, VSPs, and multiaxis ankle/foot devices.

Allergic or contact dermatitis accounts for approximately 20% of prosthetic-related dermatoses (93). Clinical symptoms range from mild lichenification or scaling to weeping eczema. A wide variety of offending agents has been identified. Detergents, scented emollients, creams, and talcs used for skin care should be considered as possible allergens and discontinued or changed. The use of gel sheaths and liners may result in contact dermatitis. The exact allergen is difficult to identify and is probably residual soap used in cleaning or chemical additives used in liner manufacturing rather than the hypoallergenic silicone material itself. If continued use of gel-type liners is needed, it is reasonable to switch to a different manufacturer or change to a different liner base material (e.g., silicone to urethane) to empirically attempt to resolve the rash. Neoprene, resins used in socket fabrication, and dyes and tanning agents used in leather are also potential allergens. If empiric trials of different materials cannot identify the allergen, patch testing on the forearm may be helpful.

The use of gel sheaths, liners, and suspension sleeves can create a moist warm environment that may contribute to contact dermatitis, bacterial, and fungal skin infections. Attention to proper skin care and liner cleansing can minimize skin disorders. At times, wearing natural fiber socks or a nylon sheath between the liner and residual limb can absorb or wick away moisture and reduce skin problems.

**Pediatric Limb Deficiency or Amputation**

When amputations are performed in children for disease, tumor, trauma, or congenital skeletal deficiency for prosthetic fitting, a disarticulation-level amputation is preferred rather than an amputation through a long bone when the resulting level of function with a prosthesis will be similar. Approximately 12% of children with acquired amputations experience a condition known as bony overgrowth. Bony overgrowth is the appositional deposition of bone at the end of the amputated long bone. This bone growth results in a spike-like formation at the end of the bone that has a thin cortex and no medullary canal. The bone frequently grows faster than the overlying skin and soft tissues; a bursa may develop over the sharp end, or the bone actually may protrude through the skin with subsequent development of cellulitis and osteomyelitis. Overgrowth is seen most frequently in the humerus, fibula, tibia, and femur. It has been reported in the congenital limb deficiencies, but rarely. Several treatment approaches have been advocated for the management of this problem. Success is limited, but the technique proposed by Marquardt in which the distal end of the bone is capped with a cartilage epiphysis is the best of the surgical options available to manage this problem (15).

For the child with a congenital skeletal deficiency, the initial prosthesis for the UE usually is fitted when the child has attained independent sitting balance, or at approximately 6 months of age (96–99) (Table 74-8). For the LE, the initial prosthesis is fitted when the child begins to pull to a stand, which generally is between 9 and 14 months. Young children and infants usually learn to use their prostheses by incorporating them as part of play activities rather than through specific exercises. Prosthetic training periods for children may last only for several minutes at a time because of limited attention span, and they may require much longer periods of free play interspersed between actual training sessions. It is important that parents be instructed in techniques to help their children attain the necessary prosthetic skills because much of the training in the use of the prosthesis will occur in the home rather than in the clinic. It is also important to understand when working

<table>
<thead>
<tr>
<th>Amputation Level</th>
<th>Age for Prosthetic Fitting</th>
<th>Developmental Milestones</th>
<th>Prosthetic Prescription</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transradial</td>
<td>6–7 mo</td>
<td>Sitting balance, reaches across mid-line for bimanual object manipulation</td>
<td>Body-powered—passive mitt, self-suspending socket</td>
</tr>
<tr>
<td></td>
<td>9–15 mo</td>
<td></td>
<td>Externally powered for greater grip strength, single control site (voluntary opening, auto close)</td>
</tr>
<tr>
<td></td>
<td>24–36 mo</td>
<td></td>
<td>Change to two site control for voluntary opening and closing</td>
</tr>
<tr>
<td>Transhumeral</td>
<td>6–7 mo</td>
<td>Same as for transradial</td>
<td>Body-powered—passive mitt and elbow, activate elbow at 18 mo.</td>
</tr>
<tr>
<td></td>
<td>24–48 mo</td>
<td></td>
<td>Externally powered terminal device, when terminal device control mastered, activate the elbow.</td>
</tr>
<tr>
<td>Transtibial</td>
<td>9–12 mo</td>
<td>Child pulls to stand</td>
<td>PTB, supracondylar strap</td>
</tr>
<tr>
<td>Transfemoral</td>
<td>9–12 mo</td>
<td>Child pulls to stand</td>
<td>ICS, belt suspension, no knee unit until ages 3–4.</td>
</tr>
</tbody>
</table>
with children who have a limb deficiency or amputation of the UE that the prosthesis becomes an aid rather than a replacement (Table 74-9). If the child cannot habilitate to the prosthesis then it will be discarded.

The age at which to fit children with a myoelectric prosthesis is a controversial and complex issue beyond the scope of this chapter. This subject is reviewed in detail elsewhere (12,96,97).

### TABLE 74.9 Comparison of Children Versus Adults with an Amputation (15,99–101)

<table>
<thead>
<tr>
<th>General</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Dynamic (growing)</td>
<td>vs.</td>
</tr>
<tr>
<td>Adynamic (decelerates—aging)</td>
<td></td>
</tr>
<tr>
<td>Dependent</td>
<td>vs.</td>
</tr>
<tr>
<td>Independent</td>
<td></td>
</tr>
<tr>
<td>Untrained (life disciplines)</td>
<td>vs.</td>
</tr>
<tr>
<td>Trained</td>
<td></td>
</tr>
<tr>
<td>Irresponsible</td>
<td>vs.</td>
</tr>
<tr>
<td>Responsible</td>
<td></td>
</tr>
<tr>
<td>Malleable</td>
<td>vs.</td>
</tr>
<tr>
<td>Less malleable</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Physical</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Growing</td>
<td>vs.</td>
</tr>
<tr>
<td>Static</td>
<td></td>
</tr>
<tr>
<td>Immature</td>
<td>vs.</td>
</tr>
<tr>
<td>Mature</td>
<td></td>
</tr>
<tr>
<td>Longitudinal growth</td>
<td>vs.</td>
</tr>
<tr>
<td>Nonlongitudinal growth</td>
<td></td>
</tr>
<tr>
<td>Circumferential growth</td>
<td>vs.</td>
</tr>
<tr>
<td>Circumferential growth (dietary)</td>
<td></td>
</tr>
<tr>
<td>Circulation and tissue tolerance are ideal</td>
<td>vs.</td>
</tr>
<tr>
<td>Circulation and tolerance vary with age and health</td>
<td></td>
</tr>
<tr>
<td>Influences surgical indications, site of amputation and goals of training</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Social</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Member of family group</td>
<td>vs.</td>
</tr>
<tr>
<td>Independent person</td>
<td></td>
</tr>
<tr>
<td>Few independent social responsibilities</td>
<td>vs.</td>
</tr>
<tr>
<td>Variable responsibilities; depends on age, marriage, parenthood, etc.</td>
<td></td>
</tr>
<tr>
<td>Adjustment relatively easy</td>
<td>vs.</td>
</tr>
<tr>
<td>Adjustment less easy because of fixed social environment</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Economic</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>By family</td>
<td>vs.</td>
</tr>
<tr>
<td>By patient</td>
<td></td>
</tr>
<tr>
<td>Not self-supporting</td>
<td>vs.</td>
</tr>
<tr>
<td>Self-supporting or at least contributes to the economic welfare of the family</td>
<td></td>
</tr>
<tr>
<td>Amputation not of economic importance</td>
<td>vs.</td>
</tr>
<tr>
<td>Amputation may interfere with established economic status</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Education</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Process of obtaining basic education</td>
<td>vs.</td>
</tr>
<tr>
<td>Usually completed</td>
<td></td>
</tr>
<tr>
<td>Advanced education can be planned to include handicap and its limitation and needs</td>
<td>vs.</td>
</tr>
<tr>
<td>Age often makes long reeducation and training difficult, if not impossible</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Vocational</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Not selected or established</td>
<td>vs.</td>
</tr>
<tr>
<td>Established</td>
<td></td>
</tr>
<tr>
<td>Oriented around handicap</td>
<td>vs.</td>
</tr>
<tr>
<td>Must reorient vocationally because of handicap</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Psychological</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Because of immaturity of development, may not have the profound changes sometimes seen in the adult. Usually reflects family (parental) reaction to the amputation or deformity. In general, is not a great problem in a stable family situation.</td>
<td>vs.</td>
</tr>
<tr>
<td>Great variation. All the way from profound psychoneurosis to mature, reasonable acceptance of the disability. The impact of the amputation on the socio-economic areas of the patient’s existence are generally profound.</td>
<td></td>
</tr>
</tbody>
</table>
PROSTHETIC PRESCRIPTION EXAMPLE CASES

The prosthetic prescriptions presented for these cases are as examples only. They do not represent the standard or typical prosthetic prescriptions for these levels of amputation. We want to be clear that a specific prosthetic prescription must be tailored to meet the specific needs of an individual with an amputation. The examples presented here serve to highlight the decision process that might be followed to arrive at an appropriate prosthetic prescription.

Case 74-1: Transradial Amputation

A 24-year-old, right-handed man sustains a work-related crush injury to his right hand, resulting in a long transradial level of amputation. He plans to return to work operating a drill press.

Possible prosthetic prescriptions include:

- **Body power**
  1. Double-wall plastic laminate socket
  2. Quick-change locking wrist unit
  3. No. 7 (heavy duty, “Farmer’s hook”) terminal device
  4. Flexible elbow hinges
  5. Triceps pad
  6. Figure-8 harness for suspension
  7. Bowden single-control cable

- **External power**
  1. Double-wall plastic laminate socket with self-suspending design
  2. Otto Bock Greifer (myoelectric hook) terminal device

In this person, body power will be lightest in weight, most durable, and least expensive. If more than 6 to 7 lb (2.7 to 3.2 kg) of pinch force is necessary for functional activities, the Greifer will provide up to 35 lb of pinch force.

Case 74-2: Transhumeral Amputation

A 35-year-old, right-handed female homemaker sustains a short transhumeral level of amputation following a motor vehicle accident.

Possible prosthetic prescriptions include:

- **Body power**
  1. Double-wall plastic laminate socket
  2. Constant-friction wrist unit
  3. No. 5XA (lightest weight) terminal device
  4. Internal, alternating locking elbow with turntable
  5. Figure-8 harness
  6. Bowden double-control cable

- **External power**
  1. Double-wall plastic laminate socket with self-suspending design
  2. Otto Bock Greifer (myoelectric hook) terminal device

In this woman, because of the short residual limb, external power may be more comfortable and functional but will be heavier and much more expensive than body power.

Case 74-3: Transtibial Amputation

A 72-year-old, retired man with type II diabetes and PVD has a transtibial amputation for an infected nonhealing ulcer and gangrenous foot. Possible prosthetic prescriptions include:

- **Provisional**
  1. Total contact PTB thermoplastic socket
  2. Foam liner (soft insert)
  3. Neoprene sleeve suspension
  4. Lightweight alignable shank
  5. SACH foot

- **Definitive**
  1. Exoskeletal design prosthesis
  2. Total-contact laminated PTB socket
  3. Silicone suction suspension (3-S)
  4. Lightweight multiaxial foot

The provisional prosthesis is a lightweight design that provides a stable support base on which to learn to walk with a prosthesis. The soft liner will make modifications for changes in residual limb volume that are expected to be easier to accomplish. An exoskeletal design with multiaxis foot was chosen for the definitive prosthesis in consideration of the individual’s desire to return to his gardening activities, which required a prosthesis that was more durable and stable on uneven ground with a secure suspension system that would not be torn up when kneeling with the prosthesis.

Case 74-4: Transfemoral Amputation

A 28-year-old female day care teacher sustained an open comminuted distal femur fracture while mountain climbing and ultimately had a mid-thigh level transfemoral amputation after developing osteomyelitis. Possible prosthetic prescriptions include:

- **Provisional**
  1. Total contact thermoplastic ICS
  2. TES belt suspension
  3. Hydraulic knee unit
  4. Lightweight dynamic-response foot
  5. Cosmetic foam cover

- **Definitive**
  1. Total contact carbon fiber, reinforced ischial containment suction frame socket with thermoflex liner
  2. Thigh rotator
  3. Swing and stance phase-control hydraulic knee unit
  4. Split-toe flex-foot
  5. Cosmetic foam cover

The provisional prosthesis, an endoskeletal design with a non-suction socket, belt suspension, and hydraulic knee unit, was chosen to allow easy accommodation for anticipated major changes in residual limb volume that were expected to occur quickly with prosthetic use while at the same time recognizing
this individual’s high level of physical activity. The cosmetic cover was added to the provisional prosthesis, not usually done, recognizing her work with small children and her desire not to scare them with the prosthesis. The changes in the definitive prosthesis reflected the individual’s desire to eliminate the belt suspension, achieve a more secure suspension, and accommodate her recreational and competitive sports activities. The thigh rotator was added so that she could sit on the floor and work with the children in her class.

IN MEMORY OF ANDREW J. GITTER, MD (1957 TO 2003)

Dr. Andrew J. Gitter was the initial author of this chapter. Due to his untimely death, he was unable to participate in this revision. The core ideas and proper drawings of prosthetic devices are his legacy to his students and colleagues. Dr. Gitter was appointed as an associate professor at the University of Texas Health Science Center in rehabilitation medicine on February 7, 1997, having previously served on the faculty of the University of Washington Medical School. Dr. Gitter was awarded tenure in September 2001.

Dr. Gitter received his B.S., magna cum laude, in computer engineering from Rochester University in 1979, and he received his M.D., cum laude, from the University of Michigan Medical School in 1980. After completing his residency and an NIDRR Research Fellowship at the Department of Rehabilitation Medicine at the University of Washington, he joined their faculty.

Dr. Gitter was the University of Washington, Department of Rehabilitation Medicine Faculty Teacher of the year in 1991, and in 1999 he was honored as The University of Texas Health Science Center at San Antonio, Department of Rehabilitation Medicine Faculty Teacher of the year. He received the American Board of PM&R Earl C. Elkins Award in 1991 for the highest written board exam score. In 1992, Dr. Gitter won the PM&R Education and Research Foundation Award for best scientific paper published by a physiatrist in practice for less than 5 years, and in 1996 he won the PM&R Education and Research Foundation award for best scientific paper published by a physiatrist in practice for more than 5 years.

Dr. Gitter was the Chief of Rehabilitation Medicine at the Audie L. Murphy Memorial Veterans Hospital and he also served on the staff of Reeves Rehabilitation Center at University Hospital. Dr. Gitter was actively involved with prosthetic research and he was known worldwide for his expertise in gait analysis. He established the first Gait Analysis Laboratory at the Veterans Hospital. Dr. Gitter had an active medical practice and he was particularly revered by his patients from the amputee clinic.

Dr. Gitter is survived by his wife Brenda, and daughters Anna Caitlin, Maria Sofya Catherine, and Yulia Katrina; parents, Richard and Elizabeth Gitter; in-laws, Harlan and Lorena Green; brothers and their wives, Joseph and Linda Gitter, Theodore and Nancy Gitter, Thomas and Karen Gitter; and sisters and their husbands, Diane and Thomas Mueller, Carol and Kenneth Casolari.

Dr. Gitter was the consummate physician, teacher, researcher, husband, father, and friend. He will be missed by all those who respected and loved him.

INTERNET RESOURCES

Company
- Animated Prosthetics www.animatedprosthetics.com
- Becker Orthopedics www.beckerorthopedic.com
- CPI www.college-park.com
- Daw Industries www.daw-usa.com
- Endolite www.endolite.com
- Evolution Liners www.evolutionliners.com
- Fillauer www.fillauer.com
- Freedom Innovations www.freedom-innovations.com
- Hosmer www.hosmer.com
- Kingsley www.kingsleymfg.com
- Liberating Technologies www.liberatingtechnologies.com
- Living Skin www.livingskin.com
- Motion Control www.uraahrm.com
- Ohio Willow Wood www.owwco.com
- Ossur www.ossur.com
- Otto Bock www.ortobockus.com

Amputee Athletes
- Active Amp.org www.activeamp.org
- American Amputee Soccer Assoc. www.ampsoccer.org
- Challenged Athletes Foundation www.challengedathletes.org/caf/
- Disabled Sports USA www.dsusa.org
- International Paralympics Committee www.paralympic.org

Other
- American Academy of Orthotists and Prosthetists www.oandp.org
- Amputee Coalition of America www.amputee-coalition.org
- Amputee Resource Foundation of America, Inc www.amputeeresource.org
- Barr Foundation www.oandp.com/resources/organizations/barr
- International Society for Prosthetics and Orthotics www.i-s-p-o.org
- National Limb Loss Information Center www.amputee-coalition.org/nllc_about.html
REFERENCES


Upper Limb Orthotics

This chapter will enlighten and educate the practicing clinician on the prescription, fabrication, and use of upper limb orthoses. Historically, upper limb orthotic prescription has been complicated by a lack of universal nomenclature, and a variety of methods used to obtain orthoses (e.g., custom-made by an occupational therapist or a certified orthotist, or simply ordered direct from a catalog and fitted to a patient by any medical professional). This chapter will help clarify the common uses and designs of upper limb orthoses, and promote the use of universal terminology for the prescription of these devices.

DEFINITIONS AND NOMENCLATURE

The term orthosis refers to a device that is applied externally to the body to support or improve the function of that body segment. Over the past decade, there has been a strong effort to name all orthoses by the joints they cross, and then specifying any design features related to those joints. Unfortunately, over the past 50 years, upper limb orthoses have been named after the designer or the facility at which they were developed. Better universal terminology would simply include the five common joints of the upper limb including finger, hand, wrist, elbow, and shoulder. The options or control features at each joint could simply be described in a fashion similar to that currently used in lower limb orthotics. The four common options for a joint include fixing the joint in one position, flexion/extension block, flexion/extension traction, and flexion/extension assist. The term “block” refers to a limitation in the range of motion of the joint created by the orthosis. The term “traction” refers to an external force applied across a joint for the purpose of stretching soft tissues (e.g., contracture). The term “assist” refers to an external force applied across a joint to substitute for weak muscles. This is certainly easy to understand as it relates to finger, metacarpophalangeal (MCP) joint, and elbow where we primarily have a hinged joint design. Additional design features will need to describe thumb positioning due for positioning. Fortunately, the vast majority of upper limb orthoses incorporating the shoulder are designed to fix the shoulder in a predetermined position following surgery or trauma. There are upper limb orthoses which stabilize a bony segment following fracture or surgery, but do not cross any joint. These orthoses can be named by the segment involved. All the features described here can be incorporated into a prescription template as shown in Figure 75-1.

All upper limb orthoses can be classified into one of three categories: static, dynamic, or hybrid. Static orthoses have no movable joints incorporated into the design. However, a static orthosis may allow active joint motion in one direction, but block motion in another direction (static with block). A static orthosis may also be changed or adjusted to alter the motion allowed or alter the pressure across a joint for stretching purposes (progressive static).

Dynamic orthoses have movable joints that can limit motion (block), increase motion through traction, or substitute for weak muscles using supplemental force (assist).

Hybrid orthoses will incorporate features of both static and dynamic orthoses into one device. This type of orthosis will cross multiple joints with the intention of limiting or stopping motion at some joints, but allowing or augmenting motion at other joints.

Therefore, a proper prescription for an upper limb orthotic device should include the simple definition of the basic platform by the joints or segments that are involved, followed by special design features controlling each of those joints. The prescription should also include materials for each of these components and any special padding indicated. As with any good orthotic prescription, the diagnosis, disability, prognosis, and duration of need should be indicated on the prescription. There should always be communication between the prescribing physician and the professional who fabricates or fits the upper extremity orthotic device to ensure clear understanding of the goals of the device.

UPPER LIMB ORTHOTIC GOALS

The primary purpose of all upper limb orthoses, and the rehabilitation program related to their prescription, is to regain or to preserve prehension of the hand. The upper limb is vastly different from the lower limb because of the unique and critical functioning of the hand. The lower limb simply maintains our
body weight as we move through space, and lower limb orthoses primarily prevent us from falling. Conversely, it is the role of the shoulder, elbow, and wrist to position the hand properly in space to provide the essential function of gross motor grasp and fine motor skills. It is the preservation or restoration of this hand function that we strive for with upper limb orthotics. With this in mind, there are five common goals for upper limb orthotics:

1. Substitute for weak or absent muscles
2. Protect damaged or diseased segments by limiting load or motion
3. Prevention of deformity
4. Correction of contracture
5. Attachment of other assistive devices

The first goal is a very common indication for upper limb orthotics. In this case, we substitute for weak or absent muscles at the wrist, elbow, or shoulder that fail to properly position the hand, or we substitute for weak musculature within the hand itself that fails to provide proper prehension. Common clinical examples would include cervical spinal injury, brachial plexus injury, or peripheral nerve injury to the median, ulnar, or radial nerves.

The second goal is to protect damaged or diseased segments such as those commonly seen in cases of surgical repairs, trauma, rheumatoid arthritis, etc. With trauma or surgical repair, the orthosis is designed to control loading across damaged bony segments or strained/strained soft tissues, to promote proper healing. This goal is commonly achieved through a series of progressive static or dynamic orthoses, each allowing increasing loads or movements across a joint. In the case of rheumatoid arthritis and other progressive diseases, such as scleroderma, loads or motions are controlled with the goal of slowing the progression or natural course of the disease. In these conditions, orthoses are also used as an adjunct in pain control when inflamed joints must be temporarily immobilized and then slowly progressed over time.

The third goal of upper extremity orthoses is the prevention of deformity. There are many clinical situations of upper and lower motor neuron disease or injury (brain injury, stroke, spinal cord injury, brachial plexus injury, peripheral nerve injury) where proper positioning of the upper extremity is critical to prevent contracture or deformity. In these cases, there is no actual disease or injury to the segment included in the orthosis, but there has been proximal neurological injury creating the risk of deformity or contracture.

The fourth goal of upper limb orthoses is the correction of contracture, which may have occurred as a result of disease or immobilization. Very commonly, clinicians are forced to immobilize upper limb segments or joints following a fracture or other significant injury to promote soft-tissue and bony healing. Subsequently, normal range of motion at these joints must be regained through progressive stretching orthoses. This can be achieved through a series of progressive static or dynamic orthoses that are modified on a regular basis. The aggressiveness of the orthotic treatment program is determined by the degree and duration of contracture, and appropriate identification of the soft tissue involved.

The fifth goal of upper limb orthoses is to provide a base for attachment of other assistive devices. The simplest example of
this would be a universal cuff, which wraps around the hand to provide positioning of a spoon, fork, or other eating utensils. Universal cuffs have also been commonly used for writing or keyboard devices or other grooming devices. In contrast, the clinician should recognize that the balanced forearm orthosis (BFO) is simply a wheelchair-mounted accessory, which then attaches to a forearm trough. The BFO provides enormous assistance to the patient with spinal cord injury to carry out activities of daily living (ADLs), such as feeding, grooming, and even access to household technologies (light switch, telephone, television remote control).

BIOMECHANICAL PRINCIPLES

There are five critical concepts that must be understood to appreciate the proper design and fabrication of upper limb orthoses. Failure to understand and integrate these principles into the upper limb orthosis may produce less than optimum functional outcome or even injury to the patient. The five principles include:

1. Three-point control concept
2. Tissue tolerance to compression and shear forces
3. The biomechanics of levers and forces
4. Selection of materials
5. Static versus dynamic control

The three-point control concept is the basis of nearly all upper and lower limb orthotic designs (Fig. 75-2). Generally, a strong force is applied at a joint and a counterforce applied proximal and distal to that joint. The location of the force and counterforce may be clearly identified as a specific loop or bar in the orthotic design. However, in many instances the force and counterforces may be obscurely hidden into the design of the orthotic device. The precise point of application and the magnitude of the force and counterforce are critical to achieve the goal of controlling that joint.

The tissue tolerance to both compressive and shear forces must be understood if the orthosis is to be designed and fabricated safely. There are more than 30 pressure-sensitive bony prominences in the wrist, hand, and fingers alone. Avoiding prolonged, excessive pressure over these bony prominences will preserve skin integrity and patient comfort. Pressure duration curves adopted from other rehabilitation fields (such as seating systems) have indicated that higher pressures of 100 to 300 mm Hg are tolerated for only 2 to 4 hours continuously, whereas lower pressures of 20 to 50 mm Hg are tolerated for up to 12 hours a day continuously. If high pressures are anticipated at a joint or bony prominence, then a proper wearing schedule should clearly delineate the duration of time and the frequency of use throughout the day or night. Although a longer stretching time will reduce a contracture more rapidly, the risk of skin breakdown increases steadily with wearing duration. The clinician should also remember that distribution of pressure over a larger surface area is better tolerated than a small focused pressure. It is also important to remember the natural position of the transverse and longitudinal arches of the hand, especially in static positioning orthoses.

The biomechanical forces applied in upper limb orthotics are most analogous to a class one lever, commonly known as a seesaw. Most individuals recognize that a seesaw has a force applied at both ends with a central fulcrum. Rehabilitation specialists should also understand that a relationship exists between the magnitude of the pressure applied (force) and the distance to the fulcrum (lever). In essence, the farther from the fulcrum or from the joint, the less pressure needed to generate a fixed force across the joint. If tissue tolerance is a concern due to the required magnitude of pressure, one can simply move the point of pressure application farther from the joint to decrease the magnitude of the applied pressure. In other words, by increasing the lever distance, less pressure can generate the same force across the joint (Fig. 75-3). The angle of pull of force is also very important to produce effective control of the identified joint, and to prevent damage to collateral ligaments. If the angle of pull is not perpendicular to the joint, the orthosis may overstretch collateral ligaments and create mediolateral instability at the joint. As an example, stretching a flexion contracture across the MCP joint or interphalangeal (IP) joint requires a direction of pull perpendicular to the finger. If the pull is at an
angle to the finger, then less effective stretching of the flexor tendon is achieved and unwanted stretching of collateral ligaments may occur (Fig. 75-4). The clinician must also understand that forces will impact the design of the device and help determine the materials that will be used to provide structural stability to the orthosis. For example, a long outrigger attached to a small platform on the dorsum of the hand may provide an inadequate base of support and lead to bending or deformity of the outrigger or the orthosis. A more stable outrigger and longer base of support will provide good biomechanical stability to the orthosis and provide the adequate stretch necessary.

The selection of materials for upper limb orthoses depends on the flexibility, strength, and durability of the material necessary to achieve the proper outcome. Most upper limb orthoses are fabricated of thermoplastics for the structural design. Low-temperature thermoplastics are commonly used because they can be easily shaped or formed to the patient’s limb without the need for high-temperature ovens. They can also be easily modified by a common heat gun. However, these devices may also deform under exposure to common heat sources such as the sun or radiators. High-temperature thermoplastics may be indicated for certain longer-term or high-stress devices. Various metals are still used for parts of upper limb orthoses where a lightweight, strong, and compact design is indicated. This includes aluminum frames, joints, and spring wire materials. Ultimately, carbon fiber materials can be used for very strong and very light designs, but manufacturing these orthoses requires specialized equipment and training. Foam materials are commonly used as padding to improve tissue tolerance in high-pressure areas.

Finally, the overall biomechanics of any upper limb orthosis can be defined as static, dynamic, or hybrid. A traditional static orthosis will simply stabilize or fix one or multiple joints. Generally, a static positioning splint for a flaccid limb will create very low tissue pressures and very low forces at joints. This is tolerated well and can be worn nearly continuously without concern of skin breakdown. However, static positioning splints for an upper limb with an increased tone will certainly generate much higher tissue pressures and the design should incorporate additional padding in these situations. A wearing schedule is also much more important when increased tone is encountered. A dynamic orthosis allows or enhances movement across a joint. Dynamic orthoses generally have fairly low tissue pressures unless external forces are applied such as stretching of a contracture, or in the case of an abnormal increased tone from upper motor neuron injury. In cases of a dynamic orthosis with an external force applied, the biomechanical stresses within the orthosis and the tissue compression will be high.

**ANATOMICAL PRINCIPLES**

Proper upper limb positioning requires an understanding of multiple anatomical issues, particularly when a joint is to be immobilized. The wrist should be immobilized in slight extension and neutral pronation/supination. This position facilitates hand prehension activities to reach the face and midline trunk for ADLs.

The IP joints of the fingers should be immobilized in extension, but the MCP joints should be immobilized in flexion to maintain the length of the collateral ligaments. The IP and MCP joints should be mobilized as soon as possible to prevent contracture and adhesions of the long flexor and extensor tendons.

The thumb should be immobilized opposite the fingers in palmar abduction and extension. The web space should be maximized to maintain both gross grasp and fine motor pinch (Fig. 75-5).

The hand itself has two transverse arches (proximal and distal metacarpals) with two different radii. These arches must be preserved to maintain proper finger positioning. As each finger is flexed individually, its fingertip points to the scaphoid bone, because of the distal arch. An orthosis that provides traction in finger flexion must also follow this same angle to the

![FIGURE 75-4. Traction angle. The angle of pull is correct on the index finger, but incorrect on the middle finger, because it is not perpendicular to the MCP joint.](image)

![FIGURE 75-5. Proper positioning of the wrist, hand, and finger when immobilized. (Courtesy of North Coast Medical.)](image)
scaphoid. As mentioned earlier, traction across a finger or any segment should be perpendicular to that segment and follow the anatomical angle of the joint involved.

**COMMON UPPER LIMB ORTHOTIC DESIGNS**

**Finger and Thumb Orthoses**

Simple static orthoses for the fingers or thumb are commonly used to treat fractures, collateral ligament sprains, and burns of the digits. These can be partial or complete circumferential designs providing both flexion-extension control and mediolateral stability across the IP joints. Static positioning of the IP joints should be in extension to maintain the full functional length of collateral ligaments. Static finger orthoses with a flexion or extension block allow motion in one direction, but not the other. The best example of this is the finger ring orthoses, commonly used in rheumatoid arthritis (Fig. 75-6). The boutonniere’s deformity creates flexion at the proximal IP (PIP) joint and hyperextension at the distal IP (DIP) joint. This can be controlled using the ring orthoses to block flexion at the PIP joint. The swan-neck deformity causes hyperextension at the PIP and flexion at the DIP joint. The same ring orthoses can be reversed to encourage flexion and block extension at the PIP. A variety of dynamic orthoses across the IP joints of the fingers are used for the purpose of stretching flexion contracture at the IP joint (Fig. 75-7). Traction is placed across the contracted joint using spring wire or rubber bands. A progressive static orthotic program can accomplish the same result with regular adjustment or alteration of a static orthosis to stretch the contracture across an IP joint. This can be accomplished through tightening of the screw, adjusting of a thermoplastic design, or altering a Velcro strap (Fig. 75-8).

**Hand-Finger Orthosis**

The most common reason for fabricating hand-finger orthoses is to gain control of the MCP joint of the fingers or thumb. Historically, the traditional static hand-finger orthosis consisted of a short opponens orthosis fabricated from metal, which wrapped around the medial or lateral side of the hand, preserving the arch of the hand. This would then act as a platform for outriggers or additional features, which would control the thumb, the MCP joints, or the fingers. A traditional short opponens orthosis would include a C-bar to maintain the web space between the thumb and the other fingers, and an opponens bar to position the thumb opposite the fingers for gross grasp and fine motor pinch (Fig. 75-9). This is most commonly used for median nerve injuries where control of the thumb and opposition are lost. An MCP extension block can also be incorporated into this orthosis to prevent MCP hyperextension (claw-hand deformity) which occurs in both median and ulnar nerve injuries. This modification allows the remaining intrinsic hand muscles to function as effectively as possible. The same short opponens design can be replicated using plastics with a circumferential design to...
applied close to the MCP joint and opposing counterforces proximal and distal. The common “knuckle-bender orthosis” is used to stretch the extension contracture at the MCP joints when collateral ligaments have been allowed to shorten due to immobilization (Fig. 75-13). Dynamic hand-finger orthoses can also be used for more vigorous stretching of flexion contracture at the PIP joints. These orthoses would incorporate a thermoplastic circumferential platform at the hand with attachment of MCP block to prevent hyperextension, and an outrigger to create an extension stretch across the PIP joint (Fig. 75-14). Once again, it is important to remember that the direction of pull must be perpendicular to the segment and that the angle of pull must follow the joint motion. As the flexion contracture improves, the positioning of the outrigger must be adjusted to accommodate for this change. Failure to provide MCP extension block may simply create hyperextension across the MCP and failure to stretch the PIP contracture.
in neutral or flexion position. The MCP joints are blocked in flexion and the IP joints may be allowed to go to full extension. Often, flexion traction is applied across the fingers to promote a protective positioning of the finger and eliminate tension across the repair (Fig. 75-16). As healing progresses, the MCP block is eliminated to allow further motion without allowing full extension at the wrist until the repair is completely healed. The wrist-hand-finger orthosis for extensor tendon repair includes a similar but contradictory design where the wrist is positioned in a static manner at neutral or extension position with a flexion block at the MCP and IP joints of the involved fingers. Often extension traction is applied across the MCP and IP joints to alleviate tension across the repair (Fig. 75-17). As healing progresses, further flexion is allowed at the MCP and IP joints. Finally, full wrist flexion is allowed once healing is complete.

Wrist-Hand-Finger Orthosis

Today, static wrist-hand orthoses are most commonly used for the treatment of carpal tunnel syndrome. This orthosis includes wrist positioning in neutral (i.e., 0 degrees of extension) and allows full freedom of thumb and finger movement (Fig. 75-15). Carpal tunnel syndrome commonly results from the overuse of the wrist and hand causing inflammation within the carpal tunnel. Therefore, immobilization of the wrist for part or all of the day and night often helps to resolve symptoms. Other uses of static wrist-hand orthoses include traumatic wrist sprain, or wrist inflammation due to diseases such as rheumatoid arthritis. Static wrist-hand-finger orthoses are also commonly used as the first step following injury or repair to flexor or extensor tendons of the hands. A short period of immobilization is often followed by limited motion to prevent adhesions of flexor and extensor tendons, and joint contracture. With flexor tendon repair, the wrist is commonly positioned in neutral or flexion position. The MCP joints are blocked in flexion and the IP joints may be allowed to go to full extension. Often, flexion traction is applied across the fingers to promote a protective positioning of the finger and eliminate tension across the repair (Fig. 75-16). As healing progresses, the MCP block is eliminated to allow further motion without allowing full extension at the wrist until the repair is completely healed. The wrist-hand-finger orthosis for extensor tendon repair includes a similar but contradictory design where the wrist is positioned in a static manner at neutral or extension position with a flexion block at the MCP and IP joints of the involved fingers. Often extension traction is applied across the MCP and IP joints to alleviate tension across the repair site (Fig. 75-17). As healing progresses, further flexion is allowed at the MCP and IP joints. Finally, full wrist flexion is allowed once healing is complete.
Radial nerve injury also often requires a wrist-hand-finger orthoses to assist wrist and finger extension. This is accomplished with dynamic orthoses, with rubber bands or spring wire providing extension at the wrist and outriggers with rubber bands providing extension at the fingers. An outrigger or C-bar must also be incorporated to maintain the thumb in opposition (Fig. 75-18).

The tenodesis prehension orthosis or flexor hinge orthosis is a dynamic wrist-hand-finger orthosis incorporating active wrist extension movement to regain gross grasp and fine motor pinch of the thumb and fingers. This is used in C6-level quadriplegia where some wrist extension strength is maintained with little or no long finger flexor strength or intrinsic muscle strength of the hand. The Rehabilitation Institute of Chicago design incorporates a non-elastic cord, which crosses the wrist, hand, and MCP joint to facilitate three-jaw chuck pinch with active wrist extension (Fig. 75-19). The more traditional, and more elaborate, flexor hinge design uses an adjustable rigid rod that facilitates prehension. The size of the opening of the hand is adjustable through changes in the length of the rigid rod (Fig. 75-20).

Finally, simple static wrist-hand-finger orthoses can be used for positioning of the hand following stroke, brain injury, or brachial plexus injury. As mentioned previously, the MCP joints should be positioned near full flexion and the IP joints in full extension to prevent contracture of the collateral ligaments. The thumb should also be positioned opposing the fingers and maintaining the web space (see Fig. 75-5).

**Elbow Orthoses**

Dynamic elbow orthoses are commonly used for flexion or extension contracture across the elbow due to immobilization. If burn scars are involved, this should include total contact across the burn scar area for compression treatment of the

---

**FIGURE 75-17.** Dynamic wrist-hand-finger orthosis with flexion block at MCP joint and fingers, and traction at fingers.

**FIGURE 75-18.** Dynamic wrist-hand-finger orthosis with stabilization of the wrist and extension of the fingers and thumb following radial nerve injury. (Courtesy of North Coast Medical.)

**FIGURE 75-19.** Dynamic wrist-hand-finger orthosis using Rehabilitation Institute of Chicago design for prehension. (Courtesy of North Coast Medical.)

**FIGURE 75-20.** Dynamic wrist-hand-finger orthosis with traditional metal flexor hinge design.
burn. Various spring-loaded elbow orthoses are commercially available, which can be easily adjusted by the patient or therapist to steadily increase tension across the elbow as stretching progresses (Fig. 75-21). Spring-assisted elbow orthoses are also used to augment elbow flexion when weakness or paralysis of the biceps muscle exists. Static circumferential orthoses across the elbow are often used for fractures at or near the elbow. Fractures of the radius and ulna should incorporate the elbow, wrist, and hand to control pronation and supination. An adjustable lock, such as the dial lock, can be incorporated into the elbow-wrist-hand orthosis to steadily progress motion at the elbow in flexion or extension, and still control pronation and supination (Fig. 75-22). Humeral fracture management may include a circumferential orthosis across the entire humeral segment and forearm. This would initially start as a static orthosis and slowly progress movement at the elbow as healing of the humeral fracture occurs.

**Shoulder Orthoses**

Flexible arm slings of various sorts have been used for a variety of problems including clavicular fracture, acromioclavicular joint injury, proximal humeral fracture, and glenohumeral subluxation due to hemiparesis. These do not clearly fit into the categorization of static or dynamic design due to their inherit lack of identifiable joints or rigid design features. These orthoses commonly consist of fabrics and straps that encompass the mid- to proximal humerus, and then extend across the shoulder anteriorly and posteriorly to the opposite axilla (Fig. 75-23). The primary goal is to maintain glenohumeral integrity and to limit motion across the AC and glenohumeral joints. A true static shoulder-elbow-wrist-hand orthosis, such as the airplane splint or arm abduction orthosis, can prevent movement across the glenohumeral joint by stabilizing the arm in abduction (Fig. 75-24). This is done for shoulder dislocation, shoulder surgery (e.g., rotator cuff repair), and some proximal fractures. Care must be taken to mobilize the glenohumeral joint as soon as possible to prevent adhesive capsulitis.

Other specialized orthoses in this category include the BFO for patients with C5-level spinal injury. This device...
includes a forearm trough which is suspended on a series of brackets and swivels mounted on the wheelchair. This eliminates the weight of the arm and allows the patient to use elbow flexion and shoulder adduction/abduction for feeding and limited ADLs. A universal cuff or hand orthosis is used to attach a swivel spoon or other device, which would then allow independent feeding by the patient (Fig. 75-25). Finally, a combination of prosthetic and orthotic components can be used for the patient with brachial plexus injury to create a shoulder-elbow-wrist-hand orthosis to regain some limited prehension at the hand once the elbow is locked and stabilized. A dual control cable to a figure-8 harness can use bicipital abduction to position the elbow and activate a prehension orthosis at the wrist and hand when there is complete paralysis at the elbow and hand. If the hand is not usable due to contracture, then a prosthetic hook can be applied at the palmar surface of the hand to obtain prehension (Fig. 75-26). The elbow can be locked in place using an additional strap which is activated with shoulder movements or by reaching across with the opposite hand. The paralyzed limb can then be used as a helper or stabilizer for the remaining functional arm.

**EMERGING TECHNOLOGY**

The concept of neural integration has recently been incorporated into upper limb orthoses. Commercially available devices are now being used to create prehension or gross grasp in the paralyzed limb. These devices detect a myoelectric signal from an intact proximal muscle, and then send a signal to distal paralyzed muscles. The paralyzed muscles are stimulated to contract via surface-based electrical stimulation. Stimulating electrodes over the finger flexor and extensor muscle groups create grasp and release maneuvers. The orthotic device itself supports the wrist as a wrist-hand orthosis.

The usefulness of these devices is variable, and is dependent upon physical factors such as muscle tone, target muscle excitability, and the presence of contractures. Good strength at the shoulder and elbow greatly enhances the functionality of these devices.

As can be seen, most common upper limb orthotic designs can be appropriately described using universal terminology and a prescription template, rather than abstract names. Any new or unique design features or materials can be added to this same template. Communication between the prescribing physician, patient, and therapist/orthotist will ensure fabrication and fit of a proper upper limb orthosis. Adequate instruction to the patient, and appropriate therapy and follow-up will help to prevent complications and achieve a favorable outcome.

**SUGGESTED READINGS**


Lower limb orthoses, with their many variations, are among the most commonly prescribed biomechanical devices intended to assist walking in individuals with gait dysfunction, caused by musculoskeletal or neuromuscular diseases (1). Walking is one of the most desirable self-reported goals for patients entering a rehabilitation program and the preponderance of lower limb orthotic devices in clinical use reflects this desire and the fact that present orthotic technology often does a reasonable job of restoring gross physical capacity for level walking and standing.

**ROLE AND USE OF ORTHOTICS**

Orthoses are devices applied to the external surface of the body to achieve one or more of the following: relieve pain, immobilize musculoskeletal segments, prevent or correct deformity, and improve function. Orthoses provide a direct support component to the braced limb segment and limit the range of motion of a joint. This mechanism largely gives rise to reduction of axial loads and thus perhaps aids in joint pain relief. It may also be used to partially or fully immobilize the joint, and in this manner produce deformity prevention or correction and also improve function. In conjunction with the above two mechanisms, a more difficult mechanism to visualize is the fact that orthoses modify the total static and dynamic force/moment distributions in the braced and more distal segments ("in" joints or about a joint center) and provide a substitute power source for weak muscles while maintaining level surface walking safety and efficiency.

Overall, orthotics can be divided into two major categories, namely, corrective and accommodative devices. In the near future, orthotic technology will also provide artificial strength. Corrective devices are meant to improve the position of the limb segment, either by stretching a contracture or correcting the alignment of skeletal structures (1). Accommodative devices are meant to provide additional support to an already deformed tissue, to prevent further deformity, and ultimately to improve function. Orthoses can be further classified as static (resting) or dynamic, the latter permitting movement of the involved joint(s) while controlling the direction or alignment of the movement and, at times, providing a substitute power source for weak muscles. The spectrum of orthotic devices available is quite broad, ranging from a simple plastic device applied across one joint to a much more complex device made of a variety of materials and crossing multiple joints (2), some also include control microprocessors. Knowledge of the disease or disorder being addressed with an orthosis, functional anatomy, biomechanics, orthotic components, materials, and, finally, recognition of the anticipated functional outcome are essential for proper orthotic prescription (3,4).

**TERMINOLOGY**

In the past, the lexicon of terms used to describe orthoses was very confusing; often, clinicians used different terms to describe even the most basic device. Orthoses or parts of devices were given names that might describe their purpose, the body part to which they were applied, the inventor of the device, or the town or institution where they were developed. To facilitate communication and minimize the use of acronyms, a logical, easy-to-use system of standard universal terminology for orthotic devices was developed in the early 1970s by the International Committee on Prosthetics and Orthotics (3). Orthotic devices are named by the joints they encompass in correct sequence followed by the word orthosis. For example, an orthosis that crosses the ankle and the foot is named an ankle-foot orthosis (AFO). One that crosses the knee, ankle, and foot is called a knee-ankle-foot orthosis (KAFO). The intended biomechanical function or material may complement the terminology (i.e., dorsiflexion assistance plastic AFO). More elaborate names or eponyms are typically given to such devices (3), but such language should be avoided whenever possible.

**MATERIALS**

The orthotic components chosen depend on the functions they fulfill, but most orthoses consist of three basic elements: interface components, structural components and if intended to be articulated, the type of joints. In orthoses of newer design, it may be impossible to differentiate the joints from the structural and interface components, for example, plastic AFO (2).
The choice of materials in the fabrication of orthotics is expanding rapidly, and a full review of them is beyond the scope of this chapter. The reader is encouraged to expand knowledge on the subject by reading elsewhere about this topic (5). The two primary material categories are plastics and metals. Plastics have the benefits of light weight, total contact, adjustable flexibility, the ability to adjust in shape, and a more cosmetic appearance. Metal orthoses often have the advantage of increased durability and, in the hands of a skilled orthotist, built-in adjustability. With the advent of new products such as carbon graphite and extruded plastic materials, the advantages of maximal tension strength, lightweight design, and ease of adjustability are combined (6). Factors that dictate the type of material used in orthotic fabrication include the length of time the orthosis will be used, the amount of forces applied across the orthosis, and the amount of axial weight loading. A complete orthotic prescription should include the joints it encompasses and suggest the desired biomechanical alignment and materials of fabrication and materials to use for closures (e.g., leather straps, Velcro, etc.), and where they should be attached (Fig. 76-1).

When the orthosis is ready, it should be evaluated off and on the patient to ensure proper construction, fit, and function. When these characteristics are achieved, appropriate training of the patient and family should be organized.

### MossRehab Lower Limb Orthotic Prescription

<table>
<thead>
<tr>
<th>Type:</th>
<th>Foot</th>
<th>Ankle/Foot</th>
<th>PTB</th>
<th>Knee/Ankle/Foot</th>
<th>Knee</th>
<th>Hip/Knee/Ankle/Foot</th>
</tr>
</thead>
<tbody>
<tr>
<td>Side:</td>
<td>Right</td>
<td>Left</td>
<td>Bilateral</td>
<td>Design: Articulated</td>
<td>Non Articulated</td>
<td>Floor Reaction</td>
</tr>
<tr>
<td>Construction:</td>
<td>Plastic:</td>
<td>Metal:</td>
<td>Graphite:</td>
<td>Other:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Closures:</td>
<td>Velcro</td>
<td>&quot;D Ring&quot;</td>
<td>Laces</td>
<td>Buckles</td>
<td>Elastic</td>
<td>Right hand pull</td>
</tr>
<tr>
<td>Ankle joints:</td>
<td>Gaffney</td>
<td>Gillet</td>
<td>Tamerak</td>
<td>Meridian</td>
<td>Lawrence</td>
<td>Scottis</td>
</tr>
<tr>
<td>Ankle control:</td>
<td>Plantarflexion:</td>
<td>Dorsiflexion:</td>
<td>Dacron strap</td>
<td>Inversion:</td>
<td>Eversion:</td>
<td></td>
</tr>
<tr>
<td>Resistance:</td>
<td>PLS Minimum</td>
<td>Moderate</td>
<td>Maximum</td>
<td>Assistance: Dorsiflexion</td>
<td>Extension</td>
<td></td>
</tr>
<tr>
<td>Foot plate:</td>
<td>Long</td>
<td>Longitudinal arch</td>
<td>Padded</td>
<td>Metatarsal pad</td>
<td>UCBL</td>
<td>Tone reducing</td>
</tr>
<tr>
<td>Straps:</td>
<td>Ankle</td>
<td>Mid foot</td>
<td>Forefoot</td>
<td>Inversion / eversion control</td>
<td>Valgus / varus control</td>
<td>Padded</td>
</tr>
<tr>
<td>Knee:</td>
<td>Single axis</td>
<td>Non Protuding</td>
<td>Polycentric</td>
<td>Offset</td>
<td>Unloader</td>
<td>Stance Phase Control</td>
</tr>
<tr>
<td>Knee lock:</td>
<td>Drop locks</td>
<td>Bail valves</td>
<td>Retainers</td>
<td>Spring loaded</td>
<td>Stance phase locking</td>
<td>Other:</td>
</tr>
<tr>
<td>Knee control:</td>
<td>Varus</td>
<td>Valgus</td>
<td>Flexion</td>
<td>Hyperextension</td>
<td>Infrapatellar strap</td>
<td>Knee Pad</td>
</tr>
<tr>
<td>Thigh section:</td>
<td>Traditional</td>
<td>Quadrilateral</td>
<td>Narrow ML</td>
<td>Glutea</td>
<td>Icshial WB</td>
<td>Extended lateral wall</td>
</tr>
<tr>
<td>Hip control:</td>
<td>Flexion</td>
<td>Abduction</td>
<td>Adduction</td>
<td>Silesian</td>
<td>External hip joint</td>
<td></td>
</tr>
<tr>
<td>Shoes:</td>
<td>Orthopedic</td>
<td>Canvas</td>
<td>Soft extradeepth</td>
<td>High top</td>
<td>Molded</td>
<td>Velcro</td>
</tr>
<tr>
<td>Other:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Special Instructions:**

---

**Signature** ___________________________ M.D./D.O.

**FIGURE 76-1.** Example of an orthotic prescription form.
Orthotic management is best accomplished in the majority of patients with the use of plastic molded orthoses. They provide more intimate contact with better distribution of the controlling/corrective forces over a larger surface area. These devices generally fit inside the shoe, tend to be more cosmetic, and hence better accepted. The patient is permitted to exchange shoes as long as constant heel height is maintained to avoid altering the intended dynamic alignment of the device. Plastic materials are lighter in weight and easier to maintain. The recent availability of adjustable ankle joints that can be attached to plastic orthoses has eliminated a major disadvantage of these devices when compared to metal/leather device. Lack of sensation and fluctuating edema are relative contraindications to the prescription of plastic molded devices. If the patient has adequate visual perception, cognition, and good social support, he or she can compensate for these two deficits and receive the added benefits provided by plastic braces. Two major types of plastic materials are used in orthotics: thermosetting and thermoplastic. Thermosetting plastics include formaldehyde, epoxy, and polyester resins that are typically used as laminates in a similar manner than that used for prosthetic socket fabrication. They require heat produced by a chemical reaction to harden a multilayer of weaved fibers of different products but do not soften with subsequent heating. Thermoplastics soften when they are heated, making the material moldable. Subsequent heating will soften the material for further molding, and lowering the temperature hardens the material once again. High-temperature thermoplastics, such as polyethylene, polypropylene, copolymers, ortholene, thermo plastic elastomers and vinyl polymers, require heating to temperatures greater than 150°F to mold. Fabrication of an orthosis made of any of these materials requires an exact plaster mold of the desired body part. The heated plastic is then applied to the replica for proper molding under vacuum. These materials are durable, and they have a “good elastic memory,” returning to their original position after flexible deformation. Adjustments to the biomechanical alignment and fit of an orthosis require the intervention of an orthotist. Based on the trim lines of the material, orthosis provide support and may also give a spring-action force (Fig. 76-2).

Metal/leather orthoses continue to have a definite place in the treatment of many patients such as those with fluctuating edema, soft tissue wound, hyperhydrosis or intolerance to plastics. Traditional orthotic devices use metals to provide strength and durability. Straps and padding are made of leather. The metals primarily used are steel and aluminum, mostly in alloy forms with various other metals to further increase their strength and to resist corrosion. Although metal orthoses are heavier and are cosmetically less appealing to many patients, their durability and adjustability allows them to accommodate for longitudinal growth and the changing needs of the patient (Fig. 76-3). The choice of orthotic material depends on the clinical purpose and the clinical characteristics of the patient.

Proper biomechanical alignment of any orthosis is critical to maximize the ambulatory capability of a patient. Biomechanical malalignment can and does prevent a borderline patient from becoming a functional ambulator; those patients with better recovery or less deficit may be able to compensate for inadequate orthotic alignment. In those patients for whom orthotic joint motion is required, close correlation between the orthotic and the anatomic joint centers of rotation is mandatory to avoid a discrepancy in the axis of motion which could likely produce pain, joint swelling, skin breakdown, and other preventable problems.

Adjustability of alignment is a significant advantage to current practice patterns where patients are transferred to inpatient rehabilitation programs much earlier than in the past, and the length of stay in the rehabilitation programs has decreased significantly. Predicting the final rehabilitation outcome of these patients so early in their recovery may be very difficult. The ability to adjust the biomechanical alignment of the orthosis with simple tools or to convert a controlling force into an assistive one in order to respond to the patient’s needs would be impossible with a purely corrective orthosis.
Shoes are built around a positive model or replica of the weight-bearing foot, which is called a last (11). The last, which is made of solid rock maple or plastic materials, determines the fit, walking comfort, appearance, and style of the shoe. Usually the last has a slight fore foot in-flare. Other common lasts include the broad-toe last with a straight medial border that extends from the heel to the toe; the juvenile symmetric straight last, which can be bisected into nearly equal right and left halves; and the orthopedic last with special features designed to accommodate various structural and anatomic problems (e.g., varus, valgus).

During fabrication, the insole is nailed to the last, the lining is tucked to the inner sole rim, and the reinforcements (i.e., counter, toe box) are attached. The upper of the shoe is softened by humidity for easier molding and fitted snugly to the last to conform to its every detail and then nailed or glued to the inner sole. Finally, the outer sole and heel are attached. The Goodyear welt construction of shoes is a method used in production of high-quality shoes in which the upper is sewn to the sole. This method provides a perfectly smooth inner surface, comfort, and a strong shoe that retains its shape and is easy to modify and repair. Unfortunately, these shoes tend to be bulky, heavy, and less flexible.

The upper is that part of the shoe that is above the sole. It is most commonly made of leather, although any soft and durable material may be used. Leather is found to be most comfortable because it allows evaporation and absorption of moisture and molds well to the shape of the foot. Deerskin is significantly softer than leather. It stretches more easily and is generally more accommodative, but is not nearly as durable as leather.

Upper materials are important because they affect the comfort and fit, and performance and durability of the shoe. Shoe uppers are also frequently made of man-made materials such as fabric, nylon, canvas, and plastic but in general they are more difficult to stretch or modify. Recently heat moldable uppers and other elastic plastic materials (e.g., neoprene) have been introduced to the fabrication of shoes. The material is capable of easily stretching around anatomical deformities such as bunions or hammer toes but a problem with this material is its reduced capacity to dissipate heat and moisture. They are an especially useful alternative to custom shoes in patients with severe fore foot deformities who otherwise would need costly custom shoes.

The upper consists of the vamp, quarters, and lace stay. The vamp is the anterior portion of the upper, which covers the toes and the instep. The tongue, a strip of leather lying under the laces, and the throat, the opening at the base of the tongue, are parts of the vamp. Anteriorly, the vamp has a reinforced toe box or toe cap to maintain appearance and to protect the toes against trauma. The lace stay or the portion containing the eyelets for laces is usually part of the vamp, but it may also be part of the quarters. The two quarters make up the posterior part of the shoe. The quarters are usually reinforced by the heel counter, which stabilizes the foot by supporting the calcaneus and giving structural stability to the
CHAPTER 76  |  LOWER EXTREMITY ORTHOTICS, SHOES, AND GAIT AIDS

Types and Styles

There are innumerable shoe types and styles, although basic designs are relatively few (Fig. 76-4). The basic designs are determined mainly by the shape of the upper, particularly the design of the toe and the height of the quarters. The low-quarter shoe, also known as Oxford shoe, is characterized by the quarters finishing approximately 1 in. below the malleoli with no restriction of ankle or subtalar movement. In high-quarter shoes, the quarters may extend up to barely cover the malleoli, as in the chukka shoe, or extending to the lower third of the tibia or higher, as in boots. This style prevents piston action during walking and back-and-forth sliding of the foot. In high-quarter shoes, the quarters may extend up to barely cover the malleoli, as in the chukka shoe, or extending to the lower third of the tibia or higher, as in boots. This style prevents piston action during walking and back-and-forth sliding of the foot. In addition, it provides some degree of medial-lateral stability at the ankle and subtalar joints, and some restriction of plantar/dorsiflexion ankle motion. The most common opening style is the blucher type, in which the lace stay is not directly fastened to the vamp. This style gives a wide opening for the foot for easy insertion and greater adjustability over the mid foot. For ease of access the surgical opening shoe allows exposure of the entire foot by opening up to the toes. The bal-type (Balmoral) throat, which has the face stay attached directly to the vamp, does not provide such easy foot access. Shoe closure usually is accomplished by cotton laces, which thread through two or more pairs of eyelets, although closure can also be achieved by buckles, zippers, Velcro flaps, or elastics.

The clinician needs to be aware that the height of the heel has a direct impact on the foot and ankle positions and affects the proximal joints and general posture of the trunk. When a brace is manufactured and aligned, this is done with a particular shoe heel height as part of the design (4). High heels, especially those with a tapered, narrow striking point, make the ankle and foot more unstable and thereby contribute to foot stress, ankle injuries, and falls. Some studies have linked the use of such shoes to abnormal forces in the more proximal joints (12).
modified by adding a variety of alterations including changes to the heel, the addition of metatarsal pads, rockers, and other modifications.

**Sole Materials**

Materials used in the construction of shoe soles vary greatly and have significantly different properties with regard to weight, durability, shock absorption and attenuation, flexibility, and support.

Leather soles are extremely durable and once it was the choice soling material but now it is reserved primarily for high-end dress shoes. Leather tends to be stiff and heavy, and offers little or no shock absorption when compared to other materials and can be slippery in wet conditions. Leather soles do allow for relatively simple attachment of AFOs.

Hard rubber is a great alternative to leather, especially when used to attach an AFO as it has some degree of impact absorption at heel strike and it is less slippery when used to walk in wet environments. Crepe is a cellular rubber compound that typically offers a great combination of shock absorption and traction. A shoe with a crepe sole can still have an AFO attached to it, although it is not as easy as with a leather or rubber-soled shoe to fabricate. Crepe is the most commonly used material for shoe modifications.

Vibram® is a dense microcellular rubber. It has many qualities of crepe but is more durable and can be more shock absorbing. This material can also be used with a shoe in the construction of an AFO.

Ethylene vinyl acetate (EVA) is a more recent introduction in the industrial manufacture of most types of shoe. It is shock absorbing, moldable, lightweight, and flexible and can have a multitude of colors. Sometimes this material can be difficult to use in AFO fabrication due to lack of adherence to commonly used cements. In recent years athletic shoes (Fig. 76-5A) have changed dramatically in design and materials used in their fabrication. Different sports may require different design features for optimal performance and comfort. The sole of a good athletic shoe is stiff at the heel and at the shank but very flexible at the forefoot, where it should bend easily at the ball. The outer sole usually is made of highly durable rubber compounds that provide a good grip on the ground, whereas the inner sole is designed to closely fit the contours of the foot. Between the outer and inner soles, fluid, gel or cavitated materials are placed for cushioning and shock dampening. The outer sole is often designed to flare out laterally at the heel and towards the mid foot to improve the mediolateral stability at the ankle and to ensure the foot is flat as it strikes the ground in the normal, slightly supinated position. Rising up from the rear sole to a height of approximately 0.5 in. (1.2 cm) is a heel cup, which further increases mediolateral control and provides hind foot stability. The rigid and non-compressible heel counter provides added ankle and foot stability, but for athletes with a tendency for recurrent ankle sprains, there is the option of using high-top shoes. The upper is reinforced in the mid foot area for maximum stability and to resist excessive side-to-side motion. Such reinforcement is obtained by adding stiffening bands, stabilizing lacing systems, or motion control straps.

Due to their natural increase in coefficient of friction, rubber soles have more traction than leather shoes when walking on carpeted or dry surfaces. Sometimes it is desirable to reduce friction, and application of an external leather toe slider covering the front 1/3 of the sole can be of great benefit to facilitate limb advancement when limb clearance is impaired.

**Fitting**

The first requirement of a shoe is that it fits and does not cause pain, skin problems, or deformities. Ideally both feet should be measured in length and width, and the shoes tried on both feet in case of size discrepancy, in which case the shoe size should be chosen to comfortably fit the larger foot while some filling material may be used in the smaller one. If more than a one size difference exist, using split size shoes is encouraged (14,15). Footwear preferably should be purchased at the end of the day, when the feet often are slightly swollen. When shoes are fitted, each shoe should be judged individually in a fully weight-bearing position. The shoe should fit snugly enough not to fall off but be loose enough to adapt to the size and shape of the foot,
which changes with climate, ambient temperature, time of day, body position, and standing weight bearing. Because the foot changes shape with weight bearing and some pathological conditions, shoes initially should be carefully tested for fit, both in length and width, not only by standing but by walking for several steps and stopping fast. In length, the shoes should extend at least a half-inch (1.2 cm) beyond the longest toe, usually the hallux or the second toe. The heel-to-ball distance of the foot and the shoe should be equal. Thus, the first metatarsal joint should be located at the inner curve of the shoe, and on toe dorsiflexion, the shoe should bend easily at the toe break that should run directly across the ball of the foot. The widest part of the shoe should coincide with the broadest part of the foot, leaving enough free space medial and lateral to the heads of the first and fifth metatarsal bones, respectively. The shoe widths measured at the ball are available in different sizes, ranging from A (narrow) to E (wide). Each size represents a 0.25-in. (0.6 cm) increase in width.

The transverse arch of the foot should function normally, weight should be evenly distributed, and no slidding of the fore foot within the shoe should occur. There should be no gap between medial and lateral quarters and the heel counter should close around the heel bone without bulges, allowing only a small amount of pistoning. The height of the quarters should be sufficient to hold the shoe securely on the foot. Some pistoning usually is unavoidable in a modified shoe with a rigid sole and heel counter. The height of the vamp should be adequate to prevent pressure or irritation over the toes and the instep. If the material is soft and pliable it should be possible to fold some material from the top of the toe box to confirm appropriate space for the toes. Shoes are not usually fabricated in different depths, although extra-depth shoes and shoes with removable insoles are available to accommodate foot abnormalities and shoe inserts. The real proof of fit, however, is if the shoe is comfortable after hours of continuous wear or walking.

If there is a pathological condition with loss of sensation in the feet, new shoes should be worn using a gradually increasing schedule starting at 2 hours/day only, after which shoes and socks should be removed and the feet carefully inspected. Before donning, each shoe should be carefully inspected visually or with the hand for irregularities or foreign objects on the inside, which could injure the foot.

**Shoe Modifications**

Stock shoes may require minor or major modifications by various methods to support the abnormal foot during weight bearing, to reduce pressure on painful areas, and to limit motion or accommodate deformed, weak, unstable, or painful joints. For these purposes, the clinician may select a special type of shoe, order certain alterations in the construction of the shoe, or apply corrections directly to the foot. For this approach to be effective, the clinician needs to make an accurate diagnosis of the problem, have a clear understanding of why corrections are needed, and write a specific prescription that is best accompanied by a simple drawing to clarify the request (Fig. 76-5B).

Although certain simple external modifications may be applied easily to many types of commercial shoes, welt shoes are more suitable to work with, especially for major internal modifications, because the process of removing and reattaching the sole to the upper does usually not alter the structure of the shoe.

**Wedges and Flares**

A flare is an extension attached, either medially or laterally, to the sole of the shoe intended to provide mediolateral ankle stability. The flare can be added to the heel only, or it can extend the entire length of the shoe. A flare is not intended to correct a deformity but rather only to control motion. Widening the base of support with a flare (usually lateral sole flare) may provide a more stable platform for the foot. Indications for a flare include posttraumatic subtalar instability or in patients with hind foot contractures and tone disorders that increase the tendency to roll the ankle (16).

A wedge is used to help accommodate a rigid deformity or correct flexible deformities of the hind foot. A shoe with a lateral wedge has more material under the lateral border of the foot than the medial border.

Flares and wedges may be combined in severe cases but care should be taken to avoid medial flares that exceed ½ in. (1.2 cm) as the patient may trip with his other foot. For lateral flairs in excess of ½ in. (1.2 cm) a buttress should be included for better control (Fig. 76-6).

**Lifts or Elevations**

Elevations of the shoe can be used to address a variety of clinical problems. Most commonly, they are used to compensate for an acquired or congenital limb-length discrepancy or a limb-length discrepancy created by a contralateral shoe modification. When added to the contralateral shoe of a patient with a...
neurological problem who has impairment in limb clearance, it can facilitate the swing phase of the affected limb.

Elevations are applied either under the heel only or under the whole foot. They can be applied internally or added to the outer sole of the shoe. A heel-only elevation is appropriate for accommodating a fixed equinus position or reducing the strain on the Achilles tendon. A buildup less than 1/2 in. (1.2 cm) can be easily added inside the shoe. Greater corrections will need to be distributed between the inside and outside of the shoe to avoid pushing the foot out of the shoe.

Increasing the width of the base of the shoe as the heel height increases should be considered in order to reduce the possibility of mediolateral instability. A heel-only buildup greater than 1 in. (2.5 cm) is not recommended unless accompanied by a high top shoe or an AFO to reduce the likelihood of ankle instability. A buildup that extends from heel to ball (shoe lift) is the more practical and biomechanical appropriate intervention to compensate for a limb-length discrepancy or when attempting to decrease a swing phase foot drag (17). Creating a small rocker may be necessary if the lift stiffens the sole excessively. Buildups can be added to most shoes with little difficulty.

Rigid Shanks
An extended rigid shank is traditionally a strip of spring steel or carbon graphite of varying rigidities that is placed between the layers of the outer sole or in some cases inside the shoe, extending from the heel to the toe brake of the shoe. An extended rigid shank, as its name indicates, extends to the tip of the shoe to make it rigid, reducing bending stresses in the midfoot and forefoot and it must be used in conjunction with a rocker sole. Some of the clinical indications for this shoe modification include metatarsalgia, symptomatic hallux rigidus, and in some arthropathies or Charcot-related joint deformities or pathologies.

Rocker Soles
The rocker sole is one of the most commonly prescribed shoe modifications (11,18). As implied by the name, the basic function of a rocker sole is to rock the foot from heel strike to toe-off without metatarsal bending. Roller soles can effectively be used to reduce pressure under the metatarsal heads and can assist gait by easing and increasing forward propulsion in midto terminal stance. It can reduce enough motion at the first metatarsophalangeal joint, mitigating the pain associated with hallux rigidus. It can be used to replace some of the motion lost due to the use of an extended steel shank. This type of rocker is a common feature of athletic shoes. If the rocker sole is expected to offload an area of plantar pressure, then the apex of the rocker must be proximal to the pressure area.

The actual shape and height of the rocker sole depends on the specific problem to be treated and the expected biomechanical effect of the rocker sole.

A heel-to-toe rocker sole, typically, is thicker than a mid rocker sole and has a more severe rocker angle at both the heel and the toe. It is designed to significantly reduce pressure under the first to fourth metatarsals and to aid in propulsion at toe-off. It also decreases heel-strike forces on the calcaneus and talus, and reduces the ankle motion. This type of modification is not very effective in reducing pressure under the fifth metatarsal head (18,19). Care should be taken to assure that the rocker does not produce mid-stance instability and to encourage the wearer to walk with shorter steps to enhance the efficacy of the modification.

Cushion Heel
Another useful shoe modification is the cushion heel, which consists of a wedge of shock-absorbing foam that is sandwiched between the heel and the sole of a shoe. The purpose of a cushion heel is to increase shock absorption and reduce the knee flexion forces occurring at heel strike resulting in a stable stance phase. A cushion heel may be indicated for patients after ankle fusion or following a calcaneal fracture. It can also be of help in patients with quadriceps weakness or patients with other forms of early stance phase knee instability.

Foot Orthoses
In the hands of an experienced clinician, custom orthotics are an effective tool in the management of most of the abnormal biomechanics of the foot. Orthotics are designed, in principle, to be corrective or accommodative devices (9).

Corrective devices are meant to improve the position of the foot by changing the alignment of skeletal structures. Rigid materials such as thermoplastics, acrylic laminates, and carbon graphite composites are frequently used for this. Cork and polyethylene are example of materials used for semi-rigid orthotic devices that are corrective. Accommodative devices made of soft open or closed cell foams alone or in combination with semi-rigid materials are meant to provide cushioning and support to an already deformed foot structure. They prevent further deformity while ultimately improving function and controlling the direction and/or alignment of the joint movement.

Appropriate posting, longitudinal or transverse arch buildups, heel lifts, pressure relief areas, or other special modifications and the use of various materials in the same orthotic device can be integrated in the design and combined at the time of fabrication. There are several factors to consider when selecting materials used in custom orthotic fabrication; these include the type of activity, the length of time to be used, the amount of forces to be applied across the orthosis, and the amount of axial loading. Also the amount of cushioning that is desirable without introducing mediolateral ankle instability should be considered.

One example of this careful application is the use of thermoplastic elastomer (TPE) for the longitudinal arch, Pelite as a molded substrate and Poron for shock absorption in the same device (Fig. 76-7). Custom foot orthoses (FOs) are superior to prefabricated ones because they can be designed to provide corrective forces to the affected joint(s) with more optimal control. In addition, they can be configured to match precisely the patient’s anatomy (of particular importance in the presence
of impaired sensation) and activity requirements, and they can be manufactured from a variety of materials singly or in combination.

When prescribing a FO, knowledge of the disease process in question, functional anatomy, biomechanics, materials, and finally, recognition of the anticipated functional outcomes are essential for proper prescription. Consideration should be given to the fact that external loads act as dynamic forces that may require different orthotic prescription and materials for different activities.

During the gait cycle, the foot is a dynamically changing lower limb segment, whereas the FO is a relatively static device. Careful follow-up and adjustments, as needed, are essential for long-term success with this treatment modality. Corrective FOs made of rigid materials can be extremely difficult to fit and require meticulous attention to detail. Progressive wearing periods of several weeks with close monitoring of the skin integrity is encouraged.

Few hypotheses regarding the use of FOs have been investigated in controlled studies; hence, much of the design variance encountered is founded on clinical experience and judgment.

FOs are the foundation for lower limb orthotic management. Not only are they suitable for treating many of the common problems encountered in daily practice, but they serve as a basic component in each and every more proximal orthosis design (e.g., AFO, KAFO or hip-knee-ankle-foot orthosis [HKAFO]).

### Ankle-Foot Orthoses

AFOs, because of their mechanical lever arm, can more effectively control the ankle complex and when appropriately designed, can influence the knee joint as well.

The clinician must choose from among metal/leather, plastic, carbon graphite, or a hybrid design assembly incorporating some or all the mentioned materials for an AFO. In general, plastic or hybrid systems predominate in North America and Europe because of the greater degree of client acceptance and control they offer. The choice of materials in the fabrication of orthotics is expanding rapidly and a detailed review of them is beyond the scope of this section. It is very unfortunate that insurance restrictions, in general, prevent open access to some of the recently introduced orthotic technology such as the use of composite or graphite and new advanced design devices. The older-style metal and leather orthoses usually are reserved for selected applications, such as when minimal contact with the soft tissues is desirable due to fluctuating edema, wounds or for heat-sensitive individuals who cannot tolerate the more intimately fitting plastic contours or when fitting large or heavy individuals.

One of the most significant factors in material selection is cosmetic appearance, which strongly affects device acceptance.

A thermoplastic or laminated supracondylar AFO, also known as a patellar tendon bearing floor reaction orthosis, is indicated for patients who lack knee and ankle muscle power but have normal hip extension strength, full knee extension range of motion, and no significant flexor spasticity. A molded footplate and a solid non-articulated ankle design immobilize the foot and ankle in slight equinus, which produces a knee extension force during stance phase. Potential genu recurvatum is controlled by the ligamentous structures of the knee joint or a supracondylar anterior shell and a counteracting posteriorly placed popliteal shell. Blocking plantarflexion of the foot can limit recurvatum as well.

Successful use of AFOs based on electrical stimulation of weak or paralyzed muscles has been recently reported. In the United States, several commercial versions of such orthoses that incorporate neuromuscular electrical stimulation (NMES) have become available for clinical applications. As with any other therapeutic intervention, to achieve their intended use, these devices require appropriate patient selection, understanding of the disease process, and gait biomechanics.

Equinovarus deformity is the most common pathologic lower limb posture observed in the population after central nervous system injuries. This abnormal posture results in an unstable base of support during stance phase. The contact with the ground occurs with the fore foot first, and weight is borne primarily on the lateral border of the foot; this position is maintained during the stance phase. Heel contact may be limited or missing. Limitation in ankle dorsiflexion prevents forward progression of the tibia over the stationary foot in stance phase, causing knee hyperextension and interference with terminal stance and preswing where lack of a propulsive phase is evident. During the swing phase, there is a sustained plantar-flexed and inverted posture of the foot, possibly resulting in a limb clearance problem. The lack of adequate base of support results in instability of the whole body. For this reason, the correction of the abnormal ankle-foot posture by orthotic means is essential. The use of an AFO to control the abnormal posture of the ankle during stance and swing phases should be attempted. An ankle inversion strap or pad should be used to assist in controlling the ankle inversion attitude. The orthosis should be attached preferentially to an orthopedic shoe and the orthotic ankle should include a plantar flexion stop to control ankle plantar flexion during swing and stance.
phases (4). If ankle clonus is triggered during the stance phase, a dorsiflexion stop will need to be used as well to prevent the stretch response triggering this phenomenon. The stop should be set just before the clonus appears. A molded foot plate with a padded long plastic foot-plate, in combination with a toe strap and an extra-depth shoe with high toe box, can be used as an option to accommodate abnormal flexed toe posture (16).

Knee-Ankle-Foot Orthoses
Below the knee, the components of the standard KAFOs are the same as those of metal or plastic AFOs, except that the uprights extend to the knee joint, where they join the thigh uprights (Fig. 76-8). A free flexing knee joint is indicated when mediolateral instability or genu recurvatum may be present but knee extension strength is adequate for weight bearing.

If knee extensors are weak, and knee buckling into flexion occurs, a posterior offset knee joint may be indicated. In the absence of knee flexion spasticity or contracture, an offset knee joint provides a mechanical stable knee during stance but allows knee flexion during the swing phase. This type of joint can provide stance phase knee stability if weight bearing is applied to the limb by placing the orthotic knee joint axis behind the anatomical knee joint. This results in the anterior placement of the line of gravity in relation to the knee. Frequently, this will need to be combined with an ankle joint that is limited in dorsiflexion. It is important to note that this is not a failsafe system since knee instability may occur if the floor is uneven, the knee is not fully extended or loaded or the ankle is allowed to dorsiflex more than 10 degrees (2). In such case presentations, a knee lock may be a better choice. The drop-ring lock is most commonly used, it is placed on the proximal section of the upright bar and drops over the joint when it is fully extended. A spring-loaded pull rod may be added to the ring to facilitate locking and unlocking, especially when the patient is unable to flex at the hip to reach the knee or lacks hand or finger strength and dexterity. A cam lock knee with a spring-loaded cam that fits into a groove in full extension is easier to release but still gives good stability and may be used in patients with flexor spasticity. A bail lock (i.e., Swiss lock) is a lever bow that snaps into locked position on full extension and unlocks automatically when pressed upward. In the presence of a knee flexion contracture, an adjustable knee joint that makes use of either a fan or dial lock may be appropriate. Even when mechanically locked, the anatomical knee with weak extensors would bend slightly on weight bearing if not stabilized by straps above and below the patella or by a patellar pad or strap, a soft leather pad covering the knee cap and fastened with four adjustable buckles to the uprights works well. The thigh uprights are connected by a rigid, padded, upper thigh band with an anterior soft closure. This band should be 1.5 in. (3.7 cm) below the ischium, unless ischial weight bearing is prescribed. Usually, a second, rigid lower thigh band may also be used with soft anterior straps.

The Scott-Craig orthosis eliminates the lower thigh and calf bands and the hip joint typical of a HKAFO used in spinal cord injuries, which facilitates donning and doffing and reduces its weight (Fig. 76-9) (23). It consists of two uprights with four rigid connections: posterior rigid upper thigh band, bail-type knee lock, rigid anterior upper tibial band with soft posterior strap, and at the lower end, a stirrup with a rigid sole plate built into the shoe extending to the metatarsal heads. It is connected to the uprights by double-stop ankle joints that are
Because the anatomic knee joint has a changing axis of rotation, polycentric designed knee joints have the ability to maintain a better alignment during knee movement than single axis knee joints and should be prescribed if the knee is to be allowed to flex in the swing phase.

Plastic Knee-Ankle-Foot Orthoses

Plastic molded or laminated KAFOs may incorporate standard ankle and knee components, but the uprights and bands are made of lightweight laminates or thermoplastics that closely fit the limb (Fig. 76-10). The thigh piece consists of a posterior quadrilateral-shaped shell with or without an ischial weight-bearing or gluteal seat or a narrow mediolateral shaped thigh section if it is to be used bilaterally for improved comfort. The thigh section is closed anteriorly by a plastic band and a Velcro strap or a pair of padded straps. A suprapatellar or pretibial shell provides knee extension force, which eliminates the need for patellar strap and supplies mediolateral knee stability. At the lower end, the uprights are connected to a molded plastic footplate to be worn inside a shoe.

Lightweight modular KAFOs have been designed for quick and easy assembly and provided for children with Duchenne muscular dystrophy in order to extend their walking ability (24). Such a KAFO consists of a plastic thigh piece and an AFO, both available in several prefabricated sizes. The two components are joined at the knee by a metal joint system with an automatic ring or bail locks.

Knee Orthoses

Knee orthoses (KOs) are prescribed to control genu recurvatum and to provide mediolateral stability. With improvements in design and application, KOs have gained recognition by many as a treatment and prevention modality. As such they may be used during sports and other physical activities to functionally


adjusted to place the orthosis in 5 degrees of dorsiflexion for optimum balance (1). The shoe sole is perfectly flat from the heel to the metatarsal bar, where it becomes slightly rounded to the toe. Properly adjusted, the orthosis should stand balanced on its own. It is a stable orthosis that biomechanically functions as the standard KAFO but with fewer encumbrances from straps and relays heavily on the subject’s anatomical structures such as the ligament of Bigelow.

FIGURE 76-10. Plastic-laminated KAFOs viewed from side (A) and front (B).
fabricated and prescribed for high-level physical activities and athletics.

**Hip-Knee-Ankle-Foot Orthoses**

HKAsOFs consist of the same components as described for the standard AFOs and KAFOs, with the addition of an attached lockable hip joint and a pelvic band to control movements at the anatomic hip joint. The movement in the mediolateral plane is controlled with the single axis design while the sagittal motion may be controlled usually with a ring drop lock. The pelvic band, which may be unilateral or bilateral, encompasses the pelvis between the iliac crest and greater trochanter laterally, curves down over the buttocks, and then passes up again over the sacrum. The indications for prescribing a pelvic band have been controversial because several studies indicate that it increases the lumbar excursion and displacement of the center of gravity during ambulation, and thus energy cost may be greater. For most persons with paraplegia, pelvic bands probably are not necessary if they are candidates for Scott-Craig orthosis, although they may improve the standing balance, especially if flexor spasticity is severe.

The Louisiana State University reciprocating gait orthosis (Fig. 76-12) consists of bilateral KAFOs with posterior offset knee joints, knee locks, posterior plastic ankle-foot and thigh pieces, a custom-molded pelvic girdle, and special thrust-bearing hip joints, coupled together with a cam and a thoracic extension with Velcro straps (27,28). The cam mechanism provides hip stability by preventing simultaneous bilateral support the unstable knee or during the rehabilitation phase following knee injury or surgery. However, there are many different categories of KOs leading to the confusion of many prescribers. The use of KOs for the prevention of knee injury in athletes is controversial (6,25).

The multitude of KO designs is, in part, propelled by the growing field of sports medicine and the desire to reduce the prevalence of knee injuries. Most KOs consist of two uprights with polycentric free or adjustable range knee joints, and thigh and calf cuffs. The Swedish knee cage (Fig. 76-11) has no knee joint and prevents recurvatum but permits flexion. The three-way knee stabilizer orthosis looks similar and gives good control of structural knee instability in the lateral, medial, and posterior directions and is indicated for osteoarthritis, flexible genu valgum, and genu varus, and for recurvatum of moderate severity. The standard KOs have short lever arms and may not be effective when strong forces are required for control. Most KOs tend to slip or migrate down the leg during movement, resulting in discomfort from pressure or lack of joint alignment. Following a recent publication by Chew and colleagues, KOs can be broadly classified as unloaders, prophylactic and patellofemoral orthoses (6). No long-term study is available to support the benefit of the use of the unloader brace in osteoarthritis but its effectiveness in short-term use has been documented by Pollo in 2002 (6,26).

Numerous KO designs with longer lever arms and, often, rotational control components have been commercially fabricated and prescribed for high-level physical activities and athletics.  

**FIGURE 76-11.** The Swedish knee cage is an example of a KO.  

**FIGURE 76-12.** Louisiana State University reciprocating gait orthosis.
hip flexion, yet allows free hip flexion coupled with reciprocal extension of the contralateral hip when weight shift during a step is attempted. Using two crutches, persons with paraplegia are able to ambulate with a four-point reciprocal gait pattern. This orthosis also has been tested and used in conjunction with a functional electric stimulation system to facilitate ambulation by persons with paraplegia.

Newer orthotic systems that allow the knee joint to flex in the swing phase of walking and provide stability in the stance phase (stance phase stabilized orthosis) are slowly becoming more common. The system may use mechanical or electronic systems to achieve knee locking in stance and unlocking in swing phase. Increased weight, cost, and maintenance are the main disadvantages. While normalization of gait pattern with likely reduction of energy consumption in the well-trained subject are the principal advantages (29).

Powered HKAFOs that support independent hip and knee movements through computer-controlled motors in the swing phase and joint stability in the stance phase are undergoing clinical trials (30). This device affords independent sit to stand transfers, a reciprocal gait pattern on level surfaces, ramps, and curbs (Fig. 76-13).

Fracture Orthoses
Orthoses of different designs have been used in the management of fractures (31). By definition, a plaster-of-Paris cast applied to a fractured limb is an orthosis that provides rigid immobilization while healing occurs. The term “fracture orthosis,” however, refers to a concept of management based on the hypothesis, supported by considerable clinical evidence, that mobility of adjacent joints does not impede healing of fractures in most cases, that functional activity stimulates osteogenesis, and that rigid immobilization of fractures is not a prerequisite for healing. Even when this management concept is applied, all fractures initially are immobilized either by traction or in conventional casts while the acute pain and swelling associated with the injury subsides and early healing takes place. In general, such immobilization should be maintained for at least 3 weeks, but no more than 6 weeks, before the fracture orthosis is applied. The initial immobilization is done to minimize leg shortening. Problems associated with fracture orthoses are increased angulation of the bone and refracturing, both of which are rare. Fracture orthoses have been used most often to treat fractures of the shafts of the tibia and femur when internal fixation is unnecessary, contraindicated, or refused by the patient and when healing is significantly delayed. Fracture orthoses are contraindicated when satisfactory alignment of the fracture cannot be obtained or maintained.

Initial efforts to use orthotic devices for lower extremity fractures were inspired by lower extremity prostheses. Three basic components are required for fabrication of a fracture orthosis: a cylindrical containment, footplate, and joint mechanisms. The cylindrical component closely fits the fractured limb to provide a hydraulic mechanism that will promote stability for the bony structures and resist shortening. The vertical load from weight bearing is offset by lateral and oblique forces from an essentially incompressible fluid chamber that is created by the encasing cylindrical orthotic component. Most proximally on the orthosis, a weight-bearing surface may be provided, such as a patellar tendon-bearing socket or ischial or gluteal seat, to reduce the vertically directed forces on the fracture site. This mechanism is far less important in the distribution of weight-bearing pressures than the hydraulic mechanism mentioned. The cylindrical components usually are made of plaster-of-Paris cast or low-temperature thermoplastics (Orthoplast). The second major component of a fracture orthosis is a footplate, which is to be worn inside a shoe. The footplate usually is prefabricated and made of plastic, although custom-made footplates occasionally are made. The footplate usually is attached to the cylindrical component by simple plastic hinges rather than metallic joints. Similar joint mechanisms may be used for the knee, connecting the tibial and femoral sections.

 Appropriately designed and fitted, fracture orthoses allow functional ambulation with progressively increasing weight bearing. Absence of pain, good callus formation, and lack of gross motion at the fracture site indicate that the fracture is well stabilized. Fractures of the distal tibia can be treated with a below-knee patellar tendon weight-bearing orthosis (Fig. 76-14). Fractures of the proximal tibia, especially those involving the knee articular surface, generally require a thigh piece that is connected to the leg portion with a polycentric knee joint (32). Fractures of the mid- or distal femoral shaft are managed more successfully with fracture orthoses than are

![image](FIGURE 76-13. ReWalk robotic powered orthosis.)
fractures of the proximal femoral shaft because of the latter’s strong tendency to produce varus angulation and malalignment. The femoral component of the orthosis resembles that of the quadrilateral above-knee prosthesis with an ischial seat and a three-point fixation contour that resists varus angulation (Fig. 76-15). The femoral component is connected to the tibial section by freely moving plastic joints or metal hinges that allow sagittal motion. The tibial section, in turn, is attached to the footplate by joints, as seen in the tibial fracture orthosis (see Fig. 76-14).

Functionally, all orthotic devices aim to control motion, correct a deformity, and compensate for muscle weakness. Orthoses should be prescribed based on the desired biomechanical function (2). A functional problem-oriented based prescription, such as “orthosis to compensate for ankle dorsiflexion weakness,” ensures that the communication with the orthotist is accurate, producing the best solution for a particular patient (4). Further specifying the desired materials or specific joint type, function or motion limitations should be pursued only when clear understanding of gait biomechanics and knowledge of available orthotic components exist. If this knowledge is not readily available, the prescription should be discussed with the orthotist. Physicians should aim at expanding their knowledge of such information to issue the most appropriate prescription.

A complete orthotic prescription should consider the type of footwear to be used, include the joints it encompasses, suggests the desired biomechanical alignment and materials of fabrication for the orthosis. In the very active adult, accommodative orthoses may need to be replaced often to adjust for material deterioration while in the child
CHAPTER 76 | LOWER EXTREMITY ORTHOTICS, SHOES, AND GAIT AIDS

2077

CHAPTER 76 | LOWER EXTREMITY ORTHOTICS, SHOES, AND GAIT AIDS

2077

equinovarus deformity is the most common pathologic lower limb posture seen in neurological rehabilitation. This abnormal posture results in an unstable base of support during the stance phase. The contact with the ground occurs usually with the fore foot first, and weight is borne primarily on the lateral border of the foot, this position may be maintained during the stance phase if spasticity is present. Limitation in dorsiflexion prevents forward progression of the tibia over the stationary foot, causing knee hyperextension and interference with terminal stance and preswing where lack of propulsive phase is noted. During the swing phase there is a sustained plantarflexed and inverted posture of the foot, potentially resulting in a limb clearance problem. The use of an AFO to control the abnormal posture of the ankle during the stance and swing phases can be considered. If an inversion deformity is present, an ankle inversion strap or pad is used to control this problem. The orthosis should be attached to an orthopedic shoe and include a plantar flexion stop to prevent excessive ankle plantar flexion. If ankle clonus is triggered during the stance phase, a dorsiflexion stop can be used as well to prevent the stretch response. When cognition and sensation are not impaired, the use of a molded plastic ankle-foot orthosis (MAFO) with inversion control wing buildup is preferred. A long plastic foot-plate with soft padding, in combination with a toe strap and an extra-depth shoe with high toe box, can be used to accommodate abnormal toe posture (Fig. 76-17).

Application of an Orthosis to Limb Instability During Walking

Knee flexion during early stance phase results in limb instability. In central nervous system problems, patients exhibit this
problem more commonly in the early phase of neurological recovery, with flaccidity and weakness affecting the limb. The same functional problem may be evident in peripheral nerve injuries. The patient is unable to control the knee flexion movement during the early stance phase and is unable to walk without upper limb assistive devices. If external means of knee control are provided through a KAFO, KO, or AFO during early stance, the patient most likely will be able to maintain limb stability and to ambulate. A KAFO with off-set knee joints, a floor reaction AFO set in a few degrees of plantar flexion, or a shoe with a cushioned heel can improve knee stability. It is important that a knee extensor strengthening exercise program be implemented, as knee weakness is promoted by the muscle disuse that the orthotic interventions may cause (7).

**GAIT AIDS**

Canes, crutches, and walkers are often prescribed to assist walking in persons with joint, muscle, and sensory functional losses (Fig. 76-18).

Canes serve many functions. They are an important component in the management of persons with balance deficits. Cane use can increase the anteroposterior and mediolateral base of support. They provide an important safety-related function, providing information related to the position of the limb in space, and assist the individuals with vision-impairment such that they can use the device to scan the environment. Increased joint force reaction has been associated with subjective discomfort at a joint, this is likely further compounded by weakness. A cane, when properly positioned, can decrease the amount of muscle force necessary to stabilize a joint. This in turn leads to attenuation of joint reaction forces and decreased pain symptoms. Muscle weakness may also be offset with cane use, where gait deficits such as excessive knee or hip flexion may be overcome. Pathological conditions affecting the upper limbs may interfere with the use of canes or crutches, and may warrant the prescription of specially designed gait aids.

In the patient with hemiplegia, the cane should be held in the less-affected hand, and the patient instructed to advance the cane and weaker leg simultaneously to attain a three-point gait pattern. When ascending stairs the same patient advances the less-affected (stronger) leg first. The reverse pattern is recommended when descending stairs, with the weaker limb plus cane leading.

When treating hip joint pain or hip muscle weakness, the cane should be held on the opposite side whilst for the knee it may be held on the same side.

Canes consist of a handle, shaft, and base (see Fig. 76-18). Of course, improper fit severely diminishes the effectiveness of the device. Canes are most commonly made of hard woods or aluminum. Designs vary, but all should be fitted with a deeply grooved 2- to 5-cm rubber tip for safety. Several types of tips are available to be attached to the cane shaft. In general, the total height should equal the length from the base of the heel to the upper border of the greater trochanter (35). A single-point cane weighs less, which is important for individuals with decreased functional reserve or upper limb weakness. It is less intrusive, and may be easier to negotiate in cramped spaces. Where stability is compromised, a wide-based cane may provide a greater base of support. These designs consist of three or four short legs attached to a single upright. Three- or four-prong canes have the added advantage of remaining standing when they are released. These canes are often heavy and awkward to use, and temporary use is thus preferred. Cosmesis may also be an important consideration with wide-based cane use.

The most common handle is the "C"-type. It is inexpensive, but this type of handle may be uncomfortable to use.
and difficult to grasp for individuals with hand pathology. Additionally, the line of support differs from the cane's line of action, which produces a bending moment that must be overcome with muscular effort. An alternative approach is a “functional” or ortho-eze handle. A functional handle better conforms to the natural angle of the hand and is centered over the cane shaft (reducing the bending moment). This type of cane is perceived as more comfortable and provides better support while reducing the strain on the wrist and fingers. However, functional handles increase the cost of the device.

The rationale for crutch prescription is similar to canes. The clinical deficits, however, are usually more pronounced. Good upper extremity strength, joint integrity, and ROM are essential to maximize benefit with the crutch use. Shoulder flexors and depressors, elbow and wrist extensors, and finger flexors must have normal or good strength. Most commonly, crutches are prescribed to decrease loading, with muscle weakness or joint pain. Where “push-off” is limited in locomotion, crutch use may aid in forward propulsion. Crutches provide sensory feedback and increase the base of support. Yet they are less commonly prescribed for balance deficits compared to canes.

Crutch types include axillary, Lofstrand, Canadian elbow extensor (i.e., triceps crutch), and forearm support (platform crutch). Axillary crutches are the most common. They are typically made of hardwood or aluminum, with two upright shafts and a series of connecting links. Connections include a padded upper support, middle hand piece, and lower extension. The extension piece and shaft, have numerous, regularly spaced holes for height adjustment. A soft, rubber tip (similar to canes) is attached at the bottom to increase the crutch contact surface area and provide additional stability. Again, proper fit is essential. Total length and handle height are the main considerations. Axillary crutch length should approximate the distance from anterior fold of axilla to: (a) bottom of heel (plus 3 to 5 cm) or (b) a point 6 in. in front of the anterolateral border of the foot.

Drawbacks include increased assistive device mass and restriction of hand use. Some patients may have trouble coordinating gait pattern with crutches. Additionally, patients who lean on the axillary crutches may experience neurological compression injury in the axillary area (i.e., crutch palsy).

The Lofstrand crutch consists of a molded hand piece, a molded forearm piece bent posteriorly just above the hand piece, and a single, length-adjustable aluminum tubular shaft. The forearm piece extends 4 to 5 cm below the elbow. A forearm cuff with a narrow anterior opening is attached at this site. Lofstrand crutches are, typically, lighter in weight, and the hand piece can be released without losing the crutch. The result is that the hands are freed up to perform some activities, which is a significant advantage. Good upper extremity strength and adequate trunk balance are minimum requirements for safe use of Lofstrand crutches.

The triceps crutch extends just above the elbow, and has a single aluminum upright attached to the bilateral uprights. This crutch is rarely prescribed, but may be of benefit in individuals with triceps weakness. Forearm support (platform) crutches may be prescribed when clinical conditions involving the hands, wrists, or forearms prevent safe or comfortable weight bearing. Examples include weakness of triceps or grasp, or arthritic conditions involving the upper extremity joints.

A walker is a frame device that is inherently stable. There are different types of walkers, the most common of which are made of aluminum, have four posts, and may have small front or rear wheels. The person pushes or lifts the walker in front of himself or herself while walking. Walkers have advantages and disadvantages, all of which should be considered when prescribing such an assistive device.

Walkers provide a wider base of support than either canes or crutches, resulting in increased stability. They may be prescribed for patients requiring maximum assistance for balance or with severe sensory deficits, ataxic conditions, and persons with substantial fear of falling. They may be useful during rehabilitation, but care should be taken to assure that the patient does not become too reliant or emotionally dependent on the stability provided by the walker.

Proper use requires bilateral grasp and arm strength. They are bulky in appearance, may be difficult to maneuver in tight quarters, and interfere with the development of reciprocal gait. The use of a walker may tax the endurance of the user's arms. To cope with fatigue, four-wheeled walkers exist and may be outfitted with a seat to make it possible to have intermediate periods of rest. This practice is not risk-free, and the walker should ideally be fitted with hand brakes.

Most walkers are too light for individuals with incoordination. When coupled with good upper extremity strength, weight may be added to this type of walker to attenuate this problem. Walkers can be instrumented with wheels to promote better gait timing without having to stop each cycle to advance the walker. Wheels help patients with incoordination or weakness of the upper extremity, where lifting crutches or walkers and placing them forward may be difficult. The rolling (wheeled) walker does not require as much strength and balance to maneuver, compared to non-wheeled walkers, as the user does not have to lift it from the floor. However, the instability introduced by the wheels, if allowed to roll uncontrolled or to swivel, may prove dangerous.

**SUMMARY**

We have reviewed the basic principles of orthotic intervention, shoe modifications as they apply to the rehabilitation of patients with upper and lower motor neuron pathology. The majority of patients surviving a stroke and other brain insults achieve limited ambulation, and the use of orthoses or other assistive devices attempt to optimize their functional capabilities and encourage improved functional level and quality of life. Orthoses are devices that are applied to the external surface of the body to achieve one or more of the following: relieve pain, immobilize musculoskeletal segments, reduce axial load,
prevent or correct deformity, and improve function. Orthotics for the lower limbs is an integral part of the rehabilitative medical management of the patient population we serve.

REFERENCES


12. Kerrigan DC, Johansson JL, Bryant MG, et al. Moderate-heeled shoes for the lower limbs is an integral part of the rehabilitative medical management of the patient population we serve.


Spinal Orthotics

An orthosis is a mechanical device that applies forces to the body in an effort to support, limit, and stabilize moving parts, assist and improve motion, correct and align deformities, and prevent and protect susceptible areas. The location, direction, and magnitude of these forces vary with the components and design of the orthosis. The word orthotic is derived from the Greek word “orthos,” meaning straight, normal, or true. Orthotics, also called braces or splints, have been known since the ancient Egyptian era. Their use continues till today; however, changes and advances in materials and techniques of fabrication, interrelated with newer surgical procedures and medical treatments, have expanded their applications.

An orthotist is a professional who designs, makes, and, with the referring physician, helps prescribe the proper orthosis for a patient. Orthoses can be prefabricated to fit a large variety of patients or custom molded to a specific patient. Their effectiveness is directly dependent upon the proper fit and alignment of the components as well as proper use and compliance by the patient. The physician and orthotist should prepare a detailed prescription together in order to avoid delivering the wrong device. Once delivered each should be present to “check out” the device in order to ensure proper fit.

The prescription of an orthosis requires an understanding of the pathology of the disorder to be treated and must take into account the goals to be achieved. Knowledge of anatomy, biomechanics, and kinesiology, and an understanding of the indications (positive effects) and limitations (negative effects) of the orthosis are paramount when prescribing such devices.

One should not get confused by the myriad classifications and names for orthoses. Some are named for their founder, others for the location where they were developed, and still others for the parts of the body to which they are applied. The most standard way to name an orthosis is by the joints that it encompasses and the motion it controls (1). In this chapter we review the different types of spinal orthotics, indications for their use, positive and negative effects, and problems that may interfere with the ongoing rehabilitation of the patient. Braces for scoliosis and osteoporosis are not covered in this chapter (see Chapters 32 and 33, respectively).

MECHANISM OF ACTION

Spinal orthotics are prescribed for a variety of reasons (Table 77-1). They are designed to protect the spinal column and supporting structures (ligaments and muscles) from loads and stresses that can cause pain or progression of angular and translational deformity. The physiologic mechanisms responsible for this protection are control of motion, trunk support, and spinal alignment (2).

Control of Motion
The control of motion depends on the flexibility of the device. Numerous studies have attempted to quantify the degree of motion restriction. The most minimal limitation in gross movement likely occurs because the device acts as a physical and kinesthetic reminder to the patient not to move in harmful ways. Braces such as soft collars or lumbar corsets are examples of these. These types of devices also serve to provide warmth and heat to the patient, which may reduce spasm and pain. More restrictive braces act to limit intersegmental spinal motion and further, inhibit flexion-extension, lateral bending, and axial rotation.

Trunk Support
Trunk support is achieved by an increase in the intra-abdominal pressure. An increase in the thoracoabdominal pressure reduces the demand on the spinal extensor musculature and the vertical loading on the thoracolumbar spine and the intradiscal pressure.

Spinal Alignment
Spinal alignment is achieved by the application of the three-point force system inherent in all bracing. The corrective component ideally is located midway between the opposing forces above and below it. These systems shift forces from diseased areas to more healthy segments and prevent unopposed forces from causing deformity.

At each spinal level, orthoses require different designs to achieve their desired function. The desired physiologic effect must be decided upon when prescribing an orthosis so that the least restrictive device capable of completing the job is ordered. For example, if trunk support by thoracoabdominal containment is sufficient to reduce the compressive forces on the spinal column and stress on the musculature, then joint motion stabilization should not be required (2).

Before prescribing an orthosis, one must begin with the indication and develop a goal and then decide which orthosis will achieve the desired goal. Once that goal has been achieved and the device is no longer needed, the device should be
DISCOMFORT MAY BE RELATED TO STRAP TIGHTNESS, THE CONDITION, WILLINGNESS TO TOLERATE A SNUG FITTING APPLIANCE, AND OVERALL COMFORT. DISCOMFORT MAY BE RELATED TO STRAP TIGHTNESS, COMPLAINTS OF CONFINEMENT, OR INCREASED PERSPIRATION CAUSED BY THE BRACE. WITH PATIENTS WHO CAN VOLITIONALLY ADJUST THE STRAPS, THE EFFECTIVENESS OF THE BRACE MAY BE COMPROMISED IF THEY LOOSEN THE STRAPS. INDIVIDUALS WITH A SHORT STOUT NECK AND NO DEFINED CHIN ARE HARDER TO FIT WITH A CERVICAL COLLAR. PENDULOUS BREASTS, SHORT TRUNK, THORACIC KYPHOSIS, OR AN OBSENE ABDOMEN MAKE IT DIFFICULT TO COMFORTABLY FIT CERVICOThorACIC OR THORACOLUMBOSACRAL APPLIANCES. TWO BRACES MAY NEED TO BE GIVEN TO THE PATIENT SO THAT ONE CAN BE WASHED ON A REGULAR BASIS IN ORDER TO MAINTAIN HYGIENE. SKIN UNDER THE BRACE NEEDS TO BE CHECKED AND WASHED DAILY. WHILE IN THE BRACE, AN EXERCISE PROGRAM SHOULD BE IMPLEMENTED, IF POSSIBLE. ONCE THE BRACE IS DISCONTINUED, A MORE AGGRESSIVE STRENGTHENING AND STRETCHING PROGRAM IS INITIATED IN ORDER TO PREVENT THE NEGATIVE EFFECTS OF DISUSE. IN ADDITION, THE PATIENT OR CAREGIVER MUST BE INSTRUCTED IN THE DONNING AND DOFFING OF THE ORTHOSIS, ITS WEARING SCHEDULE, WHETHER OR NOT THE PATIENT NEEDS TO SLEEP AND SHOWER IN THE BRACE, AND THE LENGTH OF TIME THE ORTHOSIS IS RECOMMENDED. FOLLOW-UP OF ITS CONTINUED USE IS REQUIRED BOTH FOR THE PHYSICIAN AND THE PATIENT.

### TABLE 77.2 Categories of Orthoses

<table>
<thead>
<tr>
<th>Cervical (CO)</th>
<th>—soft or rigid head cervical (Philadelphia, Aspen, Miami, Newport)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cervicothoracic (CTO)</td>
<td>—Halo, SOMI, Minerva, any rigid collar with an anterior or posterior extension</td>
</tr>
<tr>
<td>Thoracolumbosacral (TLSO)</td>
<td>—custom-molded body jacket, CASH, Jewitt</td>
</tr>
<tr>
<td>Lumbosacral (LSO)</td>
<td>—chairback, Knight, corsets/binders</td>
</tr>
<tr>
<td>Sacroiliac (SO)</td>
<td>—trochanteric belt, sacral belt, sacral corset</td>
</tr>
</tbody>
</table>

### TABLE 77.1 Indications for Spinal Orthotics

- Stabilize the spine after fracture (with or without neurological deficit)
- Limit spinal motion in cases of pain or sprain
- Support posture and prevent deformity after paralysis
- Postsurgical stabilization (with or without fracture)

Discomfort may be related to strap tightness, the condition, willingness to tolerate a snug fitting appliance, and overall comfort. Discomfort may be related to strap tightness, complaints of confinement, or increased perspiration caused by the brace. With patients who can volitionally adjust the straps, the effectiveness of the brace may be compromised if they loosen the straps. Individuals with a short stout neck and no defined chin are harder to fit with a cervical collar. Pendulous breasts, short trunk, thoracic kyphosis, or an obese abdomen make it difficult to comfortably fit cervicothoracic or thoracolumbosacral appliances. Two braces may need to be given to the patient so that one can be washed on a regular basis in order to maintain hygiene. Skin under the brace needs to be checked and washed daily. While in the brace, an exercise program should be implemented, if possible. Once the brace is discontinued, a more aggressive strengthening and stretching program is initiated in order to prevent the negative effects of disuse. In addition, the patient or caregiver must be instructed in the donning and doffing of the orthosis, its wearing schedule, whether or not the patient needs to sleep and shower in the brace, and the length of time the orthosis is recommended. Follow-up of its continued use is required both for the physician and the patient.

### CERVICAL ORTHOSES (COs)

Cervical bracing can be categorized in several different manners. In general, these devices can be subdivided into two broad categories: cervical and cervicothoracic. Cervical devices encircle the cervical spine, whereas cervicothoracic braces extend into the thoracic spine. When adding a thoracic extension piece, the cervical orthosis (CO) provides greater motion control of the lower cervical spine. To limit extension and hyperextension of the cervical spine, an intimate fit under the occiput must be achieved.

With cervical appliances, the ability to control cervical motion varies significantly from the soft collar that provides minimal restriction to the halo that offers significant reduction in movement. Several studies have examined the effects of various orthoses on mean cervical range of motion (3–8). Many of these studies used different methods to quantify the amount of restriction (e.g., radiographic analysis, goniometric assessment, and computerized spinal motion analysis). In addition, the sample size and characteristics varied from study to study (i.e., cadaver models, healthy spines vs. injured spines). Table 77-3 outlines motion restriction (3–8). For the orthotic to restrict motion adequately, it must fit properly and be worn correctly.

### Cervical Biomechanics

The cervical spine is a highly mobile structure allowing flexion, extension, lateral flexion, and rotation; thus, motion occurs in three planes: sagittal, frontal, and transverse. The atlantooccipital joint primarily permits flexion and extension, with minimal axial rotation and lateral flexion. Functionally, this synovial joint enables an individual to nod their head. At the atlantoaxial (C1-2) joint, the predominant motion is rotation. Having no vertebral body or disc, the atlas rotates around the odontoid axis. Cervical rotation begins first at this articulation and then proceeds caudally. Approximately, 50% of the total rotation achieved by the cervical vertebral column occurs at C1-2. Between C4 and C7, maximum flexion and extension

---

**Cervical (CO)—soft or rigid head cervical (Philadelphia, Aspen, Miami, Newport)**

**Cervicothoracic (CTO)—Halo, SOMI, Minerva, any rigid collar with an anterior or posterior extension**

**Thoracolumbosacral (TLSO)—custom-molded body jacket, CASH, Jewitt**

**Lumbosacral (LSO)—chairback, Knight, corsets/binders**

**Sacroiliac (SO)—trochanteric belt, sacral belt, sacral corset**
occurs, with the greatest motion occurring at C5-6. During flexion the vertebral foramina open and with extension close. Lateral flexion (lateral side bending), however, occurs between C2 and C7 in the coronal plane. Given the configuration of the articulating facets, lateral flexion and rotation are coupled motions. As right rotation occurs, it initiates right lateral flexion and as left cervical rotation occurs, it initiates left lateral flexion. Sagittal motion occurring at C2-7 is uncoupled. The C2-4 region has the most side bending and rotation.

**Soft Cervical Collars**

A soft cervical collar is prefabricated foam rubber with a cotton stockinette covering and Velcro closures (Fig. 77-1). These closures are worn posteriorly. Depending on the patient’s dexterity and upper extremity range of motion, some can only fasten the closures anteriorly and rotate the collar around their neck while others leave the Velcro closures in the front. The manufacturer’s intention was to have these collars worn with the closures facing posteriorly. Collars range in size from small to extra large. To identify the correct size, circumferential neck measurements are taken. This measure corresponds with predetermined sizes. Patients tolerate this device very well. Carter and associates reported that the degree of motion restriction achieved with the soft collar was dependent on the velcro closure position (5). If the intent is to limit flexion, then the collar should be worn in the reverse position with the tabs facing anteriorly. The explanation for this is a function of the starting position of the head. Given its soft material construction, the soft collar can only provide warmth, psychological reassurance, and kinesthetic reminders to limit cervical range of motion; it cannot provide structural support. Essentially, this orthoses reminds the patient not to move. Its use may be appropriate to treat mild muscular spasms associated with arthritic changes and mild soft-tissue injuries.

Soft collars are often prescribed for the early management of whiplash injuries. The effectiveness of its therapeutic use is under scrutiny. A study evaluating whether or not a soft collar reduced the duration and intensity of the patient’s pain following a whiplash (9) showed that test patients wearing a soft collar and control patients not wearing a soft collar reported persistent pain for at least 6 weeks postinjury. Hence, wearing a collar did influence the patients’ pain. In a randomized prospective study (10), comparing the effect of early mobilization wearing a soft collar versus not wearing a soft collar, a significant finding was the number of days it took the group with the collar to return to work. Those who wore the collar took twice as long to return to work (mean 34 days) whereas the mobilization group without the collar returned to work sooner (mean was 17 days). In another randomized controlled study (11), examining the effect of early mobilization versus the use of a collar in patients sustaining a whiplash injury, found that those patients who participated in early mobilization rated their pain and disability less than those wearing a soft collar at 6 weeks after injury. At 6 weeks, the exercise group rated their pain less in the neck and shoulder and had fewer headaches. The early mobilization group was seen by a physical therapist for exercise instruction. In 2006, the same group of researchers published their findings using the same cohort except the patients were 6 months out and the results were the same. Patients who engaged in physical therapy with active exercises had less pain at 6 months than those treated with a soft collar (12). Considerations regarding the negative effects of brace wearing (i.e., psychological dependency, muscle atrophy, etc.) should be weighed under this condition.

**Hard Cervical Collars**

These rigid prefabricated COs are used for either prehospital trauma immobilization or long-term management in patients who sustained a cervical injury. Examples of collars used for prehospital emergency stabilization are the Philadelphia, Stifneck, Ambu, and the NecLoc. These collars are either a one- or two-piece design. Each of these devices is radiolucent and also CT
and magnetic resonance imaging (MRI) compatible. With the exception of the Philadelphia collar, these extrication collars are used for short-term use. The purpose of these prehospital collars is to aid the rescuer in maintaining spinal alignment and stabilization in patients with potential or actual cervical injuries during transport. Using six fresh cadaver geriatric spines, Bednar concluded that the Stiffneck and Philadelphia collars did not provide significant mechanical immobilization in the unstable cervical spine and may be ineffective in preventing displacement (13). This study questioned the effectiveness of these collars during field use. The results have to be considered carefully in light of the sample size and population studied. Nevertheless, the prehospital standard of care is to immobilize the cervical spine with a rigid collar with sandbags or foam blocks anchored to both sides of the head on a backboard.

For long-term patient management, the Philadelphia, Miami J, Aspen, Newport, or Malibu orthoses have been used. The Newport orthoses was replaced by the Aspen collar. Each of these orthoses is available in pediatric and adult sizes. Typically, these devices are prescribed for mid-cervical bony or ligamentous injuries, postoperative stabilization, or post-halo removal. Some specific clinical conditions are anterior cervical fusion, anterior discectomy, and cervical strain. If spinal instability exists, these rigid devices are contraindicated. These appliances are being used as the first line of treatment instead of traditional halo devices for the conservative management of stable upper cervical fractures in adults (14–20). Examples of these fractures include: unilateral avulsion fracture of the transverse atlantal ligament (20), Jefferson fracture, (burst fracture of C1) (14,16), Hangman’s fracture (traumatic spondylolisthesis of the axis-C2 on C3) (15,16,18), isolated lateral mass fracture of the atlas (19), and certain types of odontoid fractures (15–17). Studies analyzing patient outcomes in the aforementioned situations showed stable fracture healing, and no increased disability or neurologic compromise on follow-up examination (14,16,18–20). In addition, these devices are cost-effective, easily applied, and do not have the increased risks associated with the use of the halo. Frequent diagnostic imaging to detect possible instability is strongly recommended. Each of these devices is radiolucent and CT and MRI compatible.

The Philadelphia collar is a two-piece design constructed from closed cell Plastazote foam with molded chin and occipital support. Anteriorly it extends from the mandible to the sternum and posteriorly it extends from the occiput to the upper thoracic spine. The Miami J, Newport, and Aspen collars are two-piece polyethylene shells with internal padding. The Miami J collar offers greater customization; the anterior and posterior shells permit angle adjustability around the chin and occiput allowing for individual differences in bony anatomy. The Newport collar has superior and inferior adjustable supporting tabs that distribute the load along the occiput, upper thoracic spine, sternum, and upper trapezius. Each of these collars has an anterior opening to accommodate a cricothyrotomy/tracheotomy. Given the increasing prevalence of latex sensitivity, physicians may want to consider this when prescribing COs; the Philadelphia, Miami J, Aspen, and Malibu collars are latex free.

According to Goutcher and Lochhead, some patients while wearing a hard cervical collar, Stiffneck, Philadelphia, or Miami J, exhibit a significant decrease in maximal mouth opening (21). In their study evaluating the above listed collars, 51 male and female volunteers’ inter-incisor distance was measured with and without a collar. In an unpredictable manner, several subjects displayed a reduced inter-incisor distance that was less than or equal to 20 mm. This reduced excursion could obstruct a physician’s visualization of the glottis with a laryngoscope. For some patients, their results suggest that attempting tracheal intubation with a collar in place could be difficult. Should a patient require an emergency tracheal intubation, the authors recommend maintaining manual cervical stabilization with the removal of the anterior shell of the collar before attempting intubation.

When prescribing cervical collars with removable padding such as with the Miami J and Aspen, a second set of replacement pads should be included. The patient needs a second set to replace soiled and odorous pads, to allow moist pads to air dry after patient showering or perspiring, and when the pads show wear. Since these pads can be cleaned, it is not necessary to discard them. For specific cleaning directions, have the patient consult the manufacturer’s guide. Proper cleaning prevents skin irritation. Patients with long hair should be advised to wear their hair outside the collar to prevent irritation.

When assessing fit or proper donning of the COs, look at the patient’s face to determine if the chin is centered in the anterior piece. If the chin extends beyond the collar edge, it is too small. If the chin falls inside the collar it is too large. These visual markers indicate whether or not the device is sized or donned correctly. Patients should be familiar with proper donning and doffing procedures.

Furthermore, patients and caregivers should be educated regarding its effects on ambulation. Since the collar forces the patient to look straight ahead, patients should be warned that they will not be able to look down to see their feet or what is below them. As a result the patient must be alert to tripping hazards, that is toys on the floor, cracks in the sidewalk, pets, loose rugs, etc. Since geriatric patients are prone to tripping and falling, the physician should consider ordering the home care therapist to evaluate the home environment for potential environmental hazards and to formulate safety recommendations.

Barry and associates found that wearing a cervical collar affects the driving performance in healthy men and women (22). In a prospective randomized study, 23 volunteers wore a Philadelphia collar while driving and the following parameters were measured: velocity, acceleration, cervical rotation, and driver’s blind spot. Their results showed that drivers were able to stop, turn, and control the vehicle but drove slower. The driver’s blind spot was larger with the collar on than without it. A larger blind spot does affect merging and lane change. Overall, the researchers found that the drivers’
exhibited greater caution. Based on their design and sample size, the authors could not conclude that wearing a collar leads to an increased incidence in motor vehicle accidents. Patients should be advised about the increased blind spot associated with wearing a collar.

Pressure ulcer formation is a potential complication of rigid collar use. Fragile or insensitive skin is particularly vulnerable to ulceration. Common areas susceptible to damage are the occipital protuberance, mental protuberance of the mandible (chin), clavicles, and ears. These wounds may be the result of pressure, shear, or moisture accumulation. A poorly fitting orthotic could exert an external pressure greater than the acceptable skin pressure of 25 to 32 mm Hg; when this occurs tissue ischemia ensues, resulting in an ulcer. In addition, shear forces can arise due to facial hair and skin sliding over the collar surface, or from positional changes. For example, when a patient moves from supine to a semi-Fowler position in preparation for getting out of bed, or if the patient slides down toward the foot of the bed, shear forces can develop. Since beards increase the shear forces, it is suggested that patients shave regularly. Since constant collar wear increases the local skin temperature, excessive perspiration in and around the area can occur. Constant moisture macerates the skin, inducing breakdown. Jirika et al. found that patients with moist skin were four times more likely to develop skin breakdown compared to those with dry skin (23). Provisions should be made to keep the skin clean, dry, and cool.

To assess for skin breakdown, remove the anterior shell to inspect the chin and clavicles, then refasten the straps before rolling the patient onto his or her side. Remove the posterior portion and inspect the occipital protuberance and ear lobes. When removing or applying the collar or a portion of it, the physician must maintain proper neck alignment to prevent injuring the cervical spine. Prior to discharge, patients should be advised to contact their physician if they notice any redness or pressure sores.

Plasier et al. conducted a study to evaluate the craniofacial pressures when using different hard cervical collars: Stifneck, Philadelphia, Newport, and Miami J (24). The study found that the Newport and Miami J collars had lower skin capillary closing pressures, and their open-cell foam material prevented moisture accumulation. In supine and upright positions, the Philadelphia collar exerted high capillary closing pressures leaving the tissues susceptible to injury. In another study, occipital pressure, skin temperature, and humidity were compared when wearing the Philadelphia and Aspen collars (25). Measurements were taken at two separate time intervals, zero and 30 minutes. Using paired t-tests the authors found no difference in pressure or skin temperature with the two collars. Skin humidity, however, was higher when wearing the Philadelphia collar. Skin humidity relates to perspiration and perhaps the closed cell materials used in the Philadelphia collar caused the subjects to perspire more. For patients predisposed to excessive perspiration, the materials used in the collar’s construction should be considered in order to optimize patient comfort, compliance, and minimize ulcer formation.

Additional but uncommon complications associated with the use of hard collars have included marginal mandibular nerve palsy (26), dysphagia (27), changes in intracranial pressure (28), reduction in tidal volume (29), and incomplete tetraplegia (30).

**Cervicothoracic Orthosis (CTO)**

Several prefabricated hard collars (i.e., Philadelphia, Miami, and Aspen) can be made with an extension piece placed anteriorly and/or posteriorly to transform a CO into a CTO. For bedridden patients, devices without a posterior piece are better tolerated. CTOs restrict middle and lower cervical motion. Other examples of CTOs include the SOMI and Minerva. The SOMI is named for its body attachments: sternum, occiput, mandible immobilizer (Fig. 77-2). If the mandibular support interferes with the patient’s ability to chew, or a pressure sore develops, or the patient has a prominent chin, a forehead strap attaching to the occipital support can be substituted. Since this device lacks posterior thoracic support it cannot limit extension; it does limit flexion, however. Furthermore, it may be an appropriate option for bedridden patients since it lacks posterior thoracic coverage. Because the SOMI limits flexion it may be used in cases of atlantoaxial instability with an intact dens, such as in patients with rheumatoid arthritis and in C2 neural arch fractures.

The Minerva is a total contact orthosis with fixation points at the chin, occiput, sternum, thorax, and the forehead. The jacket is made of polyethylene and lined with an open cell material. In addition to controlling the middle and lower cervical spine it can be used for the upper cervical spine below C2. In a study examining the effects of mastication on cervical motion in patients wearing a cervical orthotic, it was found that chewing while wearing a brace with a chin device increased motion above C4 (31). This has implications for patients wearing a Minerva brace while eating. Benzel et al. found that when comparing the halo to the Minerva, the average movement from flexion to extension was 3.7 degrees ± 3.1 for the halo and 2.3 degrees ± 1.7 for the Minerva, suggesting that the Minerva could be considered over the halo in injuries below C2 (32). Between the occiput and C1 the halo provided...
better stabilization (33). When comparing the rigid Miami J collar to the Minerva, the Minerva did not provide better control of the upper cervical spine (7). Greene and associates recommend using either a rigid collar or a SOMI for stable upper C-spine fractures (15).

There has been a published case report of the Minerva brace causing dysphagia with resultant aspiration pneumonia (34). The report outlined several mechanisms placing the patient at risk for aspiration pneumonia: head elevation, slight cervical extension, and limited chin tuck. These combined factors interfere with the swallowing mechanism in some patients. Mechanical changes affecting swallowing have been reported in healthy subjects wearing a Philadelphia, SOMI, and halo vests. Using video fluoroscopy, patients were presented with liquid and solid boluses. Mechanical changes such as alterations in swallowing initiation, reduced hyoid movement, narrowing of the pharynx, and pharyngeal residue were observed; aspiration was not (35). The proposed mechanism for the altered swallowing is related to the 90 degree head spine alignment caused by bracing. This study suggests that if this occurs in healthy subjects, those patients with disordered oral/pharyngeal mobility because of spinal cord injury and/or anterior cervical surgery may be at a greater risk for swallowing problems. An additional study found that mechanical dysphagia in healthy adult patients wearing a halo device was the result of cervical hyperextension (36).

Halo Orthosis

The halo cervicothoracic orthosis (CTO) provides triplanar cervical vertebral motion control. While it offers significant restriction, intersegmental vertebral “snaking” has been described (37). “Snaking” is defined as flexion of one vertebral segment with extension of the adjacent vertebral body. It has been blamed for the motion occurring while immobilized. Koch and Nickel found motion for individuals wearing a halo averaged 31% of normal in the lower cervical spine with the greatest motion at C4-5 (37). Johnson et al. reported 4% sagittal and frontal motion and 1% transverse motion (3). Lind and associates also detected motion ranging from 2 to 17 degrees, with an average of 9 degrees between C2 and C6; the largest movement occurred between the occiput and C1 (38). One study attempted to reduce this intersegmental motion by inserting a posterior pad (39). Using 30 healthy volunteers, wearing a halo with and without a posterior pad attachment, spinal motion measurements were taken in supine and upright positions. The results revealed that the posterior pad did not prevent segmental motion.

The halo has been used postoperatively as an adjunct to internal fixation for upper cervical and mid-cervical fractures and dislocations and as the primary method for conservative fracture stabilization. Although it is the most restrictive appliance, it does not guarantee alignment or fusion. It does have several advantages: it provides immediate cervical spinal stability in patients sustaining an acute fracture or subluxation; it is a nonsurgical alternative for patients refusing operative care or for whom surgical intervention is contraindicated; and it permits early mobilization without risk of compromising spinal alignment. Typically, it is worn for 3 to 4 months.

The basic halo vest components include a rigid open or closed back stabilizing ring fixed to the cranium by four skeletal traction pins with supporting rods and a superstructure that connects the ring to a fleece-lined plastic vest (Fig. 77-3). Instead of four skeletal traction pins, some neurosurgeons insert six traction pins into the cranium. Nemeth found that using six pins for fixation resulted in greater halo strength and cervical stabilization without increasing the occurrence of pin-site complications (40). More than four pin sites are used in pediatrics, patients with thin skulls, certain skull fractures or other conditions that compromise the usual fixation points. It should be noted that pin site complications are a frequent occurrence and require regular vigilant wound care. All current parts and vest designs are compatible with diagnostic imaging: magnetic resonance, computed tomography, or x-ray.
In cases where emergency access to the thoracic cavity is required, such as to administer cardiopulmonary resuscitation, the vest can be opened. To open the vest, some designs require a specialized tool while others have quick release buckles or an integrated hinge system. The rehabilitation team should be familiar with how to open the vest. Lateral, the vest has straps that adjust the tightness across the torso. One study found that the strap tension influenced the degree of spinal motion (41). Increasing the strap tension decreased the motion, especially in lateral bending. No detectable differences were noted for extension. Clinically, the straps should be examined periodically to ensure adequate fit between the vest and thorax. Over the course of rehabilitation, changes in the person’s body habitus are expected; for example, weight loss due to diuresis or changes in appetite can affect chest circumference, and strap adjustments may be indicated. Failure to monitor vest-thorax fit could compromise cervical stability offered by this external device. Trunk deformities may interfere with the patient’s ability to tolerate such an intimate vest fit. Patients and caregivers should be strongly discouraged from arbitrarily manipulating these straps without consulting the physician.

A recent design variation to the standard whole torso vest is an adjustable four-pad (4PAD) thoracic vest (42). Instead of a two-piece vest, this orthotic consists of two lateral pads, one sternal and one thoracic spine pad. These pads move independent of one another. This design does not use shoulder straps. Fuki et al. evaluated this new product and found that it compared favorably in terms of patient outcomes and in its ability to immobilize the cervical spine despite its reduced torso contact (43). These pads do not interface with the clavicles, scapula, or abdomen thus improving patient comfort and reducing the incidence of pressure sore development. Fukui et al. found that the distraction loads were less than the “standard” vest type.

When identifying precautions for the halo vest, the physician should instruct the rehabilitation team to limit the patient’s shoulder abduction to 90 degrees and avoid shoulder shrugging (37). Apparently, large distraction forces are generated with these therapeutic exercises. Such forces may pose a danger to the injured vertebral segment (37). In general, ample space should exist between the shoulders and the shoulder vest straps to prevent excessive distraction forces on the cervical spine. When this space is reduced, the shoulders press against the straps and changes in forces are created.

Care must be taken not to lift, turn, or move the patient by pulling on the plastic vest, the rods, or the superstructure. External forces applied to these areas could affect the spinal alignment and loosen pins. Pin loosening is defined as either a visible pin tip or when the pin can be manually rotated (44). In addition to external forces affecting tightness, there appears to be a natural decrease in compressive pin force over the duration of use (45). Subjectively, the patient may report clicking/grating/creaking sounds, a sensation of looseness, pain at the pin site, headache, or halo vest movement. Since pin loosening is the most frequent complication, all members of the rehabilitation team should be aware of these signs and symptoms.

Loose pins are either retightened or removed and reinserted elsewhere. Initially, the preferred pin placement is anterolateral above the frontal orbital rim and posterolateral below the largest diameter of the skull, the parietal occipital region (46). These positions prevent piercing the temporalis muscle and the frontal and temporal fossa and injury to nearby cranial nerves. If the temporalis muscle is penetrated, the patient will complain of increased pin site pain during mastication. Patients with a cranioplasty may be more susceptible to pin penetration and alternative safe positions for the pins may need to be identified prior to their insertion (47). A case report identified a patient who sustained an epidural abscess as a result of pin penetration through the temporal cranioplasty (47). In another case report, 11 years after the halo was removed, the patient developed tonic-clonic seizures (48). On investigation of the etiology, it was determined that the epileptic focus was the scar tissue caused by the pin penetration. In another case study involving a 71-year-old man wearing a Halo to manage an unstable fracture at C6-7 conservatively, epileptic seizures were developed after 8 weeks of wearing the vest. MR scan of the brain revealed pin penetration into the lamina of the skull with a subarachnoid hemorrhage (49). Vertullo et al. recommend tightening pins 24 to 48 hours after insertion and at 1 week (44). He found that it was a safe and effective method to decrease pin site complications.

Pin site infection is another common complication. To prevent infection, careful inspection for acute signs and symptoms of infection and prophylactic treatment should be ordered for the rehabilitation nurses. Olson recommends cleansing the sites using sterile cotton-tip applicators with antimicrobial soap and normal saline (50). Other topical cleansing agents such as povidone-iodine (Betadine), hydrogen peroxide, and alcohol have been associated with pin corrosion, interference with bacterial autolysis, ineffective infection reduction, and disruption of the healing process (50). Topical ointments can obstruct fluid drainage and increase the risk of infection; hence, prophylactic use is discouraged (51). Crusts surrounding the pins should be gently removed with a saline-soaked cotton-tip applicator. If left alone these crusts could cause fluid buildup predisposing the wound for infection (51). Pin site care should be performed daily. When crusts or drainage are present, it should be done more frequently (50). If drainage or erythema is present then laboratory testing for culture and sensitivity should be ordered to evaluate for an infectious organism. Patients being discharged to home with a Halo vest should be educated regarding proper pin site care. Those patients who are noncompliant with pin site care and develop an infection without follow-up with their physician may develop pin-site myiasis (maggots), a rare complication (52). To date there has been only one documented case in the literature.

In a case study by Rosenblum and Ehrlich (53), a 26-year-old man was treated with a halo secondary to a C6 burst fracture. One week after admission to rehabilitation, his right rear halo pin loosened and it was retightened. Pin care orders were for hydrogen peroxide three times a day. Four weeks later the right pin was loose with mild crusting. Subsequently, the patient presented with signs and symptoms
consistent with psychosis. MRI results were consistent with a brain abscess at the temporal-parietal junction. Bone and brain biopsy revealed *Staphylococcus aureus*. The patient was treated with systemic antibiotics. His psychosis resolved. This illustrates that changes in mentation may be a possible indicator of brain abscess in patients wearing a halo. Furthermore, other signs and symptoms such as scalp/pin cellulitis, headache, eye pain, fever, and seizures should alert the physician to consider brain abscess due to halo pins in their differential diagnosis. In another case study a brain abscess was diagnosed after the fixator was removed. Ray and associates treated a 53-year-old male post halo removal wearing a rigid collar who presented to their hospital with severe headaches, confusion, vomiting, and left hemiparesis (54). CT with contrast revealed a right parietal cerebral abscess. The abscess had developed under the previous halo pin site. In the presence of neurologic symptoms coupled with a past medical history of halo fixation, physicians must consider a cerebral lesion such as a brain abscess.

A prospective study evaluating bone mineral density changes in the cervical spine for patients wearing a halo found that the vest produced local osteoporosis in the immobilized vertebra (55). The reduction in bone mineral density, however, was not related to the level or type of cervical injury, age, or gender of the patient. At the 5- to 6-month follow-up, the local vertebral osteoporosis was mostly reversible.

In preparation for discharging a patient to the community with a halo vest, the patient and caregivers should be aware that the halo has an effect on balance and gait. In a study evaluating balance impairments in young healthy subjects wearing a halo, the authors discovered that their subjects’ balance was compromised (56). Kazuichiro et al. found that wearing a halo vest altered gait kinematics, stride length decreased whereas stride time increased (57). The vest also reduced normal motion between the shoulder girdle and the trunk, the pelvic girdle and the trunk and decreased hip motion. Since there is less motion between the shoulder girdle and the trunk, arm swing is reduced. These gait changes are compensatory strategies for impaired balance. It was concluded that wearing the halo vest affects the gait speed causing the patient to walk slower. This certainly has implications for geriatric patients who frequently have underlying equilibrium impairments. In fact, their cervical injury could have been a result of a fall. Intrinsic and extrinsic factors related to prevention of falls should be explored prior to discharge. Home modifications to eliminate potential fall hazards should be discussed (e.g., removal of scatter rugs, installation of stair rails, wearing proper footwear, etc.). A home visit conducted by the inpatient rehabilitation team or a home care therapist should occur to evaluate the discharge environment and make safety recommendations to prevent falls in patients wearing a halo vest. In light of the slow walking speed, physicians should advise the inpatient rehabilitation and home care therapist to educate the patient regarding safe street crossing procedures and to conduct community ambulation training. Numerous other complications have been associated with the use of the halo device (46–62). Table 77-4 outlines complications associated with the halo orthoses.

### TABLE 77.4 Complications Associated with Halo Vest Use

- Loss of spinal reduction while wearing a halo
- Failure to develop spinal stability after wearing the halo
- Pressure sore development
- Pin loosening
- Pin site pain
- Pin tract infection
- Brain abscess
- Local osteoporosis
- Acute equilibrium impairments
- Forehead scarring
- Dysphasia
- Pin penetration of the skull
- Halo ring migration
- Cranial nerve palsy
- Brachial plexopathy
- Reduction in vital capacity
- Periorbital edema
- Epileptic seizures

In response to the complications caused by the insertion of Halo pins into the cranium to achieve skeletal fixation, an alternative design was proposed: the Lerman Non Invasive Halo (LNHI). It is compatible with diagnostic imaging, including x-rays, CT, and MRI. The design consists of a padded anterior carbon composite chest plate with two posts attached to an open ring, the open ring corresponds to the traditional halo ring, with a nonallergenic silicone material adhering the ring to the skin, occipital, and mandibular supports with posterior straps that secure the anterior thorax plate in place. This device has been used in adults and children. In pediatrics, the indications for use are: immobilization of a stable fusion, reduction of C1-2 rotatory subluxation, postoperative cervical positioning following sternocleidomastoidrelase for congenital muscular torticollis, and in patient who fail standard Halo use (63). Complications associated with its use are facial swelling and skin breakdown. In a small study, three adult patients, who sustained a traumatic nondisplaced or minimally displaced fracture and were neurologically intact, were safely managed with this device during their acute hospitalization (64). No long-term outcome studies are available. As a whole there is very little published research about this orthotic device; however, it holds promise as an alternative to skeletal fixation.

### THORACOLUMBOSACRAL ORTHOSES

The thoracolumbosacral orthosis (TLSO) is a spinal orthosis that provides fixation at the extremes of the pelvis and the shoulders in an attempt to immobilize the thoracolumbar spine in varying directions. All trunk orthoses produce their desired effect by applying anterior abdominal compression, restricting trunk/intervertebral motion, and supporting/aligning the
spine, as previously discussed. Although exact mechanisms are unclear, abdominal compression results in increased intracavitary pressure, which leads to a reduction in the lumbar lordosis and a decreased load on the vertebrae and intervertebral discs. Nachemson showed reduction in intradiscal pressures by 25% with both a corset and a rigid brace (65). Studies to assess the brace efficacy are limited by the methods used to assess motion limitations, (i.e., x-ray, fluoroscopy, EMG, strain-gauge), the patient population compared (healthy, back pain, whiplash, fracture, or postoperative) and the outcome measure (pain relief, fusion, functional improvement, ROM). Nonetheless, the ability of different devices to produce a desired effect is believed to depend on how restrictive they are in design and in materials. The effectiveness in limiting intersegmental motion of the thoracic, lumbar, and sacral spine in all three planes has been studied repeatedly though not as extensively as that of COs. Dorsky and associates found various types of orthoses effective in controlling lateral side-bending as compared to flexion-extension motion (66). Cholewicki et al. compared three TLSOs, the Aspen, the Boston Body Jacket, and the CAMP in normal subjects and found no significant difference in restriction of motion or amount of passive stiffness, but noted that the Aspen was significantly more comfortable (67). Krag et al. also found no significant difference between the Boston and the Aspen on most motion restriction but again the Aspen was rated more comfortable (68). The custom-molded TLSO resulted in the most gross body motion restriction and the corset resulted in the least (69). All TLSOs have been shown to be more effective in reducing motion at the upper levels than at the lower levels, with greatest motion at the lumbar segments (70–72). With three-column injuries, none of these braces were able to adequately limit motion (73).

Myoelectric activity of erector spinae muscles, while wearing various spinal orthotics, has been investigated with inconsistent results. Lantz et al. compared the lumbosacral corset, the chairback, and a molded TLSO. Overall, the TLSO reduced the myoelectric signal activity the most (74). Cholewicki showed only a small reduction in lumbar EMG trunk muscle activity with the LSO (75) and Krag showed no difference (68).

Thoracic and Lumbosacral Spine Biomechanics

The thoracic spine can be thought of best by dividing into upper (T1-4), middle (T5-8), and lower (T9-11) segments and the lumbar spine as the thoracolumbar junction (T12-L1), mid-lumbar (L2-4) segment, and the lumbosacral junction (L5-S1). The 12 thoracic vertebrae are limited in motion in all directions by their attachment to the ribs and orientation of the facet joints; they are further limited in extension by overlapping of their spinous processes. As one moves in a craniocaudal direction, the range of sagittal plane flexion-extension motion increases. The coronal plane (lateral) flexion-extension motion and axial rotation increases to the maximum degree at the lower thoracic and thoracolumbar junction, and then decreases again. At the thoracolumbar junction, the curvatures of kyphosis and lordosis change direction, the facet joints change direction from frontal to sagittal plane, the gravity line bisects the T12-L1 disc and there is the weakest muscular protection. As a result, this area is considered the most mobile segment and is prone to traumatic injuries. The lower lumbar segments, L4-5 and L5-S1, are more susceptible to herniated discs and spondylolisthesis (76).

The TLSOs can be divided into categories based on the direction of motion they control: flexion, flexion-extension, flexion-extension-lateral, and flexion-extension-lateral-rotary control orthoses.

Flexion Control Orthosis

The flexion control orthosis is also referred to as an anterior hyperextension brace, which functions to extend the thoracolumbar region. It is prefabricated and lightweight. The Jewett and cruciform anterior spinal hyperextension (CASH) braces are examples of this type of orthosis (Figs. 77-4 and 77-5). Both use a three-point pressure system without any abdominal compression. The Jewett brace consists of a metal anterior and lateral frame. Attached to the frame are two lateral pads, a sternal pad, a suprapubic pad, and a posterior thoracolumbar pad. The pads exert pressures over a small area and therefore may cause discomfort. The CASH orthosis consists only of an anterior metal frame that is shaped like a cross with the pads attached at the ends. The horizontal bar is attached at the midpoint. It is easier to don and doff than the Jewett and is better tolerated. Both are designed to prevent only flexion and do not limit lateral and rotary movements. They have been used to aid in the treatment of osteoporosis and anterior compression fractures. However, they have not been found to decrease kyphosis and have been found to cause excessive hyperextension forces on
subclavicular pads are added. Called the cowhorn orthosis, it has the added benefit of limiting trunk rotation and flexion in the thoracic and upper lumbar spine; however, it can also cause a compensatory increase in motion at the lower lumbar spine and lumbosacral junction.

Plastic Body Jacket

The custom-molded plastic body jacket is the orthotic apparatus of choice when maximum immobilization is necessary (Fig. 77-8). Fabricated from polypropylene through a plaster cast taken from the patient, body jackets provide total contact to soft tissues with reliefs over bony prominences, thereby distributing forces over a large area for greater comfort, support, and motion control. The superior and inferior portions of the anterior section restrict flexion in the thoracolumbar and lumbosacral segments. The superior and inferior forces from the posterior shell and the forces from the abdominal area in the anterior component resist extension. The upper and lower portions of the lateral aspects of the jacket limit lateral trunk motion. Abdominal compression also alleviates pressure on the vertebrae and discs and reduces myoelectric activity of erector spinae musculature.

Proper fit requires that the inferior border of the anterior shell rest one-half inch above the pubic symphysis, the trim follows the inguinal fold when the patient is sitting, and the superior border encompasses the sternal notch. If the anterior-inferior border is too long, the patient will have difficulty sitting forward and performing sit-to-stand transfers, as the brace would block the forward motion required to transfer. If during transfer training, as the therapist attempts to guide and position the trunk forward in preparation to transfer, the patient complains of pain in and around the pubis, the orthotist must be contacted to trim the brace. Posteriorly, the shell should extend from the spine of the scapula to the sacrococcygeal junction with the upper gluteal mass contained and the axillae and trochanter free (77). Extra precautions must be taken when fitting a jacket to an insensitive patient; however, since the jackets are total contact in design, they are better tolerated. Plastizote lining and ventilation holes make the jackets even more comfortable and breathable. They can be washed and modified as needed; they

the posterior elements, inducing fracture. They are contraindicated in unstable fractures or in cases where extension must be prohibited, such as spondylolisthesis.

Flexion-Extension Control Orthosis

The Taylor brace, designed in 1863, for the treatment of Pott’s disease, consists of two thoracolumbosacral posterior uprights attached inferiorly to a pelvic band and superiorly to an interscapular band that also serves as an attachment for axillary straps (Fig. 77-6). Anteriorly, there is a corset for abdominal compression. The axillary straps extend over the shoulders and pass under the axillae, buckling at each end of the interscapular band. This brace limits the trunk extension, primarily in the mid-to-lower thoracic and upper lumbar areas with a compensatory increase in motion at the upper thoracic, lower lumbar, and lumbosacral junction. The straps must be tight in order for the brace to be effective and thereby may cause shoulder pain limiting the patient’s compliance. In addition, the straps may restrict end-range shoulder motion.

Flexion-Extension-Lateral Control Orthosis

The Knight-Taylor brace is a combination of the Knight (described later) and the Taylor brace (described above). It adds a pair of lateral uprights to the design in an attempt to limit lateral trunk motion (Fig. 77-7).

Flexion-Extension-Lateral-Rotary Control Orthosis

This is similar to the Knight-Taylor brace except that the interscapular band is extended anteriorly and superiorly, and
can even be modified to accommodate a drain. The jackets can be made into lumbosacral orthoses (LSOs) and cervical thoracolumbosacral orthoses (CTLSOs) when the upper thoracic spine needs immobilization. They should be worn at all times when on an incline greater than 30 degrees, though this is often modified by the surgeon (e.g., to allow the patient to shower). Foam designs with rigid stays may be beneficial in those cases where postural forces exacerbate curves and positioning is the main desired goal. Restrictions of activity and brace wearing must be individualized for each patient's needs.

**LUMBOSACRAL ORTHOSES**

The LSO may be either flexible or rigid. The flexible LSOs refer to corsets, belts, and binders. While not effective in restricting motion to a significant degree, they can elevate intra-abdominal pressure, thereby unloading the spine and supporting structures (77), and provide inhibitory kinesthetic feedback and warmth. The rigid LSOs are categorized in the same fashion as the TLSOs, namely by the motion they control; in this case flexion-extension, flexion-extension-lateral, and extension-lateral. A pneumatic orthosis, called the Orthotrac, is a newly designed LSO being used now for low back pain patients. Dalloio studied 41 patients with radicular low back pain using this pneumatic vest and showed a 78% reduction in pain (78). More studies need to be undertaken to assess its efficacy.

**Flexion-Extension Control Orthosis**

The chairback brace is a rigid LSO that consists of two posterior uprights attached to a pelvic band inferiorly and a thoracic band superiorly, with an abdominal support fastened to the posterior uprights with straps. The thoracic band should rest just below the inferior border of the scapula and the pelvic band at the sacrococcygeal junction. Its forces serve to restrict flexion and extension and also provide abdominal compression when the abdominal support is tight.

**Flexion-Extension-Lateral Control Orthosis**

Also known as the Knight brace, it further limits motion with lateral uprights added to the chairback brace (Fig. 77-9). The lateral bars pass over the iliac crests so they must be fitted carefully to avoid pressure on these bony prominences.

**Extension-Lateral Control Orthosis**

The Williams orthosis limits extension and lateral motion but encourages flexion (Fig. 77-10). Its effect is to reduce the lumbar lordosis. It consists of pelvic and thoracic bands joined by a pair of lateral uprights (no posterior uprights), which decreases the lateral motion. The thoracic and lateral bars are directly attached to each other but the pelvic band is attached to the lateral upright with oblique lateral bars that pivot at the top but secure rigidly at the bottom. The abdominal support is elastic and an adjustment strap between the pairs of uprights acts as a lever to pull the oblique uprights more posterior and decrease the lordosis.

**Corsets and Binders**

When treating low back pain, flexible LSOs are the most prescribed spinal orthoses. These are garments made of typically prefabricated fabric, which are closed in the front with Velcro or laces and encircle the lumbar and abdominal areas. Designs vary and include the lumbar/abdominal binder, the Warm-N-Form, and the lumbosacral corset (Figs. 77-11 and 77-12). The latter can be further modified to include posterior stays and shoulder straps.

The binder is elastic and must be wrapped tightly around the lumbar and lower abdominal area in order to elevate intra-abdominal pressure. A common error in donning the binder is placing it above the diaphragm. The addition of a thermoplastic insert (Warm-N-Form) molded to the patients’ lumbar curve and inserted into a posterior pocket may provide increased support and feedback.
sacroiliac joint separations, decrease sacroiliac joint pain and pregnancy–related (79) and postpartum pain.

CLINICAL USES FOR SPINAL ORTHOTICS

Spinal bracing is a common modality used for treatment in a variety of conditions affecting the spine (Table 77-5). There are no definitive recommendations when choosing a particular orthosis for a given condition and considerable variability exists among practitioners. An in-depth discussion of each condition identified above goes well beyond the scope of this chapter. Nevertheless, certain general principles apply.

If motion restriction for pain relief is the goal, and stability is not in question, as in cases of arthritis, sprain, strain, degenerative disc disease, osteoporosis, and even tumors or infections without neurologic impairment or instability, then corsets, collars, and binders may be adequate. This must be balanced with mobilization and proper exercise. The length of its use must be discussed with the patient and tailored based on variables such as age, concomitant conditions, and chronicity of complaints. A review of the literature by Resnick et al, regarding the prophylactic use of lumbar braces for the prevention of low back pain, suggested Class 3 medical evidence that these supports do not reduce the incidence of low back pain or decrease the amount of time lost from work in the general population. When used in those workers with a history of low back pain, there was a decrease in the number of work days lost due to pain. For acute low back pain, their use was effective; however, for the chronic back pain population

SACROILIAC ORTHOSES

Sacroiliac orthoses include trochanteric belts, sacral belts, and sacral corsets. They are prefabricated devices that wrap around the pelvis between the iliac crests and greater trochanters. They are differentiated by the height of their superior borders, their adjustability, and materials. They may have perineal straps attached to prevent upward displacement and a sacral pad to apply pressure over the sacrum (Fig. 77-13). While minimally effective in restricting motion, they may increase the intra-abdominal pressure and provide kinesthetic feedback to maintain a neutral pelvis. These devices can help to stabilize the sacroiliac joint, support pelvic fractures and traumatic

<table>
<thead>
<tr>
<th>TABLE 77.5</th>
<th>Orthotic Treatment of Clinical Conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sciatica</td>
<td>Degenerative disc disease</td>
</tr>
<tr>
<td>Degenerative disc disease</td>
<td>Herniated nucleus pulposus</td>
</tr>
<tr>
<td>Herniated nucleus pulposus</td>
<td>Radiculopathy</td>
</tr>
<tr>
<td>Radiculopathy</td>
<td>Spondylolisthesis—congenital and degenerative</td>
</tr>
<tr>
<td>Spondylosis</td>
<td>Osteoarthritis—facet joint arthritis</td>
</tr>
<tr>
<td>Osteoarthritis—facet joint arthritis</td>
<td>Traumatic fractures—compression, burst, fracture-dislocation</td>
</tr>
<tr>
<td>Traumatic fractures—compression, burst, fracture-dislocation</td>
<td>Osteoporosis—kyphosis, compression fracture</td>
</tr>
<tr>
<td>Osteoporosis—kyphosis, compression fracture</td>
<td>Spinal fusion</td>
</tr>
<tr>
<td>Spinal fusion</td>
<td>Laminctomy or discectomy</td>
</tr>
<tr>
<td>Laminctomy or discectomy</td>
<td>Inflammatory arthritis—rheumatoid arthritis, psoriatic arthritis, ankylosing spondylitis</td>
</tr>
<tr>
<td>Inflammatory arthritis—rheumatoid arthritis, psoriatic arthritis, ankylosing spondylitis</td>
<td>Infectious diseases—spinal osteomyelitis and abscess, tuberculosis</td>
</tr>
<tr>
<td>Infectious diseases—spinal osteomyelitis and abscess, tuberculosis</td>
<td>Spinal tumors</td>
</tr>
<tr>
<td>Spinal tumors</td>
<td>Paralysis—tetraplegia, paraplegia</td>
</tr>
<tr>
<td>Paralysis—tetraplegia, paraplegia</td>
<td>Spina bifida</td>
</tr>
<tr>
<td>Spina bifida</td>
<td>Scoliosis</td>
</tr>
<tr>
<td>Scoliosis</td>
<td>Cervical and lumbar sprain</td>
</tr>
<tr>
<td>Cervical and lumbar sprain</td>
<td>Torticollis</td>
</tr>
<tr>
<td>Torticollis</td>
<td>Whiplash</td>
</tr>
</tbody>
</table>
their use was not. This review also felt that if a brace was used for the chronic back pain population, rigid braces had more benefit than soft ones (80). In cases of spondylosis and low back pain, Iwamoto showed that conservative treatment with rest and lumbosacral bracing was satisfactory in controlling symptoms (81).

In cases requiring orthoses for spinal stability, flexible devices are not adequate. In cases of vertebral fracture and trauma, spinal level and degree of injury must be ascertained. Treatments may include surgical stabilization, bracing, or both. Controversy in treatment exists in deciding which fractures have enough stability to be treated in an orthosis alone versus those in which surgical stabilization is needed. Denis described instability in 1984 (82,83) based on the three-column spine with instability defined as it relates to the integrity of the middle column. While some base their evaluation of stability on the presence of injury to the posterior elements (84,85) and others on the structural integrity of the middle column’s bony as well ligamentous structures (86), Denis’ theory was supported by Panjabi et al. in 1995 (87) and is generally accepted. Other conditions considered unstable generally follow these same principles with variability due to needs of the specific patient.

First-degree instability (Denis type I) is seen in compression fractures. In these cases, only the anterior column is affected. The middle column’s structural integrity remains intact and is therefore considered a stable fracture. Conservative treatment will suffice in the majority of these injuries and tends to include hyperextension bracing or body jacket. When the compression is 30% or less, no external support is required, and early ambulation was reported to be as effective as bracing intervention (88). Orthosis had no effect on back pain or eventual disability in a retrospective review by Folman and Getstein, in stable thoracolumbar fractures with kyphosis less that 20 degrees and compression less than 50% (89). There exists, however, a subset of these lesions that can progress and lead to a posttraumatic kyphosis. According to Ferguson and Allen (90), this occurs when the compression is greater than 50% of the original body height and is the result of posterior ligamentous injury. Frequent monitoring is required. Teter et al. believe that compression fractures with kyphosis greater that 30 degrees require evaluation of the posterior ligamentous complex and if damaged then surgery should be considered even in the neurologically intact patient (91).

Second-degree instability is typical of the burst fracture (Denis type II), a compression injury of the anterior and middle columns. These fractures may behave more like fracture/dislocation injuries, making categorization in terms of “burst” variable. This has resulted in considerable controversy in treating these fractures. Nevertheless, two-column instabilities that are neurologically intact can be treated conservatively with a Jewett, CASH, or custom-molded TLSO appliance (92,93). Wood et al., in a prospective study, found that operative treatment of thoracolumbar fractures provided no major long-term advantage compared with nonoperative treatment with bracing (93). Fractures with posterior element instability (three-column injury) will likely require surgery.

Seat belt injuries (Denis type III) may heal (about 50% of these injuries) with a hyperextension TLSO if the fracture line passes through bone, the so-called Chance fracture (94). However, if the fracture passes through the soft tissues, known as a slice fracture, then surgery is required. Fracture/dislocation injury (Denis type IV) is an unstable fracture and frequently results in paralysis. This always requires surgery unless there are other contraindications (73).

After decompression and/or fusion procedures for degenerative disease, the literature varies on the need of bracing. Few studies deal with the subject and none with an adequate study design to support or refute their use. No reliable information is available on the effectiveness of bracing for improving fusion rates or the benefit of clinical outcomes. While it is possible that some fusion patients may benefit, our ability to select a patient is poor (95). Following surgery for fracture or ligamentous injury, the literature again varies. In both cases, bracing will depend on the type and location of the fracture, if present, the procedure performed, any underlying spinal pathology and mostly on the personal views of the surgeon (96).

In any of these cases, the patient may participate, if needed, in a rehabilitation program to address mobility, activities of daily living, and community reentry. The patient should also be instructed in any restrictions of activity, donning and doffing of the brace, cleaning of the brace, and proper skin care. A T-shirt should be worn under the TLSO to protect the skin, and the orthosis must be donned tightly, otherwise the stabilizing effect is diminished (73). It is usually best to don a TLSO while in the supine position in order for it to be aligned properly.

REFERENCES

PART VI | MANAGEMENT METHODS

Freedom of movement is an essential component of human independence. In the Institute of Medicine model on the enablement/disablement process (1), a wheelchair and its components are fundamental for altering the interaction a person with a mobility limitation has with the environment. Not surprisingly, studies in various populations and cultures have repeatedly shown that mobility is closely tied to quality of life (2–6). It is therefore critical that the physiatrist takes an active role in the wheelchair prescription process. Fundamental to this role is knowledge related to the understanding of the complex components that make up seating and wheelchairs.

Seating and wheelchair technology has become increasingly complex. In this chapter, the readers will find information to help them gain an understanding for differentiation among various types of manual and power wheelchairs, and seating systems. In addition, newer hybrid wheelchairs, such as power add-on units, are discussed. The importance of the wheelchair-user interface, ride comfort, durability, selection of accessories, and powered-wheelchair control devices is also stressed. The readers are introduced to wheelchair and seating measurements and a variety of cushions and postural supports.

PRESCRIPTION PROCESS

Team Approach

The complexity of wheelchair and seating components combined with the nuances of individuals and various disease processes make it virtually impossible for a single clinician to act independently when prescribing assistive technology for mobility. For this reason, it is important to involve an interdisciplinary team in the decision-making process (7). The most important team member is the patient. The opinions and desires of the patient are critical to a successful fitting but must be assessed in terms of their level of knowledge and insight. Some patients have been using a wheelchair for years and know exactly what they are looking for. For these individuals, team members act to provide unbiased information. Alternatively, a novice patient may have little knowledge of what is available and the trade-offs of each decision. For this individual, the team will need to be more directive. The family and caregiver should also provide input, as they will be the next most affected by the choice of wheelchair.

The assessment team can consist of a variety of rehabilitation professionals (Table 78-1). Currently, Medicare and the majority of private insurance companies require a face-to-face physician assessment focused on mobility before providing a wheelchair. Given the training on function and mobility, a physiatrist can be the most appropriate physician to complete this assessment. Often an occupational or physical therapist working with the physician will conduct an evaluation. Ideally, the therapist will be certified by The Rehabilitation Engineering and Assistive Technology Society of North America (RESNA) as an assistive technology practitioner (ATP). Other professionals who can be certified and may be part of the team include a rehabilitation engineer, an Assistive Technology Supplier (ATS), or Certified Rehabilitation Technology Supplier (CRTS, National Association of Rehabilitation Technology Suppliers).

Those involved in the wheelchair selection process should have knowledge about the technology available on the market. Magazine articles and commercial database sources such as ABLEDATA (http://www.abledata.com/) and SpinLife (www.spinlife.com) are good places to research devices or to direct patients who want to educate themselves. (As with any medical advice, patients should be cautioned that information gathered from TV, internet, magazines, etc., should be confirmed with their physician, therapist, or technology supplier.) People choose their seating and mobility devices based on the features available that will facilitate activities or address needs (8); therefore, it is important to be aware of the user’s preferences and the features of various devices.

Patient History

The interview process is largely the same as any standard history conducted with a patient. The necessary information such as age, past medical history, and current medical diagnosis needs to be gathered. The assessment process involves obtaining critical information about the user and his or her environment, family support, and past use of assistive technology. It is essential to determine if the individual’s physical impairment is changing rapidly or is stable. It is important to establish the diagnosis that requires the wheelchair and to assure that there are no ongoing medical problems or complications that can affect the wheelchair prescription and the patient’s health.

To properly specify a wheelchair, it is important to understand both the intentions and the abilities of the user (9). It is essential that the clinician takes into account the characteristics
of the intended user and potential caregivers to ensure that the wheelchair will be accepted and used by the user. Access to a range of heights might be important to reach objects in the home and other environments. To support vocational needs, specific requirements may exist for mobility within a laboratory, operating room, courtroom, or machine shop. Leisure activities, pursued in such places as community centers, restaurants, movie theaters, and recreational environments, often place the most demands on the wheelchair. For some users, many of the desired tasks may be accomplished simply and with existing off-the-shelf technology. Others will require custom products, and some will not be able to achieve all of their goals with existing technology.

Additional necessary information includes type of insurance, physical capabilities and if the individual is able to transport a wheelchair. Also, if the patient has been using a chair, historical information about his or her current chair should be addressed, including problems he or she may be having. The wheelchair model chosen should also be compatible with the user's public and private transportation needs (such as a bus, train, car, van, or airplane) and home/Activities of Daily Living (ADL) environment.

The surface conditions may impose restrictions on the type of wheelchair that is most appropriate. The regularity of the surface and its firmness and stability are important in determining the tire size, drive wheel location, and wheel diameter. The performance of the wheelchair is often dictated by the need to negotiate grades, as well as height transitions, such as thresholds and curbs. The clearance widths in the environment will determine the overall dimensions of the wheelchair. A scooter may not fit around a corner in an average person's home but might perform well outdoors on a sidewalk. The need to be able to operate in snow, rain, and other weather conditions are important considerations as well.

Physical Examination
With an understanding of the individual's need or desire to perform different activities, the next step is examining the user. The history likely provided significant insight related to his or her physical abilities. To verify this, a physical examination should focus on aspects of the patient that will help (a) justify the wheelchair and seating system, (b) determine the most appropriate wheelchair and seating system, and (c) assure that medical issues are appropriately addressed. Based on the interview of the patient, it may be possible to omit some portions of the examination listed below.

Often individuals require a wheelchair because of cardiopulmonary disease. For these individuals it is important to document a heart and lung examination. Attention should be paid to dyspnea on exertion and changes in vital signs with activity. These findings can be used to justify a power wheelchair, as the energy cost of wheelchair propulsion is not less than walking (10). Other common reasons for requiring a wheelchair are musculoskeletal and neurologic deficits. The clinician should document the neurologic and musculoskeletal deficits in a methodical fashion. In a patient with a stroke, for example, an examination to check for neglect or visual field deficit is important because it will impact on the ability to independently drive a chair.

Obvious examination items include strength and range of motion. For individuals with chronic arthritis problems, the examination should document the painful, swollen, or mal aligned joints. When no strength deficit is seen, it is important to document issues with coordination, tone, and proprioception. For example, a study by Fay et al. found that many individuals with multiple sclerosis were unable to effectively propel manual wheelchairs due to increased tone and decreased coordination (11). Participants showed an inability to maintain a speed comparable to community walking speed (1 m/s). This type of finding can provide justification for the device selected.

While completing the examination, the physician should be thinking about how the individual will control the wheelchair. If there is poor hand coordination, head control or switch control may be needed. In certain cases, a foot joystick may be possible.

For some individuals, the examination and history will not necessarily establish a clear need for a wheelchair. In these cases, it is important to consider alternative options that meet the individual's functional needs.

It is important to establish how long the individual will be using the chair. If from the history and examination it is determined that the deficit will be transient, then a rental chair may be appropriate. Medicare will rent chairs for their beneficiaries. This is a good option if the duration of use will be short and the device does not need to be customized. If during the history and physical examination medical issues that require intervention are identified, it may be appropriate to delay wheelchair prescription so that changes in the patient status do not necessitate changes in the prescription. In such a case a rental may provide short-term mobility.

Stability can be assessed by observing the individual in the current wheelchair or by asking him or her to sit unsupported on a mat table. Ask the patient to perform simple reaching tasks to determine the lateral and forward stability of the trunk, hand and arm strength, and hand fine motor skills. The presence of kyphosis, scoliosis, or other fixed deformities should be determined. Critical point to evaluate is hip and knee range of motion because contractures may need to be accommodated. Poor stability usually indicates the need for special attention to seating and position. Appropriate seating can enhance reach and stability, thus improving the performance of manual activities from the wheelchair. It is well known that various groups of wheelchair users, such as individuals with tetraplegia and cerebral palsy, will develop kyphosis or scoliosis over time (12).
What is less well known is whether spinal deformities can be prevented with appropriate seating. Even if prevention is not a goal, accommodation is needed for comfortable seating.

Finally, a thorough check of the individual’s skin is important. This may not be needed for individuals with cardiopulmonary disease, but it is essential for individuals with neurologic deficits or those with previous history of pressure sores. The examination should include not only the buttocks but also the feet and calves, which can be affected by pressure against a leg rest. Attention should be paid to bony prominences and previous scars. This examination will help with cushion selection and wheelchair setup. Large, previously untreated ulcers are sometimes discovered, ultimately leading to treatment before seating plans can be implemented.

For many people a few simple measurements can be used to determine the proper dimensions for a wheelchair (13). Body measurements are typically made with the consumer in the seated position. Probably the most obvious body measurements are the consumer’s height and weight. The consumer’s weight is critical to obtaining a wheelchair that is sufficiently strong as many wheelchairs are only rated to hold up to 113 kg (250 lb). The height of the wheelchair user provides information about the person’s size and can be used to check the final wheelchair measurements. For example, the sum of the sitting height, sitting depth, and lower-leg length should be close to the person’s supine height. Additional measurements and definitions are used when specialized seating and postural support systems are required. These measurements may be completed by any one of the team members.

Documentation
Providing appropriate and quality wheelchairs requires well organized and often extensive documentation. At the very least, a prescription and various insurance forms must be provided. For more expensive and complex interventions, a letter of medical necessity (LMN) is required. Table 78-2 list Medicare recommended assessment findings for power wheelchair provision. This letter can consist of two components: a cover letter and the “client/patient evaluation and intake form.” The cover letter summarizes the person’s disability, problems with existing equipment or method of mobility, evaluation procedures, conclusions, explanation of why lower-cost alternatives will not work, risks of not providing the equipment, and a line-item justification for each of the various components being recommended. The “client/patient evaluation and intake form” guides the evaluation process and captures in-depth information of the evaluation findings required to support recommendations of the appropriate seating and mobility interventions. Often, a part of this exam will be completed by an occupational or physical therapist. The therapist will perform clinical trials and simulations, which allow the patient to try many different devices.

The prescription process should include a home assessment to make sure the device selected will work in the home environment. A home assessment conducted by the supplier is required by Medicare. This requirement highlights the importance of working with qualified suppliers such as those with the ATS credentials. If a home visit is conducted, documentation of the visit should be included in the LMN.

In order for any health insurance provider to approve coverage for a wheelchair and seating system, the practitioner must establish and document medical necessity. Each funding source may have its own definition of “medically necessary,” however; in general, when it comes to wheelchairs, it is necessary to accommodate or replace a malfunctioning body part (i.e., paralysis or weakness of the lower extremities) or to reduce or manage disability. Many funding sources also require
the recommended intervention to be the least costly, reasonable alternative. Therefore, as part of an evaluation, it is helpful to document that lower-cost alternatives have been tried and were unsuccessful, and to cite specific reasons for the higher-cost choice. It is also helpful to document the potential outcomes if the person is not provided with the equipment. Examples of these risks include falls and fractures, development of pressure sores, joint contractures and musculoskeletal deformities, increased pain and discomfort, loss of function, and ultimately being more restricted to a bed or chair.

The final letter is reviewed by both the therapist and physician and cosigned. It is not appropriate for a physician or a therapist to sign the LMN for a client/patient if he or she has not evaluated the person for the device prescribed.

Follow-Up
It is important that the delivery and the final fitting of the device is documented and verified by a follow-up visit with the team or one of the team members. This visit provides the client with the support of the team to be comfortable with the final acceptance and approval of the device and supports the supplier with appropriate clinical delivery documentation. Finally, this visit can be used to refine the patient’s wheelchair driving skills to enable safe operation.

MANUAL WHEELCHAIRS
Manual wheelchairs offer many advantages over powered mobility. Manual wheelchairs are much easier to transport because of their lighter weight. No special equipment is needed to place a manual wheelchair in a backseat, and individuals with paraplegia and tetraplegia are often capable of transporting their wheelchairs independently without additional technology. In addition, manual wheelchairs generally require less maintenance than power devices, and there are no concerns related to batteries or controllers. Finally, manual wheelchairs offer a degree of physical exercise that can benefit the wheelchair user.

Depot and Attendant-Propelled Wheelchairs
The depot or institutional wheelchair is essentially the same wheelchair that was produced in the 1940s. This type of chair corresponds to the Medicare category of K0001 and, despite its numerous shortcomings, is the default chair for many insurance companies and Medicare. Today’s depot wheelchairs may be a bit lighter than the 1940s models, but the basic frame design is unchanged. Depot wheelchairs are intended for institutional use, where many people may use the same wheelchair. These wheelchairs are typically used in airports, hospitals, and nursing care facilities. They are inappropriate for active people who use wheelchairs for personal mobility, including older persons in nursing homes. Depot wheelchairs are designed to be inexpensive, to accommodate large variations in body size, to be low maintenance, and to be attendant propelled. Unlike the attendant-propelled chairs described below, depot chairs are designed neither for the comfort of the person being transported nor for the person pushing the chair. A typical depot wheelchair will have swing-away footrests, often removable armrests, a single cross-brace frame, and solid tires (Fig. 78-1). Depot wheelchairs have sling seats and back supports, which are uncomfortable and provide little support. Swing-away footrests add weight to the wheelchair; however, they make transferring into and out of the wheelchair easier. Armrests provide some comfort and stability to the depot wheelchair user and can aid in keeping clothing off the wheels. Depot chairs typically fold to reduce the area for storage and transportation. Solid tires are commonly used to reduce maintenance. Solid tires typically dramatically reduce ride comfort, increase rolling resistance, and add weight. There is very little, if anything, that can be adjusted to fit the user on a depot chair. Typically, only the leg-rest length is adjustable. Depot chairs are available in various seat widths, seat depths, and backrest heights.

Attendant-Propelled Chairs
Not all wheelchairs are propelled by the person sitting in the wheelchair. In many hospitals and long-term care facilities, wheelchairs are expected to be propelled by attendants. In addition, some individuals with severe disabilities are unable to propel a wheelchair or control a power wheelchair. For children who use attendant-propelled chairs, it is necessary to continually reassess if they may be able to use independent mobility. The primary consideration in the attendant-propelled chair is that the chair has two users: the rider and the attendant. If the wheelchair is propelled solely by attendants with no assistance from the rider, then there may be no need for the larger drive wheels (Fig. 78-2). To keep the attendant comfortable, the weight of the chair should be kept as light as possible and the push handles adjusted to the height of the attendant. If the occupant will be sitting in the chair for prolonged periods of time, then attention must be paid to comfort. For this reason, attendant-propelled chairs often have tilt-in-space as an option.
A variant of the attendant-propelled wheelchair is sometimes called a “Gerry” chair in reference to geriatric users. This type of attendant-propelled wheelchair is typically designed to make transfer out of the chair difficult. The rider is seated in a large recliner-type wheelchair. The soft padding, reclined position, small wheels, and large size make it impossible for the rider to move the wheelchair and difficult for most riders to exit the wheelchair. This helps long-term care facility to exercise control over their clients with cognitive dysfunction. There has been considerable discussion about the appropriate use of attendant-propelled chairs that significantly restrain the rider’s independence.

**Lightweight and Ultralight Wheelchairs**

The terms *lightweight* and *ultralight wheelchairs* are derived from the Medicare categories K0004 and K0005, respectively. K0004 wheelchairs must weigh less than 34 lb without footrests or armrests, and K0005 must weigh less than 30 lb without foot or arm supports. K0004 wheelchairs have very limited adjustability (Fig. 78-3). Like depot chairs, they can be sized to the user, but many of these chairs do not offer features such as adjustable axle plates, quick-release wheels, or a method to change the seat to back angle of the wheelchair. Because of the way Medicare reimbursement works, manufacturers attempt to build the best wheelchair possible under a certain Medicare reimbursable cost. Unfortunately, this practice of cost engineering does not necessarily lead to improvements in design. In addition, this Medicare policy may cause dealers to push wheelchair users toward K0001 and K0004 chairs which have higher profit margins.

The ultralight wheelchair is the highest-quality chair that is designed specifically as an active mobility device (Fig. 78-4). These chairs, which can easily cost more than $2,000, are usually highly adjustable and incorporate numerous design features made to enhance the ease of propulsion and increase the comfort of the wheelchair user. When made of titanium or high strength aluminum, this chair can easily weight less than 20 lb.

At present, it is necessary to justify the need for a K0004 or K0005 wheelchair instead of a standard K0001. Unfortunately, the process of getting prior authorization—meaning the vendor is guaranteed ahead of time to be reimbursed for the wheelchair—is cumbersome. As a result, vendors are unwilling to take the risk that a $2,000 item will be rejected by Medicare and become their problem. Currently, Medicare allows vendors to ask for preauthorization of these chairs; however, few vendors take advantage of this opportunity.
Ultraportable wheelchairs usually have a number of options and adjustments that can be made to appropriately fit the user. Following is a list of many of the components of chairs and options that are available. It is important to remember that every component adds weight to the chair. A balance must be reached between providing the optimal equipment to make the individual as functional as possible and to keep the chair as light as possible. Some of these options are also available on lightweight and depot type chairs; however, the components used on ultralight chairs are generally lighter and better in quality.

**Frames**

There are two basic frame types: folding and rigid. Within these two frame types, there are a number of different varieties. The most common type of manual wheelchair frame is the folding cross-brace frame (see Fig. 78-3). When viewed from the back of the frame, the cross members form an X with a hinge located in the middle of the X. The chair is folded by pulling upward on the seat upholstery. Cross-member folding mechanisms are simple and easy to use. However, the wheelchair may collapse when tilted sideways, and the frame becomes taller when folded. Some chairs incorporate snaps or over-center locking mechanisms to reduce the problem of frame folding while on a side slope.

The most common rigid chair is the box frame. The box frame is named for its rectangular shape and the frame tubes that form a “box” (9). Box frames can be very strong and very durable. These frames can also be collapsed to relatively small dimensions. The backrest usually folds forward, and when used with quick-release wheels, the chair becomes a rather compact shape. An alternative to the box frame is a cantilever frame that can act as a suspension element (i.e., there is some flexibility purposely built into the frame). These cantilever frames may also have fewer tubes and fewer parts and thus be more aesthetically pleasing (see Fig. 78-4).

Some manufacturers offer suspension elements on the frame. This is in part a response to evidence that vibration exposure in wheelchair users is excessive (14). Hinges are placed at the front of the seat, and elastic elements are placed at the back of the seat. The elastic elements act to provide some suspension. The flexible element for the suspension can use either metal springs or polymer dampeners. Elastic elements have not necessarily resulted in lower levels of vibration being transmitted to the user (15). In addition, the shock absorption can result in lost energy during propulsion. This can occur because the force generated by the wheelchair user during propulsion goes toward compressing the elastic elements rather than forward motion. Therefore, the decision to purchase a suspension wheelchair should depend on whether the patient prefers the drive, feel, and comfort of the wheelchair.

**Components**

A number of components can be attached to both manual and power wheelchairs. The following list is focused on manual wheelchairs, but many of the components are found on power wheelchairs as well.

**Footrests**

Most wheelchair users require support for their feet and lower legs. This support is provided by footrests. Footrests may be fixed, folding, swing-away (see Fig. 78-2), or elevating. The footrests must provide sufficient support for the lower legs and feet, and must hold the feet in proper position to prevent foot drop or other deformities. It is essential to assess limitations in knee and foot range of motion. Some users have very tight hamstrings, requiring that the feet be positioned closer to or under the front edge of the seat. This is difficult to accomplish in most configurations. Extending the knees out to accommodate the standard design of the footrest position in front of the seat results in a sitting posture with a posterior pelvic tilt and a tendency to slide forward in the seat. This is commonly seen in elderly people in nursing homes using depot-type wheelchairs.

The feet must remain on the footrests at all times during propulsion, and therefore some type of cradle is recommended. Some wheelchairs (primarily those with swing-away footrests) use foot stirrups behind the heels of each foot (see Fig. 78-2). However, for other wheelchairs, it is best to use a continuous strap behind both feet because the rider's feet sometimes come over the stirrups during active use. The frame should be selected and configured so that the feet sit firmly upon the footrests, with shoes on, without lifting the upper legs from the seat cushion. Footrests are commonly placed between 25 and 50 mm (1 to 2 in.) from the ground to ensure that sufficient ground clearance is maintained. Often, the footrests are the first part of the chair to come in contact with an obstacle (such as a door, wall, or another chair), so they must be durable.

Rigid wheelchairs often use simple tubes across the front of the wheelchair. By using a tubular rigid footrest, the wheelchair becomes stiffer and stronger (see Fig. 78-4). Rigid footrests are used during sports activities and work well for people who are very active in their wheelchair. Forward antitip rollers can be mounted to rigid footrests. This is helpful for playing court sports and reducing the risk of some forward-tipping accidents. Folding wheelchairs often use footrests that fold up and leg rests that swing out of the way to ease in transfers. Swing-away leg rests are not as durable as rigid ones. In some cases, manufacturers design swing-away leg rests that will flexibly bend on impact. This helps to absorb the energy of the impact and possibly prevent serious injury to the wheelchair rider. Elevating leg rests can be used for people who can not maintain a 90-degree knee angle or who need their legs elevated for venous return. Elevating leg rests make the wheelchair longer and heavier. This also has the effect of making the wheelchair less maneuverable by increasing the turning radius. Therefore, if elevating leg rests are needed, a power wheelchair should be strongly considered.

**Armrests and Clothing Guards**

Armrests provide a form of support and are convenient handles to hold onto when the rider leans to one side or the other.
Armrests are also helpful when attempting to reach higher places. For example, some people use their removable armrests as a tool to nudge items off high shelves. Armrests are commonly used to perform a “push-up” to assist with seat pressure relief. However, this is not the preferred method of pressure relief because of the significant stress it puts on the upper extremity (16).

There are three basic styles of armrests: wraparound, full-length, and desk-length. Wraparound armrests mount at the back of the wheelchair onto the frame below the backrest in most cases. The armrest comes up along the back of the backrest supports and wraps around to the front of the wheelchair. The major advantage of this design is that the armrest does not increase the width of the wheelchair like the other types of armrests. Wraparound armrests are popular among active wheelchair users. The most significant drawback of this design is that the armrest does not serve as a side guard to keep the rider’s clothing away from the wheels; although a removable piece of plastic called a clothing guard can be attached to the wheelchair frame to prevent clothes from getting caught in the wheels.

Full-length and desk-length armrests are similar in design, the main difference being the length of the armrest. Full-length armrests provide support for nearly the entire upper arm. They are popular on electric-powered wheelchairs because they provide a convenient and functional location for a joystick or other input device. Full-length armrests make it difficult to get close to some tables and desks. This is why manufacturers produce shorter desk-length armrests. Both of these types of armrests include clothing guards to protect clothing from the wheels. These types of armrests are mounted to the side of the wheelchair and may add as much as 5 cm (2 in.) to the width of the wheelchair.

Armrests can be fixed or height adjustable. Height-adjustable armrests may move up and down to accommodate the length of the rider’s trunk and arms. Most armrests can be moved in order to provide clearance for transferring in and out of the wheelchair and to allow a person to lean over the sides of the wheelchair. Armrests are either removed or flipped back. Both styles commonly use a latch, which is operated by the user. It is important to have secure latches on the armrests because armrests form convenient places for people to hold onto when attempting to provide assistance. If the armrests and latches are designed properly, two people can lift the rider and wheelchair by holding onto the armrests. In some cases, armrests are designed to pull out if any upward force is applied to them. These are not intended to be used for lifting the chair. It should be noted that armrests could alter the way in which a person propels a wheelchair. The hands and arms must clear the armrest in order to reach the push rim. This can force the user into excessive abduction at the shoulder, which could be a risk factor for injury.

**Wheel Locks**

Wheel locks act as parking brakes to stabilize the wheelchair when the rider transfers to other seats and when the rider wishes to remain in a particular spot. When locked, they keep the wheelchair stable to allow the rider to push things from the chair. There are a variety of wheel locks used to restrain wheelchairs when transferring or parking. High-lock brakes, which are located near the front corner of the seat, on the upper tube of the wheelchair’s side frame, are most common; however, location can vary. High-lock brakes require the least dexterity to operate. Extension levers can be added for people with limited reach or minimal strength. Wheel locks are standard equipment on wheelchairs, and they are simple to mount if the wheelchair does not come equipped with locks from the manufacturer.

Wheel locks may be push-to-lock or pull-to-lock. Most people prefer push-to-lock because wheel locks are more difficult to engage than to disengage. Riders often find it easier to push with the palm than pull with the fingers. Low wheel locks are usually mounted to the lower tube of the wheelchair’s side frame. Low wheel locks require more mobility to operate. They also alleviate the common problem that is seen with high wheel locks of the user hitting his or her thumb against the lock. This problem can be addressed for high wheel lock users by selecting retractable (i.e., scissors or butterfly) wheel locks. The retractable type of wheel lock helps to prevent jamming the thumbs and can also accommodate a wide variety of camber angles. The major drawback of retractable wheel locks is that they are more difficult to use than other types of wheel locks. The wheel lock must be positioned properly with respect to the wheel in order to operate effectively. If the wheels are repositioned, then the wheel locks must be repositioned. Tire pressure also affects the locking grip of these wheel locks. Many active manual wheelchair users will choose not to use armrests or wheel locks. Although these devices can be convenient, any equipment added will increase the overall weight of the chair, making propulsion more difficult.

**Tires**

A variety of options exist with respect to tires. The most common type of tire is pneumatic. These tires are lightweight and provide cushioning against impact and vibration from rolling over surfaces. This cushioning may increase rider comfort and improve wheelchair durability. Pneumatic tires are recommended for outdoor usage. The main downside of pneumatic tires is that they require maintenance and they can puncture. Tire pressure needs to be kept at a predetermined level because it is critical to rolling resistance, which can be related to risk of secondary injury associated with manual wheelchair use. A study has shown that when propelling on tires that were deflated by 50%, energy expenditure increased by 25% (17). Clinicians involved with wheelchair users should squeeze their patients’ tires to assure that they are keeping up with this important regular maintenance issue.

An alternative to pneumatic tires is solid inserts. These foam inserts fit into the pneumatic tire and replace the air-filled inner tube that would normally be there. They add some weight to the chair and may slightly worsen performance but are a good alternative for individuals who do not want to be
responsible for maintenance of air pressure. A less viable alternative is solid tires. These tires require no maintenance and are low in cost. Unfortunately, they make for an uncomfortable ride as all ground shocks are transmitted to the wheelchair user.

Additional Features
Many additional features are available that are unique to manual wheelchairs.

Antitippers
Antitippers are often placed on wheelchairs to assure they do not tip over backward. These can inhibit the ability to climb curbs, but they do offer a measure of safety. It is suggested that these be ordered for all wheelchairs and then have the user take them off when they are comfortable with the stability of the chair. A new antitipper is currently being developed but not yet commercially available that allows the wheelchair user to perform a wheelies and other skills. This device is folded up to allow the wheelchair user to perform his or her necessary wheelchair mobility skills; however, if the wheelchair user falls too far backward, the antitippers will extend and prevent the wheelchair user from falling backward (18).

Push Rims
A number of different push rims are currently available, and new styles are likely to be introduced into the market. Anodized aluminum rims are the current standard on most K0004 and K0005 chairs. Less expensive chairs may come with plastic push rims. For individuals with difficulty gripping the rim, alternative rims should be considered. These can include vinyl-coated rims, rims with projections (Fig. 78-5), and rims wrapped with surgical tubing. All of these rims have the advantage of increased friction, making it easier to push the chair forward. Unfortunately, this increased friction can lead to burns when the wheelchair user attempts to slow down the chair. There is ongoing research into advanced ergonomic push-rim designs that allow for easier propulsion as well as braking (19). For individuals who have good hand function, a push rim called the “Natural-Fit” (Three Rivers, LLC, Mesa AZ) (see Fig. 78-4) is an ergonomically designed hand rim that has been shown to decrease stresses placed on the upper extremity during wheelchair propulsion (19).

Wheels
The standard wheels on lightweight and ultralightweight wheelchairs have spokes. Plastic wheels can be used on low-end chairs. These increase weight and decrease performance. High-end wheels are now available with flexible spokes or with graphite and composite materials. A study performed by Hughes et al. (20) found that high-end wheels did not improve efficiency but did significantly improve comfort of the rider. These wheels can be easier to maintain than standard spokes and also offer improved aesthetics.

Caster Wheels
Caster wheels are available in a variety of shapes and sizes. Pneumatic wheels are larger than solid casters and may interfere with the footrests when turning. Pneumatic wheels offer the advantage of easier propulsion over rough terrain and increased shock absorption. Many wheelchair users are using narrow rollerblade-type wheels which allow quicker turns and reduce rolling resistance. Unfortunately, they can get caught in sewer grates or other small obstacles. Adding an elastomer shock absorber “frog-legs” in series with the caster can provide a means of shock adsorption vibration reduction (15) (Fig. 78-6).
Push Handles

Push handles, also known as canes, are attached to the back of the chair with the primary purpose of making it easier for an assistant to propel the wheelchair. Canes can also be used by the wheelchair occupant to help with pressure relief. The wheelchair user can hook the arm around the cane and pull to raise the contralateral hip, even in the absence of strong triceps muscles. Canes can also be used to hang a book bag or backpack onto the back of the chair.

Grade-Aids

Grade-aids are devices that attach to the wheel and allow it to roll forward but not back. In this manner, they can make it easier to roll up a hill. These should be considered for individuals who have upper-limb weakness and who must negotiate hills.

Manual Wheelchair Selection and Setup

As stated previously, a number of the features described above and the adjustments described below are available only for ultra-lightweight wheelchairs. Best practice dictates that any individual who will propel a wheelchair as a primary means of mobility should receive this type of chair. The optimal setup for a manual wheelchair is described in the Clinical Practice Guideline (CPG) for preservation of upper limb function (16) (Fig. 78-7). The chair should be as narrow as possible without causing undue pressure on the thighs. Rear axle position should be adjustable to fit the user, as it can affect caster flutter, rolling resistance, stability, and control and maneuverability. With a longer wheel base or more rearward axle position, the chair has more stability; however, rolling resistance is increased (21), caster flutter is increased, and downward turning on side slopes is increased. These changes are primarily related to the proportion of the weight that is placed onto the back, or larger, wheels. As the rear wheels are moved backward, more weight is placed on the front casters. These smaller-diameter wheels have higher rolling resistance. The cadence of the propulsive stroke is also higher, with more rearward axle position. This has been implicated in relating to risk of repetitive strain injury. A forward axle position has been found to be associated with improved propulsion biomechanics (21). Given these issues, the axle should be as far forward as possible, providing that the wheelchair user still feels stable. Most wheelchairs come in a factory-set position, with the axle most rearward. This should be gradually adjusted forward to maximize performance of the chair.

Axle height, or the distance between the shoulder and the axle, is an important parameter that is also described in the CPG (16). If the seat is too high, the wheelchair user will not be able to reach much of the push rim and so they will push with shorter strokes and a faster cadence. If the seat is too low, the user will be forced to abduct at the shoulder during the propulsive stroke, which may cause rotator cuff impingement. In general, while sitting upright with the hands resting on the top of the wheels, the elbow angle should be between 100 and 120 degrees for optimal mobility (22,23). Alternatively, if the arms are left to hang freely at the side while sitting on the chair, the fingertips should be just past the axle of the wheel.

Camber describes the angle of the wheel with respect to the chair. Increasing camber has several advantages: the footprint of the chair is widened, creating greater side-to-side stability; it allows quicker turning; and positions the push rims more ergonomically for propulsion (it is more natural to push down and out). In addition, by having a wider based, the area where the hands are in contact with the push rims is less likely to come into contact with the wall (24). Finally, adding camber to the rear wheels reduces effective stiffness between the rolling surface and frame, thus reducing the vibration exposure of the user.

A disadvantage to increasing camber is also that it adds width to the chair. Generally, for daily use, the chair should be as narrow as possible without substantially diminishing the handling characteristics. The wheels should be offset enough from the seat to avoid rubbing against the clothing or body. Between 2 and 4 degrees of camber is appropriate for everyday use.

There are many other aspects of a wheelchair that can be adjusted to improve fit and performance. Some of these adjustments are discussed in the seating section that follows.

Alternative Manual Wheelchairs

Two alternatives worth mentioning are chairs for amputees and chairs for individuals with hemiparesis. Wheelchairs for individuals with amputations are typically designed with the rear axle set far behind the user. This is needed because the absence of a leg causes the body’s center of gravity to be shifted posteriorly, thus reducing rearward stability. Unfortunately, all of the negative aspects of a rearward axle are present. An alternative can be to add weight to the front of the wheelchair. Unfortunately, increased weight means increased rolling resistance. There is no simple answer as to what is best, and individual patients should make this decision for themselves.

For individuals with hemiparesis or other disability that makes propulsion with a leg or both legs superior to propulsion with the arms, a “hemihigh” chair is an alternative. In this chair, there is typically one footrest or none at all, and the seat is low enough to the ground so that the feet can reach the floor. For an individual with hemiparesis, the use of the
uninvolved arm and leg can provide limited, but functional, propulsion.

Another alternative for an individual with hemiparesis is a one-arm drive chair. This chair has two push rims on one side that control separate wheels. One-arm drives are heavier than standard chairs and can be difficult to control. These limitations make them less than ideal for an elderly stroke patient. Although not popular in the United States, lever-drive wheelchairs are seen in Europe. These chairs offer a mechanical advantage by using a lever to make propulsion more efficient (25). Unfortunately, they do not offer the direct proprioceptive feedback of hand rim contact and can be difficult to maneuver in tight spaces and when traveling backward. Lever-drive chairs are heavier and often wider than standard chairs.

**Manual Wheelchair Propulsion Technique**

In recent years, much has been learned about the most appropriate way to propel a wheelchair. This has been learned by studies that have found an association between upper-limb injuries and wheelchair propulsion biomechanics (21). Manual wheelchair users should propel with long smooth strokes that minimize the cadence with which they push and maximize the length of the stroke or contact angle (16,21). The wheelchair user should attempt to impact the rim smoothly and match the speed of the hand to the rotating speed of the push rim. During the recovery stage of the propulsive stroke, the user should let the hand drop below the push rim and stay below the push rim until he or she is ready to begin propulsion again (26) (Fig. 78-8).

Wheelchair propulsion techniques can be evaluated using devices such as the commercially available SmartWheel (Three Rivers Holdings, LLC, Mesa, AZ). This is the only device on the market that can objectively evaluate wheelchair propulsion and is used at numerous centers. The device replaces one of the patient’s wheels on his or her own wheelchair. The individual can then propel over various surfaces that he or she would encounter in normal life. The wheel measures parameters including peak force, push frequency, and contact angle. A database has been created in which a clinician can compare his or her patient’s wheelchair propulsion performance against a larger group of wheelchair users (27). This device can also be helpful when making a decision if an individual should be using a manual wheelchair user (MWC) or a power wheelchair user (PWC), to evaluate propulsion mechanics, and as a wheelchair propulsion training tool.

**POWER WHEELCHAIRS**

There is strong evidence in the literature to indicate that the use of powered mobility facilitates independence, improves occupational performance, and is correlated with a higher sense of quality of life for people who cannot ambulate or propel a manual wheelchair effectively (2,4,6). However, there are some disadvantages to powered mobility devices that need to be factored into the decision of powered versus manual wheelchair mobility (Table 78-3). Power wheelchairs come in a number of configurations. Two types are those with a power base and those with integrated seating systems. In general, wheelchairs with integrated seating systems are less expensive and offer less with respect to seating options.

**Power Bases**

The power base is the lower portion of a power wheelchair that houses the motors, batteries, drive wheels, casters, and electronics to which a seating system is attached (Fig. 78-9). The base allows for the mounting of any variety or combination of

![FIGURE 78-8. Recovery pattern. This represents data from a wheelchair user. The dark circle is the push rim. The thinner lines are from a marker placed on the wrist. This recovery pattern was found to provide the best efficiency. Direction of travel is left.](image-url)
different seating systems and seat functions, including tilt-in-space, reclining back, elevating leg rests, and seat elevator.

**Drive Classification**

Power wheelchair bases can be classified as rear wheel drive (RWD), mid-wheel drive (MWD), and front wheel drive (FWD). The classification of these three drive systems is based on the drive wheel location relative to the system's center of gravity. The drive wheel position defines the basic handling characteristic of any power wheelchair. Each system has unique driving and handling characteristics. In RWD power bases (Fig. 78-10), the drive wheels are behind the user's center of gravity, and the casters are in the front. RWD systems are the traditional design, and therefore many long-term power wheelchair users are familiar with their performance and prefer them to other designs. A major advantage of an RWD system is its predictable drive characteristic and stability. A potential drawback to an RWD system is its maneuverability in tight areas because of a larger turning radius.

In MWD power bases (see Fig. 78-9), the drive wheels are directly below the user's center of gravity and generally have a set of casters or antitippers in the front and rear of the drive wheels. The advantage of the MWD system is a smaller turning radius to maneuver in tight spaces. A disadvantage is a tendency to rock or pitch forward, especially with sudden stops or fast turns. When transitioning from a steep slope to a level surface (like coming off a curb cut), the front and rear casters can hang up, leaving less traction on the drive wheels in the middle. However, manufactures have addressed these shortcomings and have updated designs of MWD that now are available with a wheelbase consisting of six wheels touching the ground; with two small wheels in the front and back. These wheels are equipped with suspension and shock absorbance systems that enhance forward stability and small obstacle climbing capability.

A FWD power base (Fig. 78-11) has the drive wheels in front of the user's center of gravity, and the rear wheels are casters. The advantage of an FWD system is that it tends to be quite stable and provides a tight turning radius. FWD systems may climb obstacles or curbs more easily as the large front wheels hit the obstacle first. A disadvantage is that an FWD system has more rearward center of gravity; therefore, the system may tend to fishtail and be difficult to drive in a straight line, especially on uneven surfaces.

**FIGURE 78-9.** Mid-wheel-drive power wheelchair with power base. The seating system can be removed and replaced with a different seating configuration.

**FIGURE 78-10.** Rear-wheel-drive power wheelchair.

**FIGURE 78-11.** Front-wheel-drive power wheelchair.
Controls
The device that is used to control a powered mobility system is called an access device. The primary function of the access device is to drive the chair. Secondary applications are for use with environmental control systems and computer access. By using the same control interface for both the power wheelchair and the secondary systems, the seating position and control selection can be optimized for multiple functional purposes. The majority of input controls are programmable, allowing changes in speed and the amount of movement to determine the direction of the wheelchair. Many power wheelchair users brace some portion of their hand against the control box and use their hand and arm coordination to operate the joystick. Gross arm function, in many cases, can be used to operate a joystick. If the user does not have the hand function or coordination to operate a joystick input device, other options are available. Other parts of the body, such as the chin or foot, can operate a modified joystick.

The cognitive abilities required to operate an alternative control typically increase as the physical and functional abilities of the user decrease. Programmable wheelchair controllers allow reduction of the maximum velocity and modification of the acceleration and deceleration rates of the wheelchair. To assist persons with more severe cognitive or visual limitations, technologies are being developed that enable the wheelchair to follow walls, navigate through doorways, and stop when other objects are contacted (28). For persons with spasticity, particularly tremor in the hand, modern controllers have filters that can be adjusted to give smoother wheelchair control. Positioning technology enables the joystick to be placed in a variety of locations to optimize the user’s ability to operate it (29).

Joysticks
Joysticks are the most common access devices for powered wheelchair systems. Most joysticks are proportional; meaning the control’s speed output to the wheelchair from the joystick is in proportion to how far the joystick is pushed from the center position. Research is being conducted on different input methods to control a joystick, one of which is an isometric joystick (30). Joysticks can be fitted with a template that only allows motion in certain directions. This is useful for people with poor motor control, such as can be seen in athetoid cerebral palsy.

The end of the joystick can also be modified for easier grip. One common modification is a goalpost shaped attachment (T-bar) with the upright section of the joystick on either side of the hand. Alternatively, a simple tennis ball cut to fit over a traditional joystick can make gripping easier.

Mini Joystick
The Mini Joystick, a type of micro proportional joystick, maybe a good alternative for a user who has lost gross arm function, and is unable to use a conventional joystick but still possesses small distal motor function within the fingers. The Mini Joystick has the same proportional and directional features of a conventional joystick; however, the tuning and the physical displacement of the joystick is scaled down and requires significant less effort to achieve small joystick displacement (i.e., wiggle of the tip of a finger). Its smaller size also permits more freedom in its placement to meet the functional/comfort needs of the user (i.e., custom fitted on a stable arm support for independent single finger control or custom fitted through the “horn” of a palmer support for thumb control). Its tuning and custom programmability capabilities allow head control drive option when faced with complete loss of upper extremity and lower extremity functions. Used as a chin interface, it requires minimal effort of facial and cervical musculature to operate the control.

Sip-and-Puff
Sip-and-puff switches are used primarily by people with tetraplegia. A sip-and-puff access device consists of a replaceable straw located near the mouth. Pulling and pushing air through the straw with the mouth controls the wheelchair. These systems can be set up in a variety of configurations. Generally, the user will sip a specific number of times to indicate a direction and puff to confirm the choice and activate the movement of the wheelchair. It is common for an auxiliary visual display to be used with sip-and-puff to provide feedback to on which command was selected. Sip-and-puff is becoming somewhat antiquated now that microproportional joysticks, head array sensors, and gyroscopic-based joysticks are available.

Switches and Buttons
An array of switches or a single switch can be used to control a chair. The more switches, the greater the motor control needed to operate the equipment. Using scanning, a single switch can control a wheelchair. In scanning, choices are presented to the user on the wheelchair controller. When the desired choice is presented, the user activates the switch to select that command. Two switches can fully control a wheelchair in a fashion similar to the one described above for sip-and-puff systems. An array of switches can be used for quicker control. There is no proportional control of the wheelchair’s speed with switched control. Thus, users are forced to travel at the same, rather slow velocity at all times.

Many of the input devices described in this section can be used for more than the control of the wheelchairs only. Two common additional uses are environmental control and computer access. In some cases, the computer is mounted to the wheelchair, but often it is at a fixed site. A critical consideration when selecting or designing a user interface is that the ability of the user to accurately control the interface is heavily dependent on the stability of the user within the wheelchair. Often custom seating and postural support systems are required for a user interface to be effective. The position and stability of the user interface are also critical to its efficacy as a functional access device.
Considerations in Selection of Electric-Powered Mobility

There are many considerations in defining the correct power wheelchair for a given individual. Important factors include the means of transportation of the wheelchair surface conditions, the chair must negotiate, need to negotiate thresholds and curbs, and clearance widths in the environment. In addition, subject preference is important, as well as maximum speed and range. Typically, a wheelchair will be used on a daily basis and, as such, has a life expectancy of 3 to 5 years.

SEATING AND POSITIONING

Seating systems can be organized into three general categories: off-the-shelf systems, modular systems, and custom-molded systems. Overlap exists between the categories, and a given seating system should be prescribed and designed specific to the user’s medical, functional, and personal preference needs. The simplest form of seating system is a linear seating system that involves a planar seat and back with fixed angles and orientations.

Medically, a system should address issues of soft-tissue management, comfort, reducing the potential for or accommodating orthopedic deformities, and maintaining vital organ capacity. Functionally, the system should address the movements and supports the user needs to reach or access objects, sit, get under tables, and perform activities of daily living. The chair must become an extension of the user’s body, much like an orthosis. This requires careful matching of critical chair dimensions to body dimensions, user abilities, and intended uses. The user’s preference as to one system over another should be paramount in the prescription process. For example, a user may choose to forgo pressure relief and comfort for a firmer seating system that provides greater stability and allows him or her to have a firm surface to slide on the seat for ease with repositioning when transferring in/out of the device.

Soft-Tissue Management

Soft-tissue management is a concern for all people who sit for prolonged periods of time and who have compromised sensation or the inability to perform weight shifts. External causes of skin breakdown or pressure sores include excessive prolonged pressure over bony prominences, friction and shear, as well as heat and moisture. Intrinsic factors include the inability to move, poor nutrition, vascular problems, and the loss of soft-tissue elasticity (31). The loss of sensation is a key factor because discomfort is the usual trigger for shifting and moving. Because the causes of pressure ulcer vary, the choice of seat cushion will vary based on the client’s risk factors and the characteristics of the cushion.

The fit of the wheelchair also contributes to pressure distribution. Footrests mounted too high increase the pressure over the ischial tuberosities. Properly adjusted armrest height allows weight to be distributed through the upper extremities. The angle of the back, relative to the seat, affects the stability and orientation of the pelvis and is reflected in how much a person will slouch, and the slouch posture affects pressure over the sacrococcygeal regions.

Material Properties of Seating Systems

Cushions are chosen based on their characteristics, which are related to the properties of the materials used in their construction. Materials specific to those used in the design and manufacture of seating systems have certain characteristics, as shown in Table 78-4. Manufacturers make cushions that possess these qualities using flat and contoured foams, air-filled bladders, combinations of air and foam, flotation, viscous fluids, contoured plastic honeycombs, custom-contoured foam, and alternating pressure systems. These cushions vary in efficacy of pressure distribution, provision of postural stability, ability to insulate or conduct heat, and the reliability of their performance over time. Finding a cushion with good airflow and pressure distribution would be important for an immobile client who perspires heavily or is incontinent. Alternately, for a client prone to pressure ulcers, a practitioner would identify cushions with optimal redistribution of peak pressures. Active manual wheelchair users may not like an air-filled cushion because it does not provide a stable base for propulsion-related activities. If all the needed features cannot be found in one cushion, trade-offs are necessary. Research evidence supports that a properly fitted pressure-reducing cushion, in contrast to a low-cost foam cushion, reduces the probability of a pressure sore (32).

<table>
<thead>
<tr>
<th>Property</th>
<th>Application</th>
</tr>
</thead>
<tbody>
<tr>
<td>Density: The ratio of mass or quantity of material to the volume of the cushion</td>
<td>A cushion filled with air will be much lighter than one composed of gel. Weight may be a problem for someone who has to lift it or propel a manual wheelchair.</td>
</tr>
<tr>
<td>Stiffness: The strength of the resistance to compression</td>
<td>Foam has low stiffness and does not resist body weight compared with a solid seat base. A solid surface provides greater pelvic stabilization; foam may allow better pressure distribution.</td>
</tr>
<tr>
<td>Thermal characteristics: The ability of the material to insulate or conduct heat</td>
<td>Dense foam cushions retain body heat. Honeycomb-designed cushions hold less heat. Gel and fluids tend to pull heat away from the body.</td>
</tr>
<tr>
<td>Friction: The ability to maintain position and to reposition if needed</td>
<td>Cushions with solid bases and slick covers make sliding in lateral transfers easier but may promote sacral sitting.</td>
</tr>
</tbody>
</table>
Pressure Mapping
Pressure-mapping technology estimates interface pressure. A thin mat with pressure sensors is placed between the client and the seating surface. The mat connects to a computer and presents data in both graphic and numeric forms. As part of a skilled clinical assessment, it can be predictive of potential risk for pressure sores (33). This technology can help a clinician decide which cushion provides the best pressure distribution for a particular client. It is important to remember that pressure-mapping devices do not measure shear forces, heat, moisture, postural stability, or maintenance of the cushion. These factors must also be considered.

Custom-Contoured Seating
Custom-contoured seating systems are necessary when all available off-the-shelf or modular seating systems cannot address the needs of the individual. This may occur in individuals with moderate to severe fixed and semiflexible structural deformities of the spine and extremities. In addition, individuals who require significant off-loading of soft tissue because of pressure sore issues may need a custom-contoured seat.

Custom-contoured seating involves a process of capturing a specific mold of a person’s body. The mold can be obtained through several methods, including liquid foam in place, plaster molds, or through computer aided design/computer aided manufacturing (CAD/CAM) with greater control and accuracy. CAD/CAM usually involves the use of a seating simulator composed of bead bags for the seat and backrest supports. The bead bags are manually and gravitationally contoured around the shape of the person’s deformities or pressure points, followed by vacuum evacuation of the air from the bags to produce a rigid mold. This mold is then scanned, using sensors that send data to a robotic milling machine for the production of a positive mold.

Custom-molded systems are not capable of correcting deformities. Careful skin inspections should be performed, and pressure-mapping systems should be used as appropriate to verify that the custom-molded seat contours are applied properly. Inappropriately applied contours can lead to pressure sores. Careful consideration of transfer technique is needed with these seating systems because proper positioning in the seat is essential to performance, and the custom contours usually make transfers more difficult.

Back Support
A back support should conform to the normal spinal curvature while allowing movement as required by the user. The typical back support in a folding-frame wheelchair is sling upholstery, not because it is good back support, but because it bends to allow the chair to fold. Sling seats provide little in the way of support. The sling-back support stretches before a propulsive force is effectively applied to the rim, resulting in inefficiencies during wheelchair mobility. In individuals with tetraplegia, the stretching of the sling back can mean that the wheelchair user adopts a more posterior tilt of the pelvis, and this may contribute to a kyphotic spine (12). Like cushions, wheelchair backs are chosen based on the client’s seating goals. For clients with truncal weakness, the stability from a contoured backrest with or without modular lateral supports is needed to maintain head and neck position. Some clients may only need the soft contouring of an adjustable tension sling backrest, whereas others with significant kyphoscoliosis may need a custom-molded backrest to enable sitting in a more upright posture. Clients with this level of weakness or deformity will most likely use this seating in a powered mobility base or an attendant-operated base.

Recline and Tilt-in-Space
Recline and tilt-in-space technologies relieve pressure, manage posture, provide comfort, and help with personal care activities. Recline helps stretch hip flexors and also assists with attending to catheters, toileting, and dependent transfers. Because reclining the seat back creates shear, the user often shifts in the wheelchair into a sacral-seated position. For a patient who is unable to reposition without help, adding tilt may help the user reposition independently. Tilt-in-space keeps the hip and knee angles constant when tilting the client back. Unlike reclining systems, the position of the user is maintained in the tilt seating system. It is essential that subjects who are unable to independently shift weight, unable to transfer independently, or have pain as a result of prolonged sitting, have a tilt-in-space and recline system on their chair (34). Tilt-in-space is also necessary for individuals with progressive disorders (34). An individual with amyotrophic lateral sclerosis may find it easy to perform weight shift and repositioning at an initial evaluation, but this can change quickly, leading to the need for modifications. Tilt and recline are also available in manual wheelchairs. For the most part, these chairs are only used for patients who require attendant control. As in recline, tilt-in-space greatly reduces pressure on the ischial tuberosities by shifting the pressure to the back. Several recent research studies support the combination of tilt-in-space and recline together. Using a combination of tilt-in-space and recline will maximize the function of both devices, enhance the ability of the individual to use the system for positioning, perform personal care tasks, reduce pressure and increase comfort (34).

Other Positioning Systems
Stand-up wheelchairs are an alternative that deserves mention. The benefits of standing for individuals normally unable to do so may include decreased bladder infections, reduced osteoporosis, and decreased spasticity (35). In addition, there are likely psychological benefits that result from the feeling of upright posture and the ability to interact at eye level (36). Certain people cannot use a stand-up wheelchair because they do not have adequate joint range of motion. Some of the benefits of a stand-up wheelchair can be obtained by using a variable-seat-height wheelchair. The most common function of variable-seat-height wheelchairs is to provide seat elevation. A seat elevator has many useful functions that will improve the quality of life of the user. The device can help individuals avoid positioning their hand above the shoulder which can
help to reduce shoulder pain (37). The individuals will be able to reach heights that would have been impossible in the past, thus increasing independence. By allowing the individual to change his or her seat height, an easier, down hill, transfer can be performed. A downhill transfer may be beneficial because it can decrease the forces placed on the upper extremity (38). Finally, a seat elevator will allow the individual to interact with peers at eye level, thus improving the individual's ability to interact in a social situation. Seat elevators are strongly recommended by the CPG on preservation of upper limb function (16) and have been advocated by RESNA (39). Finally, some newer chairs offer lateral tilt-in-space. This feature allows the user to be leaned to either side and offers an alternative for people with difficult pressure sore or pain issues.

Seating Setup
Setting up the wheelchair is critical to optimizing performance. A therapist or rehabilitation supplier working with the clinic commonly performs this. Seat height can be adjusted on most chairs. The seat height is dependent on the total body length of the user. Users with longer leg lengths will require higher seat heights to achieve sufficient clearance for the footrests. There is some flexibility when selecting seat height, even for taller individuals, because most active users prefer some seat angle or dump. Dump is achieved by tilting the seat down toward the backrest, thus the end closest to the backrest is lower than the front of the cushion. Dump allows the user to fit more securely in the chair, increasing the user's trunk stability and making the chair more responsive to the user's body movements. Seat depth is determined from the length of the upper legs. Generally, no more than a 75 mm (3 in.) gap should be between the front of the seat and the back of the knees when the person is in the wheelchair. This will help ensure broad distribution of the trunk weight over the buttocks and upper legs, without placing undue pressure behind the knee. Some gap is required to allow the user some freedom to adjust his or her position. Seat width is determined from the width of the person's hips, the intended use, and whether the person prefers to use side guards. Generally, the wheelchair should be as narrow as possible; thus, a chair about 1 in. wider than the user's hips is desirable.

Power-Assisted Wheelchairs
A developing class of wheelchairs provides a power assist when desired, but allows the user to push the wheelchair as one would with a manual wheelchair. Power-assisted wheelchairs (Fig. 78-12) have force/moment-sensing push rims that provide an additional torque to the rear axle proportional to the applied moment. Such devices have potentially important benefits, such as allowing individuals to perform tasks faster and easier. In addition, people who normally would need to use a power wheelchair are able to self-propel a power-assisted wheelchair despite obstacles such as steep ramps (40–42). Finally, when using a power-assisted wheelchair, oxygen consumption and heart rate are significantly lower (43). For individuals with upper-limb pain or tetraplegia, the power-assisted wheelchair may prove to be a good compromise between a manual wheelchair and a power wheelchair. Before a power-assisted wheelchair is recommended to a client, it is essential that the client tries the device in his or her actual environment. Subjects may report that they encountered difficulty transferring from the chair equipped with power assist and in disassembling the power-equipped chair for transport. One study has shown that performing wheelchair skills that require great control, such as performing a wheelie or navigating the chair in small spaces, is more difficult with a power-assisted chair (44). Another study of individuals using power assist in their real-life environment did not show a significant difference in the amount of time individuals used the device and the distance they traveled (45). It is important to remember that under current Medicare guidelines, a power-assisted device will not be funded unless the patient has used a manual wheelchair for 1 year. This should not prevent recommending the device, but must be kept in mind when producing the LMN.

Scooters
Scooters (Fig. 78-13) are another option for certain individuals. These devices typically have a single front wheel for steering and two drive wheels in the back. Possibly because scooters are available to help shoppers in many large department stores, they seem to have a greater degree of social acceptability than wheelchairs. This leads many clients to request these devices. Steering is accomplished via hand bars that are intuitive to users who have previously used a bicycle. Seating is provided in a chair having foam padding typical of a car seat. The backrest height ends at the level of the shoulder blades, which allows for unencumbered rotation of the trunk.

Scooters have a number of advantages and disadvantages that must be critically considered when prescribing this device.
In general, scooters are a reasonable option for individuals who retain some ability to ambulate, such as those with cardiopulmonary disease limiting the ability to walk. Scooters are a poor option for individuals with progressive neuromuscular disorders because they have few options to accommodate progressive disability. They are also a poor choice if one needs to stay in the chair all day because seating options are limited. If a scooter is found to be the appropriate device, it is important that the scooter is trialed in the patient’s home environment. Scooters require a large turning radius and often do not fit well into a person’s home.

### Power Wheelchair Training

New wheelchair users should go through a wheelchair-training program. Wheelchair users can first practice basic maneuvers in controlled environments, then transition to uneven surfaces and slope transitions (i.e., level to sloped surfaces) in both the uphill and downhill directions and maneuvering through tight environments. Once these skills are mastered, they should gradually tackle more challenging environments, such as steep grades and step transitions that exceed Americans with Disabilities Act (ADA) Accessibility Guidelines. The rider should always practice with an appropriate lap belt (46) and chest support in place, and a spotter (therapist) to assist if needed.

### PEDIATRICS

In many ways, the seating and mobility needs of children with disabilities are similar to those of adults. Comfort, stability, and function are paramount, and there is no one perfect position for every activity or situation. There is a tendency to be aggressive with children and force them to sit upright with many postural supports in the hope of preventing or delaying deformities. This tendency in the design of the seating system may force postures that are not tolerated or desired over time. It is important to seek input from both the child and the caregivers who will be dealing with the seating system on an hourly basis. Several variables in system design need to be considered with children, including developmental status, mobility, growth, age-appropriate activities, school, therapies, the environment, and family issues. Children not only change in size but their disability also often changes as they grow, even in nonprogressive disorders such as cerebral palsy.

### Pediatric Seating Systems

Pediatric seating systems can be classified in the same manner as outlined for adults in the previous section, with off-the-shelf, modular, and custom-molded systems available. The issue of growth should not be overvalued over function or other points of concern. There are some systems that can accommodate growth to a certain degree; however, these modular systems tend to be heavier and bulkier than the lighter, more compact, off-the-shelf systems or custom system that cannot change with growth. Clinicians need to consider the impact on current function of a heavier system and make their recommendations accordingly. It is possible to justify a new chair in a shorter period of time based on growth.

In certain cases, such as when a child lacks muscle control, modular systems can offer advantages. Trunk lateral supports may need to be removed while the child is engaged in a dynamic reaching activity or in therapy working on trunk control, but replaced to sit more passively when he or she fatigues. A child should also be allowed to passively sit when focused on other activities, such as schoolwork, or relaxing while riding in a vehicle or watching television. If the child has to focus all his or her energy on balance and stability, then he or she will have no energy to devote to the task at hand. If a seating system is too confining or restrictive with multiple supports, the child may fight the system to be able to move and be dynamic. Thus, the pediatric seating specialist must find a balance between appropriate amounts of support without overly restricting movement. An anterior chest harness or support may be needed for stabilization during transportation but removed to engage in schoolwork. It is important to educate

<table>
<thead>
<tr>
<th>TABLE 78.5 Scooters</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Advantages</strong></td>
</tr>
<tr>
<td>Lower cost</td>
</tr>
<tr>
<td>Easier to assemble and disassemble for transportation</td>
</tr>
<tr>
<td>Better than some chairs at rough terrain</td>
</tr>
<tr>
<td>Less perceived stigma of disability</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
all parties as to when a certain support is appropriate and when it is inappropriate. Inappropriate use of chest harnesses has resulted in strangulation when the child slid down in the seat because he was becoming uncomfortable and trying to fight the restrictions of the seating system (47). It is often difficult to find a balance between providing enough support and still allowing for some freedom of movement. Having systems that are modular and flexible, as well as educating caregivers, can help find this balance.

Custom-molded are indicated when a child’s postural deformities are so severe that they cannot be supported by a modular or off-the-shelf seating system. Custom-molded systems are not as common with children as they are with adults, as children’s deformities tend not to be as severe. As with off-the-shelf systems, custom-molded systems have no growth capabilities.

Mobility

Mobility is the precursor to all childhood development. Children need to explore their environment to know where things are and how to get them. The need for specific types of mobility bases, such as strollers, self-propelled systems, and powered systems, will depend on the age and the physical, developmental, and functional capabilities of the child, as well as the environmental and transportation resources of the family. Very young children may need a seating system that can be transferred and attached to a variety of bases for different activities. For play, children want to be close to the floor near their toys and peers. For eating, they need to be up where a caregiver can have access. Some manufacturers make seating systems that can be transferred to various mobile and stationary bases that are height adjustable. For example, one seating system can be transferred between a folding-style stroller type base, a higher eating base, and a power-wheelchair base and can also serve as crash-tested car seat.

There has been a historical tendency to push children with disabilities to ambulate with braces and other aids. The practitioner, however, needs to consider the mobility needs of the child, including the expected surfaces and distances. A manual or power wheelchair may be more efficient and actually provide greater independence. Children often abandon cumbersome upright ambulation in favor of a wheelchair when they are able to make their own choice. Consider a child with cerebral palsy trying to ambulate with braces and crutches as well as carry a bag full of books through a crowded hallway. Manual wheelchairs are sometimes a useful option for children; however, the weight of the chair may be an issue. Even an ultralightweight manual wheelchair weighing 20 lb is going to be heavy for a 30 to 40-lb child. Proper chair selection and training are essential to the long-term health of the child.

Powered mobility is important for children who cannot effectively self-propel a manual system; however, families often do not have the resources or psychological readiness either to make the necessary home modifications or to purchase an accessible vehicle that can transport a powered system. Research and clinical intuition indicate that children without cognitive disability should be offered mobility devices such as power wheelchairs or adapted ride-on power toys as early as 12 to 18 months—about the time able-bodied children begin moving on their own (48–50). Even with a power wheelchair, there is often the need for a transportable folding supportive stroller-type base, as described previously, for use in family outings.

Psychosocial and Family Issues

There are many psychosocial and family-related issues practitioners need to be cognizant of and respect. Having a child with a birth or acquired disability can devastate a family. Initially, there tends to be focus on the search for a cure to the problem, and the family may be a reluctant to accept use of a wheelchair, place a ramp in front of the home, or purchase a van with a ramp. The practitioner needs to counsel the family regarding the realistically expected short- and long-term outcomes of the child’s situation. There is also a tendency to “care for” a child with a disability, especially in certain cultures and religions. It is common for children with disabilities to develop learned helplessness and for parents to develop a codependency in the relationship with their child (51). Later in life, this can result in lack of ability to make decisions or function in society. Children with disabilities need to experience the same stages of development—including successes as well as failures—as their able-bodied peers within the norms and values of that culture. Practitioners again need to counsel families so that they can effectively use this equipment to facilitate development and promote active participation by the child as appropriate based on their current and potential capabilities.

Standards and Durability

International standards are applied to mobility devices. The standards are formulated by ISO, the International Standards Organization; the American National Standards Institute; and by RESNA (52). The standards concern many aspects of wheelchairs, including electrical systems, durability, dimensions, flammability, strength of armrests, ability to withstand an impact, and stability, to name a few. These standards can be requested from the manufacturer and can serve as a method of comparing classes of wheelchair as well as individual wheelchairs to each other. Wheelchair users should be confident of the structural integrity of their wheelchair. Failure of any component is more than an inconvenience for the wheelchair users—it is the limitation of their mobility and can be life-threatening.

Tests included in the standards protocol include the curb drop and double drum. In the curb drop, the wheelchair and a 70-kg (150 lb) test dummy are lifted 5 cm and dropped to the ground. In the double-drum test, the wheelchair and test dummy are placed on two independent rollers (rear wheels on one roller and the front on the other). Each roller has two 1 cm high by 2 cm wide slats that simulate door thresholds. The standards require that the chair survives 200,000 cycles on the double drum and 6,666 drops.
A study by Fitzgerald et al. (53) compares the results of ISO standards testing on three classifications of manual wheelchairs. The classes followed Medicare definitions of K0001, depot or hospital-type (54); K0004 lightweight (55); and K0005, ultralight (56). Using ISO equivalent number of testing cycles, curb-drop and double-drum tests were expressed as a single variable, and Kaplan-Meier survival curves were determined. The fatigue life of ultralight wheelchairs was significantly greater than both the lightweight or hospital-type wheelchairs. Because ISO testing is based on a 5-year life cycle, it was concluded that lightweight and hospital-type wheelchairs might not last for the typical 3- to 5-year period expected by health insurers. A better investment for an individual patient was shown to be the ultralight, despite its initial high cost. This analysis did not take into account the other advantages of ultralight wheelchairs described in this chapter. Similar testing is being done on power wheelchairs. In general, similar to manual wheelchairs, results indicate that higher quality power wheelchairs will be less costly over the lifespan of the chairs (57).

SECONDARY INJURY/ACCIDENTS

Unfortunately, wheelchair users are at risk for other injuries as a direct result of their wheelchair use (58). In 2003, over 100,000 injuries were treated in the emergency department. Of individuals treated, 68.9% were more than 65 years old and 65.0% were female. The highest percentage of injuries reported were fractures, contusions, and lacerations, and the leading cause of injury was from tips and falls (65% of injuries). Another important area of concern is that most wheelchairs are not crashworthy for use as seats in motor vehicles, as their seat belts and other seating components are not designed to withstand the forces occurring in motor vehicle accidents. Additional seat belts and tie downs must be used to individually hold both the patient and the wheelchair secure (59). Clinicians should be aware of this and discuss transportation directly with the patient. Additional information on the latest development of wheelchair transportation safety is available at www.rercwts.pitt.edu.

Possibly the most important area of concern for manual wheelchair users are repetitive strain injuries of the upper extremities. These injuries are so significant that some researchers have gone so far as to say that damage to the upper limbs may be functionally and economically equivalent to a spinal cord injury of a higher neurologic level (60). The two most common areas of injury are the shoulder, with rotator cuff disease (61) and degenerative arthritis (62), and the wrist, in the form of carpal tunnel syndrome (63,64). These studies have found injury rates as high as 70%. Studies have found a direct link between manual wheelchair propulsion and injury to both the shoulder and wrist (21). Clinicians must be aware that an appropriately prescribed and setup wheelchair, which is propelled in an appropriate manner, can reduce the risk of injury. In addition, for some individuals, it may be appropriate to discuss power-assisted or power wheelchairs as a means for preserving the upper limb for activities such as transfers. As mentioned earlier, the CPG for preservation of upper limb function specifically details the proper manual wheelchair setup and propulsion techniques to prevent repetitive strain injuries (16).

FUNDING/INSURANCE

For many individuals, funding can present a major limitation to the type and quality of the wheelchair they can receive through their health insurance. It is tempting as a practitioner to determine the wheelchair that insurance will cover and then work from this limitation. This path can be poor for the patient and in the long term will not lead to changes in policy. Therefore, it is important that the clinician working with the patient determines the optimal mobility device. After making this determination, the team can then assess what is the best way to convince insurance that the device is medically necessary and should be covered.

Having stated this, it is helpful to know current policies and their impact on patients. The section below is focused on the United States, but policies differ markedly from one country to another. In the United States, wheelchairs and seating systems are covered in whole or part by health insurance plans, including Medicare Part B, state Medicaid programs, commercial insurance, and managed care plans, unless the policy stipulates no durable medical equipment coverage. Other funding sources also exist. If the device is needed for work-related activities, state vocational rehabilitation programs can be utilized. In addition, the Veterans Administration for veterans with both service- and non–service-related disabilities can be a very valuable funding source.

Medicare has made significant changes to coverage policies for mobility assistive equipment (MAE, a new term they use to include canes, crutches, walkers, manual wheelchairs, scooters, and power wheelchairs) over the past several years. It is becoming increasingly difficult to get the appropriate equipment for a patient covered. Medicare funding policy for MAE is important to understand and follow, as equipment can be denied if the practitioner’s documentation does not reflect what is being requested in the policy. Medicare policy is also generally adopted by most other funding sources including State Medicaid programs and private health insurance.

The new Medicare policy approaches the need for MAE from a more functional perspective and uses a new term known as mobility-related activities of daily living (MRADLs, meaning how mobility impacts a person’s ability to participate in ADLs). In order to qualify for any MAE, a person has to have an impairment in “one or more” MRADL in the “home.” The “in the home” language does, however, continue to be a restriction meaning that a manual wheelchair, scooter, or power wheelchair will not be covered if it is “only” to perform MRADLs outside the home. Medicare defines a mobility limitation as one that
• Prevents the patient from accomplishing an MRADL entirely (i.e., independence), or
• The patient is at reasonably determined heightened risk of morbidity or mortality secondary to the attempts to perform an MRADL (i.e., safety); or
• The patient is unable to complete the MRADL within a reasonable time frame (i.e., quality).

Although these coverage policies are based more on functional necessity and then good practice, requirements for documentation in the medical record are very specific and must be furnished to the company that supplies the equipment as they need to keep records on file in the event of an audit. Failure to provide the necessary documentation to the supplier will result in the supplier refusing to dispense the equipment. Physicians are also not permitted to complete forms provided by suppliers or manufacturers as a substitute for what is documented in the medical record. Medicare policy states that the medical record can include (but is not limited to) the items listed in Table 59-2. The policy also recognizes and encourages physicians to refer patients with mobility limitations to an occupational or physical therapist for assessment in identifying the appropriate MAE intervention and to provide necessary documentation.

The qualifying criteria for the various types of wheeled mobility devices are as follows:

The patient has limited ability to ambulate even with the use of a cane, crutch, or walker. For this, it is usually sufficient to explain that an individual can only ambulate short distances or is unsafe ambulating. Placing these limitations in the context of his or her current living situation and usual activities can help persuade payers of the need.

Lightweight and ultralightweight wheelchairs. To qualify for either of these wheelchairs, it is necessary to document that the person is unable to propel a lower-cost alternative standard wheelchair or use a cane or walker in an effective manner. It also helps document the person's lifestyle situation and how the wheelchair will facilitate his or her ability to engage in activities.

Power wheelchairs and power-operated vehicles. Powered mobility devices including power wheelchairs and power-operated vehicles (POVs) or scooters have been recategorized by CMS in 2006.

To qualify for a power wheelchair or POV (or scooter), one must document that the person cannot effectively propel any type of manual wheelchair. For individuals with upper-extremity paralysis, it is obvious that they cannot propel a manual chair. However, many other individuals can require a power wheelchair. People with upper-extremity pain that limits propulsion meet this criterion, because pain and risk of aggravating injury may make them incapable of propelling a manual chair. This is also true for individuals with cardiopulmonary disease or obesity. Both of these conditions make functional manual wheelchair propulsion difficult and, in some cases, impossible. These deficits and risks should be documented and explained in the LNM.

Power wheelchairs are categorized into five groups. Group 1 power wheelchairs can only accommodate people up to a weight of 136 kg (300 lb) and cannot accommodate any seat functions such as tilt-in-space or a seat elevator. They are also not well suited to negotiate on even surfaces or thresholds more than 20 mm in height. Group 2 power wheelchairs can accommodate multiple weight capacities up to 272 kg (600 lb); however, are only designed to accommodate a single power seat function such as a seat elevator or tilt-in-space. Group 3 power wheelchairs can handle multiple weight capacities up to 272 kg (600 lb) and multiple seat functions including a vent tray and alternative controls. Group 3 power wheelchairs are similar to Groups 2 and 3 power wheelchairs; however, have features that are inherent to outdoor mobility such as larger motors and a suspension and therefore are not covered by Medicare and many other payer sources. Group 5 power wheelchairs are classified as pediatric products.

Scooters are only covered when a patient resides in a home where there is sufficient maneuvering space to operate a scooter, the patient can safely transfer in and out of a scooter, a scooter seating system will address the patient's postural needs, and the patient has sufficient upper extremity function to operate the scooter steering mechanism. Scooters are broken down into two groups. Group 1 POVs can handle multiple weight capacities and are generally designed to negotiate most level indoor and outdoor surfaces. Group 2 POVs have designed features such as larger motors and suspensions that are inherent to outdoor mobility only and therefore not covered by Medicare and many other health insurance plans.

Medicare requires that any Group 2 power wheelchair with a single power seat function, any Groups 3 or 4 power wheelchairs, or a PAPAW be provided by a company that employs a RESNA certified ATS. The beneficiary receiving the equipment must also be evaluated by a qualified medical practitioner (such as an occupational or physical therapists) with experience and knowledge in complex wheelchair seating and mobility equipment. As stated early, Medicare also requires that the supplier of wheeled MAE ensures that the patient's home and environment are suitable for the MAE and that there is sufficient maneuvering space to operate the device.

SUMMARY

As can be seen from the text, the prescription of wheelchairs is a complex and time-consuming task. Many parties need to be involved, and technology is constantly changing. In an ideal world, all wheelchairs would be prescribed using all the members of the team described earlier in this chapter, and clinicians would be reimbursed at a level that allowed for appropriate evaluation and training. Unfortunately, wheelchair clinics are costly to run; therefore, the team approach described may not be possible in many settings. It is essential that the physician signing the prescription understands the equipment well enough to explain the choices made as well as the trade-offs involved, limitations, and safety issues. Most importantly, the consumer should be able to make informed choices.
A clinician practicing without a clinic can still provide good care. The best approach is to find a therapist in the rehabilitation team who has an interest in wheelchairs. If possible, the therapist should attend meetings such as RESNA (www.resna.org), the International Seating Symposium (www.iss.pitt.edu), or Medtrade (www.medtrade.com) to learn about new technology. In addition, the therapist should consider taking the ATP examination for certification. This can be sold to the hospital as a value-added service that their institution has and others do not. The other key team member then becomes the dealer. The doctor and therapist team should request (or require) that the dealer become a CRTS. You can also ask the dealer to support the team by having equipment available for trials and by visiting the patient’s house. This team can be very effective at wheelchair delivery and can improve the function and quality of life of their patients.

FUTURE

New and exciting products are placed on the market each day. As mobility products improve, the line between needing the device because of a disability and wanting the device because it enhances mobility can become blurred. Mainstream use of assistive devices is a great thing for wheelchair users as it expands the market and lowers costs. In addition, it blurs the line between disability and normal function. In the future, it would not be uncommon for everyone to have a personal mobility device, such as Segway (65), and the only difference between individuals with disabilities and individuals without impairments is that individuals with disabilities always use their vehicle. With these advances it will be important for health care professionals to continue to lobby on behalf of their patients to increase funding for wheelchairs so that function dictates the prescription.

Note: Michael Boninger and Rory Cooper are coinventors on the Natural-Fit hand rim. As such they receive a royalty from the University of Pittsburgh related to its sales. In addition, Three Rivers Holdings is owned and operated by Dr. Boninger’s brothers. However, neither Dr. Boninger nor Dr. Cooper has ownership in the company and neither has been or are currently paid consultants.

REFERENCES


Complementary and Alternative Medicine

The practice of physical medicine and rehabilitation (PM&R) is inherently a holistic discipline which focuses on the patient as a whole human being, functioning on physical, emotional, mental, social, and spiritual levels simultaneously. PM&R is a leader in integrating the concepts of holism and a diverse spectrum of therapeutic modalities into the conventional medical paradigm. PM&R and complementary and alternative medicine (CAM) share several core principles, including patient-centered care, the concept that the wisdom of the team is greater than the sum of its parts, and clinician-patient partnership. These ideas are now more commonly being adopted in other areas of medicine. Because of the similarities between the theoretical underpinnings of PM&R and of the philosophies associated with CAM modalities, CAM fits easily into the rehabilitation model. According to the National Health Information Survey (NHIS) conducted in 2007, 83 million adults spent $33.9 billion out of pocket on CAM, and CAM costs are 11.2% of total out-of-pocket expenditures on health care. American adults are most likely to use CAM for musculoskeletal problems such as back, neck, or joint pain (1).

After World War II, the distinctive field of PM&R evolved in response to the fact that mainstream medicine was not providing adequate care and treatment for patients who had sustained injuries to the musculoskeletal and central nervous systems. In his pioneering book, *The Knife Is Not Enough*, Dr. Henry H. Kessler clearly espoused the need for what are today being called "complementary and alternative" therapies (2). He included not only physical and dietary interventions in his multidisciplinary approach, but he also recognized the important role that spirituality played in healing. This approach has marked PM&R with more openness to CAM than other specialties.

The NIH National Center for Complementary and Alternative Medicine (NCCAM) describes whole medical systems as approaches to care that are built upon complete systems of theory and practice that have evolved independently of the conventional medical approach used in North America and Europe. Examples of whole medical systems that have developed in Western cultures include homeopathic and naturopathic medicine. Examples of systems that have developed in non-Western cultures include Ayurveda and traditional oriental medicine. Each of these approaches has relevance to PM&R, but the area within the whole medical systems category that currently plays the biggest role in the field of PM&R is Oriental Medicine.

Oriental Medicine

Oriental medicine is over 3,000 years old and is based on the theory that there is a vital life force (qi) that supports physiologic functioning. Qi is concentrated in multiple pathways or channels called meridians that run longitudinally throughout the body. Health is the state in which qi is flowing abundantly and harmoniously, and illness is believed to result from an improper amount distribution or imbalance of this energy, which then precipitates physiologic change (3). Acupuncture is the insertion of thin, noncutting needles into specific acupuncture points along the meridians in order to correct aberrant energy flow. Of all modalities under the heading of Whole Medical systems, acupuncture is the most widely used in PM&R and is discussed in detail. T’ai Chi, a movement...
PART VI | MANAGEMENT METHODS

Peripheral nerves, and is not observed with nonacupuncture points (7,8). More recently, similar results using gadolinium injections and MRI have been reported in humans (9). There is substantial overlap in the locations and properties of myofascial trigger points, which react to needling with a local twitch response and predictable pain radiation patterns, and acupuncture points, which react to needling with a local ache (called "de qi") and radiation of this sensation along the associated acupuncture channels (10). However, as acupoints are not always tender to palpation and are often used for other indications aside from local pain, debate exists about whether trigger points may merely represent a smaller subset of acupuncture points called "a shi" points (11–13).

Anatomic studies in vivo and postmortem section in the upper arm have revealed that more than 80% of acupuncture points consist of a characteristic subcutaneous column of loose connective tissue, with a looped core bundle of neurovascular and lymphatic vessels that may help propagate the influence of superficial needle inputs to deeper structures (5). Decreased electrical resistance and increased capacitance have been reported between acupuncture points sharing energy channels (6). Additionally, nuclear medicine studies have shown that subcutaneous injection of technetium-99m into the low-resistance acupoints of humans and dogs results in a rapid migration of the radioactive tracer along the associated meridians; this movement is distinct from the trajectory of blood vessels, lymphatics, and peripheral nerves, and is not observed with nonacupuncture points (7,8). More recently, similar results using gadolinium injections and MRI have been reported in humans (9). There is substantial overlap in the locations and properties of myofascial trigger points, which react to needling with a local twitch response and predictable pain radiation patterns, and acupuncture points, which react to needling with a local ache (called "de qi") and radiation of this sensation along the associated acupuncture channels (10). However, as acupoints are not always tender to palpation and are often used for other indications aside from local pain, debate exists about whether trigger points may merely represent a smaller subset of acupuncture points called "a shi" points (11–13).

Acupuncture

Acupuncture has numerous applications in PM&R. It is utilized for both acute and chronic musculoskeletal and neurologic problems. Needles are placed both locally, near the anatomic location of the problem, as well as in other areas of the body. A typical acupuncture session lasts 30 to 60 minutes including examination and needle insertion. The number of sessions required depends on the intensity and chronicity of a condition as well as the age and general health of the patient. For an acute back spasm in a young and otherwise healthy individual, one to two sessions may be sufficient. For long-standing pain in lumbar spinal stenosis or motor recovery in hemiplegia secondary to cerebrovascular accident (CVA) in an older individual, up to 20 or more closely spaced sessions followed by less frequent maintenance may be required to achieve optimum results.

Recent research has begun to highlight promising avenues of exploration regarding the physiologic nature of the acupuncture channels and the points themselves (4). Histologic studies have suggested that some 80% of acupuncture points consist of a characteristic subcutaneous column of loose connective tissue, with a looped core bundle of neurovascular and lymphatic vessels that may help propagate the influence of superficial needle inputs to deeper structures (5). Decreased electrical resistance and increased capacitance have been reported between acupuncture points sharing energy channels (6). Additionally, nuclear medicine studies have shown that subcutaneous injection of technetium-99m into the low-resistance acupoints of humans and dogs results in a rapid migration of the radioactive tracer along the associated meridians; this movement is distinct from the trajectory of blood vessels, lymphatics, and peripheral nerves, and is not observed with nonacupuncture points (7,8). More recently, similar results using gadolinium injections and MRI have been reported in humans (9). There is substantial overlap in the locations and properties of myofascial trigger points, which react to needling with a local twitch response and predictable pain radiation patterns, and acupuncture points, which react to needling with a local ache (called "de qi") and radiation of this sensation along the associated acupuncture channels (10). However, as acupoints are not always tender to palpation and are often used for other indications aside from local pain, debate exists about whether trigger points may merely represent a smaller subset of acupuncture points called "a shi" points (11–13).

Anatomic studies in vivo and postmortem section in the upper arm have revealed that more than 80% of acupuncture points and 50% of meridians correspond to intermuscular and intramuscular connective tissue cleavage planes (14). Microscopic images show that the tissue resistance “needle grasp” effect (a sensation of resistance felt by the practitioner when the acupuncture needle has been advanced to its desired depth in some acupoints) that is clinically recognized upon twisting an inserted needle actually corresponds to the winding of collagen fibers and fibroblasts in the connective tissue around the surface of that rotated needle (14–17). Needle grasp of acupoints may thus propagate a mechanical stretch to the deeper fascial matrix, which ultimately invests structures such as the peritoneum of the viscer, the pleura, the perineum, and the meninges (14). Mechanotransduction of this stretch converts into a potentially gene transcription–altering intracellular signal by deforming the cytoskeletons of fibroblasts and directly influencing actin and microtubule remodeling, as shown in a murine model (18,19). Related studies have demonstrated that cyclical stretching of fibroblasts in this fashion decreases TGF-β1 production and type 1 procollagen...
Homeopathy
Homeopathy is the system of diagnosis and treatment developed by Samuel Hahnemann, MD, a German physician who viewed the symptoms of an illness as a reflection of the body’s attempts to regain homeostasis. He postulated that if we could properly interpret these symptoms, we could then direct the body’s own homeostatic mechanisms toward healing illness. Hahnemann reported that when a small test dose of a biological substance is given to a healthy person, it elicits a predictable array of signs and symptoms. Hundreds of substances have been tested in this manner (this is referred to as a “proving”) and logged in the Homeopathic Materia Medica. Homeopathic theory states that a dilute amount of this substance will help to direct the body’s innate healing mechanisms toward balance. When a patient becomes ill, symptoms are recorded and closely matched to the material that causes similar symptoms in a healthy person. The homeopath combines this information with past medical and social history, and examination in order to prescribe a homeopathic remedy. Prescription of remedies in homeopathy is highly individualized, since the mental-emotional pattern, general level of health, and particular physical attributes of the patient are all taken into consideration. The underlying theory is that the remedy will help to steer the body back to homeostasis.

Many medical doctors in the United States used homeopathy in the early 1900s, and there were 16 homeopathic medical schools in the United States at the beginning of the 20th century. Since then its popularity declined until recently and in 1997, 3.4 million visits to homeopathic providers were reported in the United States (49,50). Research in homeopathy using current double-blind placebo-controlled methods presents considerable difficulty. Two patients with the same allopathic diagnosis may have two completely different symptom complexes when the patient is examined from a homeopathic point of view. The result is that two different homeopathic remedies may be prescribed for two patients with the same allopathic diagnosis. Nonetheless, it has shown to be useful in the treatment of mild TBI (51) and some musculoskeletal problems (52–54). A meta-analysis of four trials on osteoarthritis concluded that although the clinical evidence appears promising, the small number of randomized clinical trials conducted to date, although favoring homeopathic treatment, did not allow a firm conclusion as to the effectiveness of homeopathic remedies in the treatment of patients with osteoarthritis (55).

Ayurveda
The indigenous healing system native to India, Ayurveda translates as the “science of life.” The basic principles of Ayurveda include the interconvertibility of energy and matter and the interconnectedness of all life (56). Like the field of PM&R, Ayurveda teaches that disease and healing occur on all levels: physical, emotional, mental, and spiritual. Everything in the cosmos is believed to be made up of varying proportions of the five elements—earth, air, water, fire, and space. According to Ayurveda, these elements also occur in people in varying
proportions that are grouped into types functionally classified as doshas. These proportions make up the individual’s prakriti, or individual constitution, and determine the mental and physical makeup of each individual. Illness is thought to occur when there is an imbalance in the doshas. Each dosha has corresponding symptoms and illnesses that occur during imbalance. Treatments are individualized by constitutional type and include herbs, yoga postures, diet, pranayama or breathing techniques, purification techniques, meditation, and mantras (25).

Ayurveda is not well integrated into the PM&R setting at this time. However, one tool used by Ayurvedic practitioners, yoga, is discussed in separate sections below.

Naturopathy
Naturopathy is a system that stresses health maintenance and disease prevention through patient education and acceptance of responsibility for one’s own health. Underlying the various treatments in naturopathic medicine is the belief in the healing power of nature, and the innate intelligence of the body. Naturopathic physicians are trained in the medical sciences as well as combinations of disciplines such as herbology, nutrition, homeopathy, or acupuncture. A naturopathic doctor holds the degree ND and has 4 years of training followed by an internship. Naturopathic physicians are licensed in some states and Canadian provinces, and scope of practice for some includes limited drug prescription.

MIND-BODY INTERVENTIONS
The basic premise of mind-body interventions is that the mind and body are not dually distinct entities but exist on a continuum. Thus, thoughts and emotions influence the body, and physical processes have an impact on the mental, emotional, and spiritual state. Mind-body interventions may be separated into either a mental or physical process that influences this mind-body continuum. “Mind”-based techniques including visual imagery, art and music therapy, biofeedback, hypnosis, meditation, and prayer are designed to act through the mind to alter its state and thus physical conditions including muscle tension, endorphin levels, and pain. Body-based techniques use the body as a vehicle and act to increase strength, flexibility, and body awareness and in so doing, impact upon mental/psychological/spiritual states such as level of mental focus, anxiety, or compassion for self and others. Modern day usage and supporting evidence are listed for mind-body techniques separately below.

Hypnotherapy
Hypnotherapy works at the level of the subconscious mind to assist patients in making positive behavioral changes and to decrease symptoms. A clinical hypnotherapist makes use of therapeutic suggestion while an individual is in a highly relaxed and altered state of awareness. Hypnotherapy is used to improve self-esteem, improve mood, increase feelings of self-efficacy, and control and decrease negative behaviors and fears that interfere with recovery. In case reports of patients treated with hypnosis following a stroke, observations were made of increased movement in hemiplegic limbs and improved ambulation (57–59) as well as the return of normal speech (57,59). Hypnosis was used in the treatment of headache or vertigo following brain injury in a series of 155 patients, with half reporting resolution of symptoms and another 20% experienced significant symptom reduction (60). In the area of musculoskeletal rehabilitation, hypnosis, when used, is largely for pain control. One NIH consensus review panel and one meta-analysis suggest that hypnotherapy may be a useful adjunct to the treatment of chronic low back pain and other neuromusculoskeletal conditions (61,62). Positron emission technology (PET) scanning reveals that the hypnotic state is related to the activation of a widespread set of cortical areas involving occipital, parietal, precentral, premotor, and ventrolateral prefrontal and anterior cingulate cortices (63).

Meditation
In some form, meditation has long been used in many cultures, both Eastern and Western. Originally, it was tied to religious practices and was performed with spiritual goals of enlightenment. Modern forms of meditation have separated the techniques from their religious contexts, and the goals for which meditation is used have become more concrete, for example, decreased anxiety and improved health. While there are many forms of meditation, most are intended to create a state of physical and mental relaxation. Most forms of meditation involve sitting or lying quietly while focusing on one or two stimuli, such as breathing patterns, a word, or an image.

The most robust finding on the impact of meditation is the success of meditation for reducing hypertension (64–67). Research evidence for the effectiveness of meditation in the rehabilitation process is limited to anecdotal reports (68).

Meditation has been shown to improve cognitive functioning in college students and elderly adults who have no history of head injury (69,70). For pain and musculoskeletal conditions, studies indicate that the practice of meditation as part of a comprehensive treatment program can improve self-management of chronic pain (71), fibromyalgia symptoms (72) and perception of pain in arthritis (73). Although multiple studies have been performed on the effect of meditation on various conditions, for the most part, methodological quality has not been sufficient to allow conclusive statements on effectiveness (74).

Mind-Body Movement Therapies
Described here are six of the more popular therapeutic movement modalities in use today: yoga, T’ai Chi, the Alexander technique, the Feldenkrais method, Pilates, and body-mind centering (BMC). These techniques are similar in that they aim to improve a patient’s kinesthetic ability, coordination of motion and breathing, and ease, control, and joy in everyday movement. All may be useful for a wide spectrum of patients from severely physically challenged to high-performance athletes.

Yoga and T’ai Chi, are frequently referred to as “mind-body” techniques, because they actively seek to balance both
the physical and nonphysical aspects of the person. Both are meditative in nature and are part of a larger philosophy or way of life. Each is derived from a broader system of health with an underlying model differing from those of the West. The other five techniques described were developed in the 20th century. All of the techniques take a less focal and more global approach to rehabilitation of specific conditions based on the rationale that the whole body is involved in all movement. A summary of CAM movement therapies and indications is listed in Table 79-2.

Yoga

Yoga is an ancient Indian art first brought to this country in the mid-1800s. In Sanskrit, the word “yoga” means “union” (with the Divine), and is a way of life involving a number of different spiritual practices and encompasses ethical conduct, social responsibility, nutrition, and physical health practices. The branch of yoga that is best known in the West is hatha yoga, and is often simply referred to as “yoga.” Originally intended to prepare the body for Divine experience, it is practiced in this country for the achievement of physical strength, flexibility, and relaxation through postures known as asanas (a-sa-nas). Pranayama (pra-na-ya-ma), or breathing techniques, and meditation are also often practiced along with hatha yoga. The many forms of yoga practiced in the United States are diverse in terms of focus, strenuousness, and applications. A few of the forms most relevant to PM&R are described in Table 79-2.

Yoga has been extensively studied both in India and the West, with thousands of studies reporting positive health effects, including lowering of blood pressure (75), and decreased cholesterol levels (76). Yoga has been applied to programs for rheumatoid arthritis (77), osteoarthritis (78), chronic back pain (79), cardiac rehabilitation (80), carpal tunnel syndrome (78), and improved athletic performance (81). Yoga is increasingly being integrated into the office and inpatient settings by instructors who are experienced in teaching rehabilitation patients (82). Individual sessions and classes are a convenient way to transition from physical therapy supervision to a home-based program (83).

T’ai Chi

T’ai Chi is a form of postures and movements that dates back to 17th century China. T’ai Chi is made up of a series of flowing contrasting movements, with constant weight shifting from one leg to another, changing of direction, and moving the limbs in space. Specific postures are linked together with the focus on achieving balanced and graceful movement. Since it is a slow, rhythmic and weight-bearing exercise, participants

<table>
<thead>
<tr>
<th>Technique</th>
<th>Focus</th>
<th>Indications, Precautions, Considerations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anusara yoga (Sanskrit for “flowing with grace”)</td>
<td>Alignment, smooth flow of poses, use of yoga to maintain health</td>
<td>General health, balanced combination of alignment, posture flow and heart-centered approach; rigorous certification process</td>
</tr>
<tr>
<td>Kripalu yoga (Sanskrit for “compassion”)</td>
<td>Gentle movement, strength and flexibility, awareness and care of the body, release of emotions</td>
<td>Deconditioning, neck and back pain, ROM deficits, fibromyalgia, initial rehabilitation phase of recovery from illness esp. post-cardiac, postsurgical, increasing general strength and flexibility; broad-based certification process</td>
</tr>
<tr>
<td>Iyengar yoga</td>
<td>Alignment, correct execution of poses, use of yoga to correct physical health problems</td>
<td>Back and neck pain, scoliosis, sports rehabilitation; significant use of props and modifications for physical limitations; rigorous certification process</td>
</tr>
<tr>
<td>Viniyoga</td>
<td>Tailoring practice to the individual, therapeutically linking of breath with movement</td>
<td>Severe deconditioning, general conditioning, neck and back pain</td>
</tr>
<tr>
<td>Ashtanga yoga, Bikram yoga T’ai Chi</td>
<td>Strength building</td>
<td>Athletic and mental conditioning; not to be used in serious injury, deconditioning, or rehabilitation Fall-prone patients, peripheral neuropathy neck and back pain, general deconditioning. Requires ability to stand; may be modified to patient needs</td>
</tr>
<tr>
<td>Feldenkrais</td>
<td>Exploration and choice of useful movement; application to functional tasks</td>
<td>Neurological disorders, orthopedic rehab, neck and back pain; rehabilitation of dance and musician injuries</td>
</tr>
<tr>
<td>Alexander technique</td>
<td>Posture and alignment, release of mal-adaptive movement habits, low back and neck pain</td>
<td>Neck injuries, scoliosis, low back pain, neck pain, rehabilitation of dancer and musician injuries and any activity in which proper posture is important</td>
</tr>
<tr>
<td>Pilates</td>
<td>Core strengthening, balanced strengthening, ease of movement</td>
<td>Athletic rehabilitation, general strengthening, neck and back pain, postural awareness; avoid with moderate to severe injury unless practitioner has rehab experience</td>
</tr>
<tr>
<td>Body-mind centering</td>
<td>Awareness through touch, movement, and intention</td>
<td>Developmental delay, CP, autism, stroke, TBI, SCI, neck and back pain</td>
</tr>
</tbody>
</table>
are able to improve balance, coordination, concentration, and relieve stress in a safe manner. Randomized trials report that benefits of T’ai Chi include improved cardiorespiratory function, improved strength and balance, decreased falls in the elderly (84–87), and improved psychological parameters (88). T’ai Chi has been modified for the elderly (T’ai Chi Chih), and the arthritis population and instructors are now certified by the Arthritis Foundation.

Movement Awareness Techniques
Movement awareness techniques focus less on strengthening isolated muscles and more on ease and comfort in functional movement. Each one promotes body awareness and balance and may be taught in one-on-one or in group format. The Alexander technique was developed by FM Alexander, a Shakespearian actor, in his attempts to heal his own recurrent loss of voice. Alexander found that his voice improved when he decreased his cervical tension. His technique attempts to improve posture and movement by focusing on developing a balance between the head and neck in static and dynamic situations, as well as proper breathing (89). Most sessions are one-on-one, with the practitioner directing the student verbally and with light touch into postures that will help him or her to experience the state of being in proper alignment. Alexander practitioners are consulted for any patient with postural difficulties or chronic and occupational back and neck pain. Clinical reports include effectiveness in increasing respiratory capacity (90) in Parkinson’s (91) or for improved performance in athletics and the performing arts.

The Feldenkrais method is a movement therapy that capitalizes on the ability of the human nervous system to develop new engrams or movement patterns. It was developed by Moshe Feldenkrais, PhD, a physicist and judo black belt who studied Alexander technique, yoga, psychology, and other disciplines and applied them to injury rehabilitation. This method teaches the patient to break down and examine functional movement habits and to choose new patterns based on their usefulness and efficiency. The technique helps patients to learn to use efficient and pain-free movement. Feldenkrais sessions are conducted in two formats. Awareness through movement is a verbally directed group lesson in which movements are gentle and slow, within comfort range. Lessons are based on a function (getting up from a chair, rolling from supine to prone). The second variation, Functional Integration, is a one-on-one, hands-on technique in which the practitioner gently guides the student through various movement sequences. Feldenkrais has been used with various conditions, including cerebral palsy, hemiplegia, and multiple sclerosis. There are reports of the successful use of the Feldenkrais method in multiple sclerosis (92), orthopedic injuries (93), Parkinson’s disease (94), and neck and low back pain (95,96) and evidence that it is cost effective in the management of chronic pain (97). A review of the numerous studies on Feldenkrais in rehabilitation concluded that with a great number of positive case reports, further research with rigorous methodology is warranted (98).

The Pilates method was developed by Joseph Pilates during and after World War I and combines the experiences he gained as a result of his contact with dancers, athletes, and disabled soldiers. The technique is now used by many high-level athletes and dancers as well as those undergoing rehabilitation, especially for musculoskeletal conditions. This popular technique is characterized by use of proper body mechanics, stabilization of the shoulder and pelvic girdles, and coordinated breathing to promote strengthening with minimal increase in bulk. A typical rehabilitation program using a Pilates-based approach may begin not by focusing on the area that is injured but by teaching exercises to enhance awareness that the body works as a whole and that the injured part is weak link in the entire kinetic chain (99). Clinical reports of the use of Pilates-based protocols are in the area of sports training and rehabilitation (100,101).

Bonnie Bainbridge-Cohen developed BMC, a unique approach to movement analysis and reeducation (102). Cohen drew on her training in occupational therapy, and Bobath’s neurodevelopmental therapy, dance therapy, bodywork, martial arts, voice, and yoga.

The primary goal of BMC is the integration of body and mind, and the development of an ability to sense this connectedness. BMC teaches students to tune into the body and feel at the tissue and cellular levels. BMC holds that the body’s tissues have the ability to respond to messages given to them (102). Thus, body tissues can be affected by the practitioner’s hands as well as the patient’s actions, thoughts, and intentions. The result is the modulation of pain and restrictions preventing optimal movement. BMC is used for patients affected by any neurological or musculoskeletal condition including developmental issues, stroke, SCI, TBI, and orthopedic conditions.

BIOLoGICAL THERAPIES
Biological therapies refer to modalities which act by biochemical means to elicit a desired outcome. The category includes diet, herbal, and vitamin supplements. Biological therapies are appealing to the public because of their accessibility and perceived efficacy and safety. According to the most recent NHIS report in 2007 (103), the most common supplement used by Americans is fish oil, followed by glucosamine. Because of patient request for natural therapies, wide usage, increasing literature support, and their unique side effect profiles and interactions, physicians should be familiar with the range of biological treatments frequently used for musculoskeletal disorders.

Unlike conventional pharmacologic treatments which undergo rigorous testing for safety and efficacy and must be approved by the FDA before they are marketed, herbs and nutritional supplements are not held to the same quality standards. In 1994, Congress enacted the Dietary Supplement Health and Education Act (104) which defines “dietary supplement” as products taken by mouth which contains one or a combination of vitamins, minerals, herbs or other botanicals, amino acids,
concentrates, metabolites, constituents, or extracts. The Act makes manufacturers responsible for ensuring the safety and efficacy of their products by listing all ingredients and including a disclaimer stating that the product has not been evaluated by the FDA and "is not intended to diagnose, treat, cure, or prevent any disease" (104). Good Manufacturing Process or GMP is a set of standards voluntarily adopted by manufacturers, and the acronym GRAS, or Generally Regarded As Safe, reflects a level of safety that is agreed upon by experts. The FDA prohibits use of any food additive or supplement that is not GRAS (105).

The agents discussed here were chosen from a long list of supplements as those that have the most literature support and which are relevant to PMR conditions. The majority of herbs and supplements used in common disorders seen by PM&R professionals are used for pain management of musculoskeletal conditions and have some anti-inflammatory mechanism (see below); others have structural modification properties, still others have diverse mechanisms.

Before discussing supplements with patients, it is wise to discuss the effect of diet in a healing neuromusculoskeletal system. The importance of nutrition cannot be overemphasized. Because the Standard American Diet (SAD) contains significant amounts of precursors to proinflammatory fatty acids (saturated fat, vegetable oils, most nuts), with lesser proportions of foods that give rise to the manufacture of anti-inflammatory compounds (fruits and vegetables, omega-3–containing nuts, cold water fish), many patients ingest a diet that contributes to a high inflammatory load (106). Since most pain syndromes are associated with inflammation (107), an anti-inflammatory diet is a prudent choice for patients in pain or with healing neuromusculoskeletal systems. The basic principles of an anti-inflammatory diet include reduction of sources of arachidonic acid (mainly saturated fat), elimination of trans fats, which interfere with essential fatty acids (EFA) metabolism, optimization of omega 6 to omega 3 ratios, consumption of abundant and variety of phytochemicals from fruits and vegetables, and use of anti-inflammatory herbs and spices. Several popular guides to an anti-inflammatory diet exist (108,109). There is evidence that an anti-inflammatory diet supports the actions of anti-inflammatory supplements and decreases symptomatology in patients with systemic inflammatory conditions (110). There is an increasing number of reports of correlation between dietary imbalances and autoimmune musculoskeletal diseases (111,112), osteoporosis (113–115), and inflammatory conditions (116), and therefore opportunities for intervention. The physiatrist who makes herself/himself aware of these relationships and seeks appropriate help may help save the patient years of suffering.

One of the most discussed supplements in medicine today is the use of fish oil. Fish oil contains eicosapentaenoic acid (EPA) and docosapentaenoic (DHA), both precursors in the manufacture of anti-inflammatory compounds. EPA and DHA are necessary to the development and maintenance of healthy membranes, and deficits associated with aberrant receptor binding capacity, cell signaling, and neural development. The use of fish oil has been well studied in the rheumatoid arthritis population, and patients tolerate high doses (4.5 g/d DHA/EPA) with no negative incidents in 3 years of follow-up (117). Approximately 2 to 3 g/d of combined DHA/EPA is considered anti-inflammatory, and patients should be started at approximately 2 g/d and titrated upward as tolerated (117). Omega 6: Omega 3 ratios can be periodically monitored.

Glucosamine, widely used for joint pain, is the supplement that has the second highest frequency of use in the United States and is in the category of structure-modifying agents. Glucosamine is made from glucose and the amino acid glutamine and is a fundamental building block for glycolipids, glycoproteins, glycosaminoglycans, and hyalurionate. It is required for production of glycosaminoglycans, a molecule used in the formation and repair of cartilage and other body tissues. The production of glucosamine slows with age.

Over 20 trials have examined the effectiveness of glucosamine with variable results (118). Some of the important variables in such trials include glucosamine preparation (glucosamine sulfate vs. glucosamine hydrochloride), dosing schedule, and outcome measures (pain, function, radiographic change). The GAIT and GUIDE trials are the two most definitive trials, and illustrate this point well. The GAIT trial used five conditions: glucosamine HCl 500 tid versus chondroitin sulfate 400 tid versus combination versus celecoxib 200 qd versus placebo × 24 weeks in a total of 1,583 patients that were stratified into mild, moderate and severe. Despite an unusually high placebo response (60%), there was a mild response in the glucosamine sulfate combination in the moderate-severe group (119). In the GUIDE trial, glucosamine sulfate 1,500 mg was superior to acetaminophen 3,000 mg in a total of 318 patients × 24 weeks (120).

In studies evaluating the pharmacokinetics of glucosamine, the serum level peaks at 10 μm after oral ingestion, which is thought to be insufficient for clinical effect. Glucosamine may act as a carrier for sulfate. This hypothesis predicts better performance of sulfate versus nonsulfate glucosamine salts (121–123). Glucosamine may also have an anti-inflammatory effect via inhibition of the IL-1 signaling cascade and down-regulation of the catabolic effects of proinflammatory molecules present in osteoarthritic cartilage. It also inhibits production of matrix metalloproteinase-3 (MMP-3) which degrades the collagen matrix.

Two studies evaluated the effect of glucosamine sulfate (1,500 mg qd) on joint space narrowing after 3 years, and both showed an increase in joint space as compared to placebo (124,125). Furthermore, 5-year follow-up showed a 0.52 relative risk of lower limb surgery (124) and 75% reduction in knee replacement surgery (125) compared to placebo.

There has been some suggestion that glucosamine was associated with increased blood glucose levels but that has been refuted (121). No other significant side effects have been reported.
**Avocado/Soybean Unsaponifiables**

A preparation of the unsaponifiable fractions of avocado and soybean is reported to repress the catabolic activities of chondrocytes and to increase proteoglycan accumulation by OA chondrocytes. Avocado/soybean unsaponifiables (ASU) have been shown to inhibit basal MMP-3 production and inflammatory cytokine production and to have a chondroprotective effect via IL-1-repression (126). Two randomized controlled trials in patients with hip and knee osteoarthritis taking ASU found a reduction in nonsteroidal anti-inflammatory drugs (NSAID) usage. Disease modifying effects following the administration of ASU were reported in human OA of the hip (127). Dosage is 300 mg/d, standardized to 30% phytosterols; precautions include cross-reactivity in patients with latex allergies.

**SAMe**

S-adenosylmethionine (SAMe) is an amino acid precursor which is found naturally in human cells and has been used in Europe for over 25 years for the treatment of a broad range of conditions including osteoarthritis, depression, and fibromyalgia. The mechanism of action appears to be its involvement in several metabolic processes including those which synthesize proteins and detoxify substances in the liver (128). With respect to osteoarthritis, SAMe increases chondrocyte viability and function (129), likely by acting as a restorative after damage induced by TNF-α (129). It has been found in clinical trials to be as effective as naproxen 750 mg/d, indomethacin 150 mg/d, piroxicam 20 mg/d, and ibuprofen 1,200 mg/d (130). Dosage in such trials ranged from 400 to 1,600 mg.

**Antioxidants**

While a full discussion of all antioxidants that could be of benefit during rehabilitation is not possible here, a few noteworthy associations stand out: Coenzyme Q10 (CoQ10), or ubiquinone, is manufactured in the body and is an essential contributor to mitochondrial function and is depleted by statins. CoQ10 has been shown to increase the time between the diagnosis of Parkinson's disease and the need for medication in a dose-dependent manner. Optimal dosing is 1,200 mg/d, but may be effective in lower doses if given in conjunction with other antioxidants. A 50-day course of vitamin C, when given at the time of diagnosis of wrist fractures, prevented the development of complex regional pain syndrome (CRPS) in a dose-dependent manner (131). Alpha lipoic acid is a potent antioxidant that has been used to treat painful peripheral neuropathy. Earlier trials showed mixed results, but a more recent multicenter trial using IV alpha lipoic acid trial at three doses (1,200, 600, or 100 mg ALA), 82% of the group receiving 600 mg had significant relief of symptoms at 19 days. It is not known whether oral dosing will give the same effect (132).

**Anti-Inflammatory Botanicals**

Patients at times seek an opportunity to use natural anti-inflammatory and circulation enhancing supplements. Most botanical anti-inflammatory agents act in a manner similar to NSAIDs without the same side effects. Boswellia serrata (frankincense) has been found to be helpful in the treatment of OA of the knee (133) and other chronic inflammatory diseases (134) and to potentiate glucosamine action (135), probably by blocking arachidonic acid to leukotriene conversion (136). Bromelain is a sulfur-containing enzyme from the pineapple plant which has been found to inhibit PGE2 and enhance PGE1 synthesis and to have separate analgesic properties. Effective dosage in studies ranges between 100 and 900 mg/d, with no significant adverse effects (137). Zingiber (ginger) inhibits COX and LOX pathways (138) and is commonly used to enhance digestion of herbal substances in Chinese Medicine. Curcuma longa (curcumin) is used in many pain formulae and has been investigated for effects on autoimmune disease (139); Harpagophyllum procumbens (Devil’s claw) has undergone multiple placebo-controlled trials demonstrating success with and is prescribed widely in Europe for low back pain (140); Salix alba (white willow bark) has been trialed successfully and is used for migraine prophylaxis and low back pain (141); All of these compounds inhibit conversion of arachidonic acid to proinflammatory compounds. These compounds are often found in combination, with less than the maximal amount of each substance, with the rationale that the sum of multiple agents will be mutually potentiating without the need for high doses of each agent.

Vitamin D is more aptly termed a hormone and has been found to be low in patients with osteoarthritis, cancer, multiple sclerosis, and associated with muscle soreness in older women (142). Recommendations are for levels to be maintained between 50 and 100 ng/mL. Dosage to achieve this is 1,000 U/d or for monitored patients, 50,000 U weekly for 8 weeks (prescription only) (143).

The leaf of the ginkgo tree has been part of the traditional Chinese pharmacopoeia for 500 years. A standardized extract of the ginkgo leaf is widely used in Asia, Germany, and France, and is increasingly being used in the United States. There are multiple active ingredients in the extract, including amino acids, flavonoids, flavonols, and steroids. Ginkgo biloba extract has been used in the treatment of symptoms associated with cardiovascular injury, traumatic brain injury, dementia, and memory impairment (144). One placebo-controlled trial found no effect of ginkgo on memory enhancement in healthy patients over 60 (145). There have been multiple trials examining the effect of ginkgo on cognition in healthy and impaired populations (146), but at this time there is no convincing evidence that it should be recommended for routine care in the enhancement of memory in healthy patients, in the prevention of stroke, or in the reduction of cognitive decline. The usual pharmacologic regimen for the treatment of cerebrovascular disease is 120 to 240 mg in divided doses (147).

The physiatrist should advise patients that “natural” does not mean “safe” and that research on supplements is in its infancy. Just as with some pharmaceutical treatments, “nutraceutical” agents designed for therapeutic benefit may
MANIPULATIVE AND BODY-BASED THERAPIES

Manipulative and body-based therapies are employed to improve the structure and functioning of the body by manipulation of the musculoskeletal structure. Although often considered to be “unconventional” in mainstream medicine, many forms of physical manipulation and massage have been utilized for years in PM&R and osteopathic manual medicine. Typically, massage done in the context of PM&R is based on Swedish massage. A detailed discussion of the theoretical basis and nomenclature for massage is presented in Chapter 64. Less conventional bodywork concepts and techniques are briefly reviewed below. An excellent introduction to a number of bodywork, mind-body, and energy techniques applied to the rehabilitation process can be found elsewhere (148–150).

Eastern massage styles, such as acupressure, shiatsu, and tuina, are derived from the original Chinese massage/manipulation technique known as anmo (“press and stroke”). These techniques involve hand movements familiar to Western practitioners such as pressing, stroking, kneading, and grasping, as well as others such as joint manipulation, knuckle rolling, and use of the feet and knees to tread on the client that are outside the scope of Western massage therapy. As with acupuncture, smooth and adequate flow of energy (chi) is considered the basis for good health. The most widespread use of Eastern massage techniques in the United States is by massage therapists trained in Western massage technique, integrating these additional skills into their practice. Acupressure usually takes place on a treatment table, and lubricants are not ordinarily used. Case studies report it to be effective for the treatment of headache and temporomandibular joint pain, postoperative pain, and as an adjunct to pulmonary rehabilitation (151). Shisatsu means “finger pressure” in Japanese. Although the term is used synonymously with acupressure, the two are distinctly different therapies. In shiatsu massage, the session is performed on the floor on a blanket or futon rather than on a table. This allows the practitioner to use various parts of his or her own body (feet, legs, or full body weight) for efficient application of pressure over points and along meridians. Passive stretching and range of motion are also a part of the session. Tuina (“push and pull” or “push and squeeze”) is a massage/manipulation technique encompassing a variety of hand strokes, ranging from gentle and superficial movements to techniques with vigorous articulatory maneuvers. It may be roughly equated with Western manipulative techniques. Tuina is taught in the North America as a part of Chinese bodywork training in schools of oriental medicine.

Craniosacral Therapy

Craniosacral therapy is an osteopathic technique developed by William Sutherland DO in the early 1900s (126). It is a gentle, noninvasive manipulative technique applied to the spinal cord and cranium to correct disruptions in the craniosacral rhythmic activity. This activity is believed to be disrupted in orthopedic trauma, spinal cord and traumatic brain injury; so one of the major focuses of this technique has been this type of injury (151,152). Upledger reports success in treating chronic pain, chronic brain dysfunction, spasticity, and other conditions associated with spinal cord injury. A number of other conditions have also been treated with this technique; however, there is little formal research supporting its efficacy.

Rolfing

Rolfing was developed in the 1950s by Ida Rolf, a chemist who approached the body as a group of units (head, shoulders, trunk, pelvis, and limbs) connected to each other by the fascial network and with gravity (149). She used this concept in her approach to help clients achieve proper vertical alignment and efficient movement. The method is characterized by the release of restrictions in the deep fascial planes in the body, allowing shortened and tense muscles to relax and lengthen. The goal is to balance and align the body in the gravitational field. The practitioner uses a series of ten 60 to 90 minute sessions, working from superficial to progressively deeper structures, each session building upon the previous one. These may be followed by additional sessions after the client has had time to incorporate the changes into daily habit. Movement education may also be used in conjunction with hands-on work. Traditionally, the Rolfing treatment consists of a series of sessions, each concentrating on specific body regions. The intensity of pressure on muscle and connective tissue that characterizes the original style used by Rolf has been refined by many subsequent practitioners of Rolfing. In the rehabilitation setting, Rolfing has been used by therapists integrating the techniques into musculoskeletal and neurological rehabilitation. A case series of patients with cerebral palsy and varying levels of motor impairment suggested that Rolfing was of benefit in improving locomotor capacity in patients with mild CP involvement (153). Structural Integration, developed by a student of Rolf, uses a similar approach and philosophy. Increasing recognition of the importance of the fascia in structural and metabolic health (19) gives a new appreciation for the work of Ida Rolf and her colleagues.

Trager Psychophysical Integration

Trager therapy involves a combination of gentle, hands-on tissue technique movement, reeducation, and relaxation exercises (125). Developed by Milton Trager, MD, in the 1940s, it attempts to teach patients to move with ease and efficiency. The work consists of gentle rocking, stretching, and rolling movements intended to relax and enliven tense areas of the body. The Trager therapist attains a calm and focused state of mind in order to be in contact with the client’s needs and responses. The practitioner manually communicates this light
and energetic state to the client’s body so that the body may relearn a feeling of ease in movement. Tragerwork may be used for the rehabilitation of sports and musculoskeletal conditions. In a controlled trial, Trager was equal to acupuncture for relief of chronic shoulder pain in spinal cord injured patients (154). Case studies suggest its usefulness as a useful adjunct in rehabilitation of chronic obstructive pulmonary disease (155) and cerebral palsy (156).

**ENERGY THERAPIES**

Energy therapies may be divided into biofield therapeutics, or the assessment and adjustment of the human energy field, and electromagnetic therapies (magnets). Biofield therapies are discussed here. Biofield healing is the most controversial of the unconventional modalities, because it postulates the existence of a subtle energy field within and around the body that cannot be measured by standard biomedical instrumentation. Although unrecognized in modern biomedical science, the existence of this subtle energy is an integral part of almost every traditional ethnic medical system, and the concept dates back thousands of years. The energy field is known as the vital life force in Western metaphysical traditions, and as qi or chi (Chinese), ki (Japanese), and prana (Sanskrit) in Eastern traditions. All of the techniques have as a basic underlying concept—the idea that illness is a result of, or at least associated with, imbalances and blockages in this energy field.

The various techniques included in this group of therapies involve a practitioner placing his/her hands on or near the physical body and either actively or passively altering the energy in the recipient’s field. Although light physical touch may be involved, its purpose is to modulate the energy field, not to manipulate the skin, muscle, or other organ, and the mechanism of action is quite different from that usually proposed in manual manipulation techniques. It should be noted, however, that a number of techniques (e.g., Jin Shin Jyutsu [JSJ] and Polarity Therapy, described below) have been developed during the past few decades that combine physical manipulation with subtle energy healing. In traditional systems of medicine, such as Chinese Medicine and Ayurveda, massage and energy healing, while seen as separate techniques, are often used together. Although generally used to promote the overall health of the individual, these techniques have been used to treat specific diseases and medical conditions.

**Prayer and Spiritual Healing**

Prayer is discussed here as a mental process outside the context of any particular religious system. There are a number of studies that demonstrate the efficacy of prayer directed at nonhuman subjects including mice, yeast, bacteria, fungi, flagellates, plants, and seeds, and also for in vitro samples of human tissue, such as red blood cells and cancer cells (157). Comparatively, there are fewer studies involving human subjects, and the research methodology is generally weaker (158–160). Nevertheless, approximately one half of the stronger studies showed some positive effect due to prayer. One of the best controlled studies (a randomized, double-blind study in which the patients, nurses, and doctors did not know the group assignment of the patients) demonstrated a strong, beneficial effect of prayer on coronary care unit patients. The prayed-for group was less likely to require antibiotics or endotracheal intubation, were less likely to develop pulmonary edema, and were less likely to die (161). Similar results were found in a more recent study in a similar setting (162). Prayer also had a positive effect on fertility rates and pregnancy rates after in vitro fertilization (163), general health of AIDS patients (164) and patients with anxiety, depression, and reduced self-esteem (165), but not on wound healing (166) or alcohol abuse patterns (167).

**Therapeutic Touch**

Therapeutic touch (TT) is a technique first described in 1979 and developed by a nurse academician, Dolores Krieger, PhD, Professor of Nursing at New York University, and Dora Kunz, a natural healer. The practitioner uses the hands to sense and locate problems, and then serves as a conduit for universal energy, consciously transferring energy into the client’s energy field in rhythmical, sweeping motions (168). TT is a credentialed nursing skill and it is taught in over 80 universities in over 30 countries.

**Qigong**

Qigong (pronounced chee-gong) is an ancient philosophical system of harmonious integration of the human body with the universe (169). When qigong is practiced, the goal is to balance the qi or vital energy within the body, thus preventing or reducing energy imbalance which can give rise to illness. In self-administered qigong (internal qigong), the individual uses a variety of means, including breathing, visualization, and physical movements to bring the qi into balance. Qi can also be directed toward a recipient by a qigong practitioner, who consciously emits and directs the qi to another person (external qigong).

**Other Energy Therapies**

The general healing nature of the qi is such as to suggest that it can be applied with some success to almost any medical or health problem. Rigorous research is sparse, but McGee and Chow relate a number of case studies for many different conditions, including stroke, paralysis, and cerebral palsy (170). Reiki (pronounced ray-key) is a form of energy healing that usually involves a practitioner laying his or her hands slightly off of 12 specific locations on the body of the person being treated, allowing the “ki” or vital life force to flow through the Reiki practitioner into the body of the patient being treated (171). Having originated in Japan in the 1800s, Reiki has become an increasingly popular and accessible healing technique throughout the Western world. It can also be administered from a distance or self-administered. Training for the lowest level of “Reiki Master” is short in duration, making it an accessible form of energy healing. It is recommended by
induces a state of deep relaxation associated with enkephalin
delay, but little research has been done to elucidate exactly how
this process works. There is some evidence that acupuncture,
when practiced as an integrative treatment, may improve pain
management, immune function, and general health (169). The
authors, however, that only practitioners with significant
Reiki experience (5 years or more) or with training in an
additional form of energy healing work with hospitalized or
debilitated patients.

JSJ or “acupuncture without needles” is an energy heal-
ting technique which uses oriental pulse diagnosis and merid-
ian theory to correct imbalances along energy channels. It is
practiced one-on-one and is recommended for patients who
may not tolerate acupuncture needles. Practitioners place
their hands on specific combinations of acupuncture points
to release energy blockages. In comparison to acupressure,
there is significantly less emphasis on applying physical pressure
and more on adjusting the energy along the meridians. JSJ
for self-administration is also taught in a 1-day format which
allows the patient a wide range of applications. There are no
side effects, making it safe for application on both mild and
severely ill patients. A meta-analysis involving 24 randomized
controlled or controlled clinical trials involving 1,153 partici-
pants examining energy therapies such as TT and Reiki for
pain showed a modest effect on pain (172).

SPECIAL TOPICS: SPORTS MEDICINE,
FUNCTIONAL MEDICINE, CAM IN THE
INPATIENT SETTING

Sports Medicine
The role of the sports medicine physician is to return the
athlete to play as quickly and safely as possible after injury and to
prevent injury and illness. CAM therapies provide the sports
physician with valuable tools to achieve these goals while
avoiding the overuse of medications. The use of CAM is readily
accepted by elite athletes: in a study of Division I NCAA
Intercollegiate athletes, 56% of subjects reported using CAM
over a 12-month time period, higher than the prevalence of
CAM used by all adults in the United States (36%) (173).

Some of the more commonly used modalities are acu-
puncture, manual therapies (massage, osteopathic manual
therapies, and chiropractic), homeopathic preparations, nutri-
tional supplements, meditation and visualization techniques,
and movement therapies (Alexander, Feldenkrais, Yoga) for
the treatment or prevention of injuries. In this brief review,
only acupuncture, use of nutritional supplements for muscle
recovery, homeopathic remedies, and movement therapies, all
discussed more fully above, are addressed.

Acupuncture is widely available and used by athletes at
all levels for acute and chronic musculoskeletal conditions. In
the office setting or on the sidelines, it is used in acute injuries
for the rapid reduction of pain and swelling. This treatment
of injuries within the first 48 hours is referred to as a tendino-
muscular treatment (3). Used on a regular basis for nonacute
conditions, acupuncture is helpful for restoration of flexibility,
joint pain, and ligament sprains. Although tempting from the
point of view of symptom relief, treatment immediately prior
to athletic activity is discouraged. This is because acupuncture
induces a state of deep relaxation associated with enkephalin
and dynorphin release, which could decrease alertness and place
the athlete at risk for injury. A properly trained and licensed
physician/acupuncturist should perform all treatments.

Proper nutrition and adequate caloric intake are extremely
important in ensuring optimal performance. Athletes are
strongly encouraged to follow the guidelines and basic principles
of the anti-inflammatory diet (above). A sports medicine dieti-
cian/nutritionist with background in the nutrition and physical
requirements of a particular sport is an integral part of the elite
athlete’s sports medicine team, and should design a basic team
program with modifications for each athlete based on caloric
intake, gender, metabolism, and sport-specific needs. The
same parameters are used in considering supplements. Athletes
should, of course, be replete in basic vitamins and minerals; a
few of greater importance to athletic performance are outlined
here: L-Carnitine (1 to 3 g/d) and creatine (2 g/d are used as
aids to recovery from exercise (174)); creatine also helps build
muscle tissue (175). There are various minerals advocated for
use in athletes, especially those who experience greater loss due
to excessive sweat. The majority of athletes can obtain adequate
amounts with a well-balanced diet (176). Low levels of mag-
nesium and sodium have been implicated in exercise-induced
muscle cramps and poor exercise performance (176). It should
be noted that serum magnesium levels are not a good measure
of magnesium status; monocyte leukocyte magnesium or urine
levels after a magnesium challenge have been recommended as
a better measure of actual magnesium levels (177,178). Various
anti-inflammatory and antioxidant supplements described pre-
viously may be incorporated in the athlete’s regimen.

Homeopathy is described above and offers a few pow-
erful agents for the athlete: Traumeel (a mixture of homeo-
pathic agents) and Arnica Montana are used commonly for
soft-tissue sprains and soft-tissue injuries (179). Alexander and
Feldenkrais techniques and Yoga, described above, are move-
tment therapies that increasingly have a place in the treatment
of sports injuries. These are especially helpful in restoring or
implementing efficient movement patterns which help the
athlete heal and prevent injuries. The author Lisa Bartoli DO
incorporates yoga, nutrition, Alexander, and acupuncture to
improve strength, flexibility, and balance and as part of the
recovery and treatment of injuries in elite rugby players, run-
ners, and bicyclists.

Functional Medicine
Functional medicine is an emerging approach to complex and
chronic disease. The term was first coined in 1993 by Jeffrey
Bland, PhD, nutritional biochemist and previous Director of
Nutritional Research at the Linus Pauling Institute of Science
and Medicine. Functional Medicine is an emerging paradigm
that, instead of focusing on the primacy of pathology and dif-
ferential diagnosis, focuses on the antecedent physiological
processes that ultimately find their expression in health and
disease. This concept of “upstream medicine” was first elab-
orated by Leo Galland, MD, in his unpublished but widely
disseminated paper, Patient Centered Diagnosis: A Guide to
the Treatment of Patients as Individuals and published in 1997.
The Four Pillars of Healing (180). The tenets of functional medicine are on the leading edge of ideas beginning to be recognized by mainstream medicine and include:

1. Biochemical individuality of the human being leading to the need for personalized medicine that is not necessarily best served by the double-blind randomized controlled study or an assembly line approach to medicine.
2. A patient-centered rather than disease-centered approach.
3. The search for dynamic balance among the internal and external factors in body/mind/spirit.
4. Familiarity with the weblike interconnections of all human biology.
5. Identification of health as positive vitality, not merely the absence of disease.
6. Promotion of organ reserve as a means to expand the health span (181).

Functional medicine reflects a systems biology approach to health care; a comprehensive analysis of the manner in which all components of the human biological system interact functionally with the environment over time (182,183). Just as the physiatrist recognizes multiple interdependent macro aspects of function such as impairment, disability, and handicap, the functional medicine practitioner considers human function on a more fundamental level. These include processes such as communication inside and outside the cell, bioenergetics and energy transformation, issues of replication, repair and maintenance, structural integrity, elimination of waste, protection and defense, issues of transport and circulation, and the influence of environmental inputs such as nutrients and pollutants, as well as psychosocial stressors and the influence of both past physical and psychological trauma. All of these factors determine where the individual sits with respect to the continuum between vitality and disease. Nutritional medicine or “food as medicine” is a central hub of the functional medicine approach.

An example of how a physiatrist may practice functional medicine is the approach to a patient with osteoarthritis who wants to avoid surgery. In addition to recommending weight loss and physical therapy, the physiatrist would request detailed information on concurrent medical conditions, family history of inflammatory and other disease, lifestyle, environment, and nutrition habits. Laboratory values for basic chemistry, metabolic indicators, nutrients, and toxic substances would be evaluated. If suspected, gluten and other food sensitivities, which can result in an overall increase in inflammatory load, would be ruled out. Nutritional, pharmaceutical, and lifestyle interventions would be implemented to return the patient to optimal functioning on structural, biochemical, and mental-spiritual levels. This is but one example of where a comprehensive approach and dietary modification may have a profound effect on health.

Because both physiatry and functional medicine see the patient with chronic disease at the center of a complex web of interactions, physiatrists and their patients may find this approach intellectually and clinically rewarding.

Use of CAM Techniques in the Inpatient Setting

CAM therapies are increasingly integrated into standard rehabilitation care. An example of integration of CAM therapies into this setting is the Atlantic Rehabilitation Institute in New Jersey. Inpatients are offered chair yoga and T’ai Chi by rehabilitation professionals in conjunction with standard rehabilitation care, and have the option of participating in daily qigong class, art therapy sessions, sitting quietly in a dedicated meditation space, or requesting one-on-one massage or energy treatments. Outpatients are referred for yoga, T’ai Chi, stress reduction, or nutritional counseling after PT and OT to help maintain gains made during the formal rehabilitation period, or for acupuncture to complement their current post inpatient physical therapy. Advantages of using CAM modalities within a rehabilitation hospital include increased variety of modalities, ability to treat across the continuum of care, and increased patient interest and confidence. Barriers to integration of conventional and CAM therapies include lack of established standards for practitioner licensing for some modalities, absence of protocols for management of specific conditions, and inconsistent third-party reimbursement. Suggestions for minimizing such barriers include hiring practitioners with experience in both conventional and complementary management of chronic musculoskeletal conditions, informed oversight and education of practitioners, ongoing education of therapists, physicians, and third-party payers. While integration of CAM therapies into the rehabilitation setting is in its infancy, it holds great promise for expanded management options and patient satisfaction.

Conclusions

The judicious use of complementary therapies to supplement standard medical interventions has been a hallmark of the physiatric approach, often under the supervision of a multispecialty team of caregivers. The techniques discussed in this chapter represent only a sampling of the many potentially useful complementary therapies available in a rehabilitation setting. Some techniques, such as acupuncture, phytotherapies, and nutritional supplementation, already have bodies of research that suggest their usefulness in this setting. Other techniques are in an earlier stage of validation. With a historically holistic approach to patient care and the increasing demand for physician-patient partnership and for natural options for prevention and treatment of chronic disease from an aging population, the field of PM&R is perfectly positioned to continue to lead the way in integrating CAM therapies and a holistic approach to patient care into mainstream medicine.

References

1. NHIS 2007 www.nccam.gov 2/06/09;


61. Anonymous. Integration of behavioral and relaxation approaches into the treatment of chronic pain and insomnia. *NIH Technology Assessment Conference, NIH.*
69. Anonymous. Integration of behavioral and relaxation approaches into the treatment of chronic pain and insomnia. *NIH Technology Assessment Conference, NIH.*
PART VII

Scientific Advances
Knowledge is not enough; we must apply. Willing is not enough; we must do.

Goethe, 1829

INTRODUCTION

History and Background

The importance of infusing scientific research into medical education and clinical care was emphasized in the Flexner Report, published in 1910. As Abraham Flexner noted: “If the sick are to reap the full benefit of recent progress in medicine, a more uniformly arduous and expensive medical education is demanded” (1). This observation coincided with numerous scientific breakthroughs that led to repudiation of folk treatments, such as bleeding and purging, in favor of treatments based on scientific method. To meet new challenges, modern medicine developed education standards that emphasized clinical practice and scientific research (2). However, with this innovation, research and clinical practice remained somewhat isolated in separate realms until the emergence of evidence-based medicine.

Evidence-based medicine bridges the gap between clinical practice and research and provides a method to systematically select and apply research findings as part of the clinical decision-making process. Using principles and methods derived from clinical epidemiology, the founders of evidence-based medicine sought to ground medical decision making in rigorous scientific evidence to invalidate previously accepted but ineffective diagnostic tests, prognostic protocols, and clinical treatments and replace them with new, more powerful, accurate, efficacious, and safer ones (3).

Evidence-based medicine offers a framework and objective methods for making clinical decisions by embedding research in the clinical decision-making process. Consequently, evidence-based medicine promotes the translation of research into practice and provides a mechanism for ensuring that research evidence informs clinical decisions (4).

Over the last decade evidence-based medicine attracted considerable attention because it meets an identified need. As noted in the 1990 Institute of Medicine report, Crossing the Quality Chasm, an evidence-based approach to clinical practice provides a sound method for making clinical decisions; reduces idiosyncratic variations in practice; and bridges the gap between knowledge and practice (5).

Despite the potential benefits, there are concerns about the movement towards evidence-based medicine. A literature survey of 47 articles published from 1966 to 1999 noted the most commonly cited barriers to evidence-based medicine: the need to develop new skills; difficulty applying evidence to the care of individual patients; and limited time and resources (6). Over the last 10 years, an explosion of practice resources to support evidence-based practice, which are available via the Internet (Tables 80-1 and 80-2), addressed some of the perceived barriers. These web-based resources provide clinicians with evidence-based practice background information, references, and self-study guides. Clinicians have easy access to evidence summaries and reviews that meet specific quality criteria. Many reviews also include a “clinical bottom-line” that summarizes the clinical importance of the evidence presented.

Criticisms of evidence-based medicine also include the following misperceptions: evidence-based medicine devalues clinical experience, ignores patient values and preferences, and is a cookbook approach to medicine. Dr. David Sackett, one of the founders of evidence-based medicine, states: “Evidence-based medicine is the conscientious, explicit, and judicious use of current best evidence in making decisions about the care of individual patients” (7). Evidence-based medicine does not blindly adhere to specific criteria for patient care. Rather, it serves as an adjunct to clinical decision-making. Finally, evidence-based medicine is not limited to use of randomized controlled trials. Evidence-based decision making emphasizes finding the best available evidence. The evidence is then critically appraised and the relative strengths and weaknesses are considered before applying the findings to clinical practice.

Evidence-based decisions are defined as the integration of best research evidence with clinical expertise and patient values (7). A model for evidence-based clinical decisions emphasizes three elements—clinical state and circumstances, patient preferences and actions, and research evidence—along with the fourth element, clinical expertise, that overlies the other three...
Evidence-Based Practice Internet Resources for the Evidence-Based Practice Process

**TABLE 80.1** Evidence-Based Practice Internet Resources: General

**Self-study**
http://www.hsl.unc.edu/Services/Tutorials/EBM/
EBM tutorial, developed by Duke University Medical Center Library and the University of North Carolina at Chapel Hill Health Science Library, provides a good overview.

**Background information—general resources**
http://www.cebm.net/
The Centre for Evidence-Based Medicine at Oxford. The most comprehensive EBM site. Includes numerous resources and teaching materials in PowerPoint that can be downloaded.

http://www.cebm.utoronto.ca/
The Centre for Evidence-Based Medicine Website at the University of Toronto offers valuable resources including chapters from Sackett’s EBM book. Patient cases include EBP examples for nursing and therapy.

http://www.sheffield.ac.uk/~scharr/ir/netting/
Netting the evidence is intended to facilitate evidence-based healthcare by providing support and access to helpful organizations and useful learning resources, such as an evidence-based virtual library, software, and journals.

http://www.cche.net/usersguides/main.asp
Centre for Health Evidence: The site has links to full-text articles published as “Users’ Guides to Evidence-Based Medicine.” This series was originally published in the Journal of the American Medical Association (JAMA).

**Reference material**
http://www.cebm.utoronto.ca/glossary/index.htm#top
Glossary that provides definitions of EBM terms and examples
http://www.cebm.net/levels_of_evidence.asp
Summarizes levels of evidence with specific criteria for evaluating evidence

**TABLE 80.2** Internet Resources for the Evidence-Based Practice Process

**Step 1: Identifying learning or development need**
http://www.cebm.utoronto.ca/doc/edupres.doc
Word document to keep track of learning issues and develop clinical questions.

**Step 2: Ask focused question**
http://www.cebm.net/focus_quest.asp
An example of the PICO format used to develop focused clinical questions.

**Step 3: Search for evidence**
http://sumsearch.uthscsa.edu/
Organizes results: existing guidelines, systematic reviews, and primary literature.

http://www.guideline.gov
AHRO's National Guideline Clearinghouse. Guidelines are indexed under rehabilitation.

http://www.cochrane.org/reviews/
Cochrane Collaboration—database of abstracts of systematic reviews.

http://www.ahcpr.gov/clinic/epcix.htm
Reports from evidence-based practice centers.

http://www.crd.york.ac.uk/crdweb/
DARE contains over 5,000 abstracts of quality assessed and critically appraised systematic reviews. The database focuses on the effects of interventions used in health and social care.

PEDro—Physiotherapy evidence database—most reviews are critically appraised systematic reviews. The database focuses on the effects of interventions used in health and social care.

http://www.otseeker.com/resources/default.asp
OTseeker provides resources for learning about EBP and a database with abstracts of systematic reviews and randomized controlled trials relevant to occupational therapy—critically appraised and rated.

http://clinicaledvidence.bmj.com/cweb/index.jsp
Search by condition for a list of clinical questions that are linked to evidence synopses, guidelines, and references.

The NLM has a free public access search engine—PubMed. Use the “clinical queries” button to organize a search using EBP filters.

**Step 4: Critically appraise evidence**
http://www.cebm.net/cat_about.asp
Describes a process for developing critically appraised topics (CATs).

http://www.cebm.net/catmaker.asp
Tools for developing critical appraisals, includes worksheets and EBP calculator.

http://web.uccs.edu/fbecker/Psy590/escalc3.htm
Online effect size calculator

http://www.cebm.utoronto.ca/palm/ebmcacal/
Evidence-based medicine calculator (diagnosis, prognosis, intervention)

Challenges to Adopting an Evidence-Based Approach in Rehabilitation Practice

There are several challenges to adopting an evidence-based approach in rehabilitation. Evidence-based medicine was developed for use in general medicine and focuses on mortality or the presence or absence of disease. In contrast, rehabilitation medicine is a unique area of medicine that focuses on functioning, both as an outcome and an area for clinical assessment (9). A unifying model for rehabilitation, developed based on the International Classification of Functioning Disability and Health (ICF) framework, defines rehabilitation as a health strategy that aims to help people “achieve and maintain optimal functioning in interaction with their environment” (10). Because the focus of rehabilitation is so different from general medicine, it is often difficult to translate evidence-based medicine concepts into a framework that is useful for rehabilitation. Another problem is that evidence-based medicine requires
knowledge of clinical epidemiology (11), a field that deals with the solution of clinical problems using epidemiologic methods (12). Few rehabilitation professionals are trained in clinical epidemiology and must acquire evidence-based medicine skills independently through continuing education or self-study. These issues notwithstanding, the evidence-based medicine model has much to offer rehabilitation.

Over the last decade evidence-based medicine has been widely adopted by health care disciplines and is now commonly referred to by a more generic term—evidence-based practice. It is important that the rehabilitation field considers ways to infuse the principles of evidence-based practice into clinical practice. This chapter has the following aims: inspire rehabilitation professionals to adopt an evidence-based approach; help translate evidence-based medicine into terminology that is meaningful to the rehabilitation field; and provide strategies to implement an evidence-based practice approach to clinical practice.

### The Evidence-Based Practice Process

Figure 80-2 presents a schematic diagram of the steps involved in the evidence-based practice process. As the figure
Step 1: Identify Learning or Development Need

The first step in the evidence-based practice process is to identify the learning issue and select the category of research evidence appropriate for the issue. For example, the learning issue may be uncertainty about which tests or clinical examination findings to consider in determining the presence or absence of a disease, impairment, or disorder—evidence related to diagnosis. A learning issue may arise when the patient inquires about the ability to return to work or the family inquires about the likelihood of a patient being able to return to home or to her/his previous level of functioning—evidence related to prognosis. Questions about other rehabilitation interventions that may yield a better outcome present another type of learning issue—evidence related to intervention. For each category of evidence—diagnosis, prognosis, or intervention—it is important to understand specific evidence-based practice concepts and terms.

Step 2: Ask a Focused Question

Once the learning issue and the appropriate category of research—diagnosis, prognosis, or intervention—are identified, a focused clinical question is formulated. A clinical question consists of four parts: patient, intervention, comparison, outcome (PICO). Table 80-3 provides an example of the four-part clinical question. Clinical questions are constructed for each category of evidence. For example, a question related to diagnosis is framed as: For patients with shoulder pain, can clinical tests identify those with rotator cuff tears? A question related to prognosis is posed: For patients with a quadriceps contusion injury, can clinical tests predict the time to return to sports? And a question related to intervention is structured: For elderly individuals with chronic disease, can a strength training exercise program reduce disability? Thus, a well-constructed clinical question establishes a link between patient problems, research evidence, and the clinical decision.

Step 3: Search for Research Evidence

Medicine has witnessed tremendous growth in research over the last century and this growth is particularly evident in the volume of available information. Each year Medline indexes over 560,000 new articles and the Cochrane Center adds 20,000 new trials. The volume of information presents a challenge because published research includes many preliminary or exploratory studies. Knowledge that has been rigorously tested and is of sufficient importance to influence practice represents only a small fraction of the published research. Thus, we do not suffer from a lack of information but from an inability to efficiently process and organize the plethora of available information. Evidence-based practice provides useful strategies for the daunting task of keeping up-to-date on research literature. There are three strategies for conducting an effective and efficient search: (a) organize the search based on the hierarchy of research evidence quality; (b) use evidence-based practice search engines; and (c) search for evidence that has been summarized and rated for quality.

Research Evidence Hierarchy

Since the goal of evidence-based practice is to apply the best available evidence, the literature search begins by looking for the highest quality evidence. A key concept in evidence-based practice is that research evidence is not equal in value. Evidence-based practice ranks research evidence based on a hierarchy that considers the degree to which the study design reduces bias and threats to validity. Professional organizations and evidence-based practice work groups have devised specific guidelines for rating evidence that reflect the following values: controlled studies are ranked higher than uncontrolled studies; prospective studies are considered stronger than retrospective studies; and randomized studies are better than nonrandomized studies. Based on quality considerations, research evidence is categorized on a scale from Level I to V: high quality research with strong methodology is rated at the highest level (Level I) and expert opinion is ranked at the lowest level of evidence (Level V). Studies with weaknesses in research design and implementation or large confidence intervals (CIs) are ranked as Level II to IV.

Systematic reviews play an important role in evidence-based practice. In contrast to traditional opinion-based literature reviews, systematic reviews use a specific protocol to identify, critique, and summarize relevant studies that address a similar clinical question. Bias is minimized in systematic reviews because a comprehensive and reproducible process for searching the literature is outlined. Systematic reviews also assess the methodological quality of studies, which is defined as

<table>
<thead>
<tr>
<th>TABLE 80.3 Four-Part Clinical Question</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>P</strong></td>
</tr>
<tr>
<td>Elements</td>
</tr>
</tbody>
</table>

"all aspects of a study's design and conduct can be shown to protect against systematic bias, nonsystematic bias, and inferential error" (17).

It is important to consider the type of research—diagnosis, prognosis, or intervention—when evaluating the strength of research evidence. For all types of evidence, a valid systematic review of quality research is the highest level of evidence while expert opinion is the lowest. However, for the other levels of evidence, the criteria for research quality are somewhat different for diagnosis, prognosis, or intervention studies. Evidence related to diagnosis compares the diagnostic test results to a "gold" standard, which is generally an accepted proof of the presence or absence of the disease or disorder. Therefore, an important consideration for ranking evidence from diagnostic studies is to determine how the gold, or reference, standard is applied (15). Studies where the gold or reference standard is consistently applied blindly or objectively to all subjects are ranked higher than studies where the reference standard is not consistently applied. Furthermore, studies that validate diagnostic test results from a previous study are ranked higher than exploratory studies (14). Prognosis studies are concerned with examining the effect of patient characteristics on outcomes (15). Inception or prospective studies are rated higher than retrospective studies or studies using untreated subjects from randomized controlled trials (14). Intervention studies evaluate the effect of treatments on outcomes. Randomized controlled trials help to eliminate potential bias and studies using randomized controlled trials are rated higher than cohort and case-controlled studies.

Based on the hierarchy of research evidence, an efficient literature search strategy begins by looking for systematic reviews that summarize the results of many different studies and provide an overview of the body of research evidence that addresses a clinical question. If no appropriate systematic reviews are located, the search continues by focusing on specific well-designed individual studies until the best studies are located.

**Search Engines**

The evidence-based practice movement spawned the development of search engines designed to streamline the literature search process. Some search engines incorporate an evidence-based practice framework. The search is organized by the type of research evidence and the results are presented in order, based on the hierarchy of research evidence. One useful resource is SUMsearch, which is described as a "meta" search engine because it searches multiple databases. SUMsearch results are then organized according to the hierarchy of evidence. SUMsearch is a good starting point for conducting a literature search because it is an efficient strategy for viewing a summary of the available evidence in a specific area and it is especially effective for locating information on systematic reviews and guidelines (18).

Databases of systematic reviews can also be searched directly. The Cochrane Collaboration, an international group dedicated to combining similar randomized trials to produce a more statistically sound evidence through systematic reviews, is one of the best resources for systematic reviews. Complete systematic reviews require a subscription, but abstracts of the reviews are available via a searchable database at Cochrane website. The Cochrane Library includes reviews for the top ten causes of disability in developed and developing nations and, therefore, it is an important resource for evidence-based rehabilitation (19). Moreover, the Cochrane Collaboration includes the study groups dedicated to examining topics relevant to rehabilitation such as bone joint and muscle trauma, movement disorders, multiple sclerosis, musculoskeletal and neuromuscular disorders, and stroke (20).

If the search fails to yield a relevant and valid systematic review or evidence summary, the next step is to search the primary literature. PubMed, a service of the U.S. National Library of Medicine and the National Institutes of Health, provides an extensive database of research abstracts (21). PubMed is widely recognized as the best resource for information from original studies (18). PubMed recently added an evidence-based medicine search filter termed "clinical queries." This feature allows the user to conduct a search based on the category of evidence (diagnosis, prognosis, or therapy) and to request results that are "focused" to include a few of the most relevant studies (specific) or "expanded" to include a wider range of studies (sensitive).

**Evidence Summaries**

Another search strategy for busy clinicians is to search for evidence that has already been summarized and critiqued. The evidence-based practice movement fostered the development of secondary resources or "predigested" evidence summaries (see Table 80-2). Evidence summaries help to address concerns that evidence-based practice demands too much time and requires research critiquing skills beyond the ability of most clinicians. One example, the Database of Abstracts of Reviews of Effects (DARE), is a free, searchable database that includes evidence summaries relevant to rehabilitation (22). Each summary must meet specific quality criteria to be included in the DARE database. The evidence summary reviews key elements of the study and concludes with a critical commentary on the study. Discipline-specific evidence summaries are also available. The Physiotherapy Evidence Database (23) comprises abstracts of evidence-based clinical practice guidelines, systematic reviews, and clinical trials. The methods for the clinical trials are reviewed and rated on a ten-point scale (23). OTseeker is another database with abstracts of systematic reviews and clinical trials relevant to occupational therapy. The clinical trials are rated on an eight-point internal validity score and a two-point statistical reporting score (24). Online journals, such as APC Journal Club and Evidence-Based Medicine, also summarize results of clinical trials. Secondary sources present clinically relevant research that has been critically appraised and provide an important resource for busy clinicians. However, it is still important to examine the source and consider the validity and accuracy of the research summary and appraisal. Evidence summaries streamline the evidence-based practice process by simplifying the critical appraisal of evidence, which is described in the next step.
Step 4: Critically Appraise the Evidence
Once the highest level of evidence for a specific clinical question is located, the evidence is critically appraised. The first item of the critical appraisal process is to assess the validity of the research evidence. Unfortunately, specific standards for determining the validity of a study have not been established and a proliferation of critical appraisal tools has led to considerable confusion. One systematic review identified 121 published critical appraisal tools. While the tools vary somewhat, the most frequently cited areas evaluated to establish the validity of research evidence for efficacy studies were: eligibility criteria, appropriate statistical analysis, random allocation of subjects, consideration of outcome measures used, sample size justification/power calculation, study design reported, and assessor blinding (25). The critical appraisal process, which varies for different types of research evidence, is described in the following sections.

Systematic Reviews and Meta-Analyses
Similar to systematic reviews, meta-analyses use a specific protocol to locate and assess research articles addressing a specific question. However, meta-analyses go one step further by combining data to yield a summary statistic. The focus of the meta-analysis or systematic review should be clear with detailed information on the population, intervention and outcomes. Systematic reviews should meet criteria for establishing homogeneity. Specifically, the studies should be similar in terms of patient characteristics (e.g., age, type of disease/disorders), interventions used, outcomes measured, and the study methods (e.g., randomized trials, cohort studies) (26). The search strategy used to identify research articles should include multiple search engines to ensure that all relevant studies are considered (27). The process used to include studies should be outlined and the criteria used to evaluate studies clearly stated (28). Each study should be evaluated in terms of the methodological quality, precision or the width of the CI around the result and the external validity, or the extent to which the results can be generalized (27).

As the number of systematic reviews continues to grow, it is important to ensure that the information is up-to-date. Unfortunately, there is little research to support how and when to update systematic reviews, so it is important to consider if the information presented is still relevant (29).

Evidence Related to Diagnosis
Evidence related to diagnosis is concerned with determining the extent to which tests for specific conditions are able to distinguish individuals with and without a specific disease, impairment, or disorder. Results from the diagnostic test and the “gold” standard are organized in a 2 × 2 contingency table and analyzed to determine how well the diagnostic test sorts patients into the appropriate “bins,” based on findings from the “gold” standard test. Often clinical tests are compared to an invasive surgical, radiological, or electrodiagnostic “gold” standard test; therefore, substituting a less invasive or less expensive clinical test is preferable. There are four possible outcomes when a diagnostic test is evaluated (Table 80-4): A—the test is positive and correctly identifies persons with the condition; B—the test is positive but the person does not have the condition; C—the test is negative but the person has the condition; D—the test is negative and the person does not have the condition.

The ability of a diagnostic test to correctly identify patients with a disease, impairment, or disorder is termed the test’s sensitivity, which is calculated as the proportion of patients with the disorder who have a positive test result. The test’s ability to identify patients without a disease, impairment, or disorder is termed as the test’s specificity, which is calculated as the proportion of patients without the disorder who have a negative test result. Ideally, tests are sufficiently sensitive and specific so that individuals with the disease, impairment, or disorder are identified and receive the appropriate interventions and individuals without the disorder are identified and not subjected to unnecessary treatments or restrictions.

Likelihood ratios combine sensitivity and specificity and provide another way to summarize the value of a diagnostic test. A positive likelihood ratio is the odds that a patient with a condition will have a positive test result compared to a patient without the condition. A negative likelihood ratio is the odds that a patient without a condition will have a negative clinical test result compared to a patient with the condition.

<table>
<thead>
<tr>
<th>Test positive</th>
<th>“Gold” standard identifies that disease/disorder is present</th>
<th>“Gold” standard identifies that disease/disorder is absent</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>A: Correct identifies condition</td>
<td>B: Incorrect over identifies condition</td>
<td>A + B</td>
<td></td>
</tr>
<tr>
<td>Test negative</td>
<td>C: Incorrect misses condition</td>
<td>D: Correct identifies absence of condition</td>
<td>C + D</td>
</tr>
<tr>
<td>Totals</td>
<td>A + C</td>
<td>B + D</td>
<td></td>
</tr>
</tbody>
</table>

Sensitivity = \( A/(A + C) \). Does a test correctly identify when the condition is present?
Specificity = \( D/(B + D) \). Does a test correctly identify absence of the condition?
The **positive predictive value** is the proportion of patients with positive test results who have the condition and **negative predictive value** is the proportion of patients with negative test results who do not have the condition. Positive and negative predictive values provide a simple and clear way to represent diagnostic findings; however, because these values are affected by prevalence they must be interpreted with caution. A high prevalence will inflate the positive predictive values and lower the negative predictive values. Conversely, a low prevalence will lower the positive predictive value and inflate the negative predictive value. **Prevalence**, or pretest probability, estimates how common a disease or disorder is in the population. Prevalence estimates are derived from the study sample, the literature, or data from a specific clinical setting.

Several criteria must be met to establish the validity of research on a diagnostic test. First, the study should include an independent, blind comparison of the diagnostic test with a “gold” standard. It is important to critique the “gold” standard to ensure that it is, in fact, an accurate measure of the presence or absence of a disease or condition and that it was consistently applied to all test subjects. Next, consider the study design and characteristics of the subjects. Ideally, a prospective cohort design is used and the sample includes subjects who represent the continuum of severity for the disease or disorder. If a study includes subjects without symptoms or is comprised mainly of subjects with severe symptoms, the study may misrepresent the ability of the diagnostic test to identify persons with and without the disease or disorder. Before applying the results of a diagnostic study, examine the characteristics of the subjects to be sure that they are similar to the patient group for whom you intend to use the diagnostic test. Also, examine the methods section to determine if the diagnostic test was conducted in a manner consistent with its intended use. Finally, determine if the test was validated in a second, independent group of patients. If all of these criteria were met, the validity of the diagnostic study is established (30).

The next question is to determine the clinical importance of the diagnostic test by examining the ability of the test to correctly categorize patients with the condition of interest. Individual research studies may report results in terms of sensitivity, specificity, likelihood ratios, or predictive values (31). If results of the diagnostic study are not reported using the desired statistics, it may be possible to extract relevant data and calculate the variables of interest. With the use of calculators, which are available online, it is relatively easy to calculate different diagnostic variables.

Results from tests with continuous numeric scores can also be used to categorize patients for diagnostic purposes. A specific score is used as a cut-off point and patients with scores above or at and below the cut-off point are categorized based on the presence or absence of a trait, condition, or functional ability. For example, a prospective study followed elderly individuals for 1 year and used their scores on the Tinetti balance scale to identify elderly individuals who fell within the year (32). A **receiver-operator characteristic curve**, which plots sensitivity and 1-specificity of the different test scores, was examined to select a cut-off score that maximized the sensitivity and specificity of the test. As noted in Figure 80-3, a cut-off point score of 33 on the Tinetti scale yields a sensitivity of 51% and a specificity of 74%. This means that only 51% of elderly subjects with Tinetti scores at or below 33 were correctly identified as fallers, while 74% of nonfallers were correctly categorized. Increasing the cut-off point to a score at or below 36 increased the sensitivity to 70%. With a more stringent cut-off point more fallers were correctly identified. As noted in the previous example, sensitivity increases when the test is effective in ruling out the disorder. However, the increase in sensitivity was accompanied by a loss of specificity; now only 52% of nonfallers were correctly identified. This example illustrates the importance of considering the goal of the diagnostic test when selecting a cut-off point.

An example of a critical appraisal of a diagnostic study is presented in Table 80-5. In this study, clinical tests are compared to the “gold” standard that used joint arthroscopy to identify which patients had rotator cuff tears. This example shows that the clinical tests have low sensitivity (24%). Of the 200 patients with confirmed rotator cuff tears, only 48 were correctly identified with the clinical tests. The clinical tests were negative in 152 cases where arthroscopic findings showed a confirmed rotator cuff tear. Thus, a negative result on this clinical test was not effective in ruling out this condition. The sensitivity of a test increases when a **negative result** is able to rule out a disease or disorder. In this example with low sensitivity, more than 75% of the cases with negative results actually had a rotator cuff tear. In contrast, the specificity for the clinical tests is high (99%). There was only one case where the clinical tests indicated a rotator cuff tear that was not confirmed.
TABLE 8.5 Critical Appraisal: Diagnostic Study

1. Is this diagnostic study valid? (7)

Was there an independent, blind comparison with a reference (“gold”) standard of diagnosis? Prospective study using arthroscopic results as the “gold” standard. The authors completed clinical tests prior to grouping by arthroscopic results.

Was the diagnostic test evaluated in an appropriate spectrum of clients (like those in whom it would be used in practice)? Yes

Was the reference standard applied regardless of the diagnostic test result? Yes

Was the test (or cluster of tests) validated in a second, independent group of clients? Yes

Interpretation: This diagnostic study is valid.

2. Are the results of this valid diagnostic study important?

<table>
<thead>
<tr>
<th>Diagnosis of rotator cuff tear by three positive clinical tests: supraspinatus weakness, weakness in external rotation, positive impingement</th>
<th>Present</th>
<th>Absent</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive diagnosis of rotator cuff tear by arthroscopy</td>
<td>48</td>
<td>a</td>
<td>b</td>
</tr>
<tr>
<td>Negative diagnosis of rotator cuff tear by arthroscopy</td>
<td>152</td>
<td>c</td>
<td>d</td>
</tr>
<tr>
<td>Total</td>
<td>200</td>
<td>200</td>
<td>400</td>
</tr>
</tbody>
</table>

**Definition**

- Sensitivity: Ability to correctly identify someone with the disorder
- Specificity: Ability to correctly identify someone without the disorder
- Likelihood ratio (positive test result): The likelihood that a positive test result was observed for someone with a rotator cuff tear compared to someone without a rotator cuff tear.
- Likelihood ratio (negative test result): The likelihood that a negative test result was observed for someone without a rotator cuff tear compared to someone with a rotator cuff tear.
- Positive predictive value: The proportion of subjects with positive results who have confirmed rotator cuff tears.
- Negative predictive value: The proportion of subjects with negative test results who do not have confirmed rotator cuff tears.
- Pretest probability (prevalence): How common is the diagnosis of rotator cuff tear within the population (of this study)?
- Pretest odds: The odds of a rotator cuff tear before the tests are done.
- Posttest odds: The odds of a rotator cuff tear after the test has been carried out.
- Posttest probability: The probability of a rotator cuff tear after the results are obtained.
- Diagnostic accuracy: Percentage of people correctly diagnosed.

**Formula**

- Sensitivity: $\frac{a}{a + c}$
- Specificity: $\frac{d}{b + d}$
- Likelihood ratio (positive test result): $\frac{sens}{1 - sens}$
- Likelihood ratio (negative test result): $\frac{1 - sens}{spec}$
- Positive predictive value: $\frac{a}{a + b}$
- Negative predictive value: $\frac{d}{c + d}$
- Pretest probability (prevalence): $\frac{a + c}{a + b + c + d}$
- Pretest odds: $\frac{prev}{1 - prev}$
- Posttest odds: $\frac{prev \times +LR}{1 - prev \times +LR}$
- Posttest probability: $\frac{post test odds}{post test odds + 1}$
- Diagnostic accuracy: $\frac{100 \times a + d}{a + b + c + d}$

**Example**

- Sensitivity: $\frac{48}{48 + 152} = 0.24 = 24\%$
- Specificity: $\frac{199}{1 + 199} = 0.995 = 99.5\%$
- Likelihood ratio (positive test result): $\frac{0.24}{1 - 0.24} = 0.76$
- Likelihood ratio (negative test result): $\frac{1 - 0.24}{0.24} = 0.995$
- Positive predictive value: $\frac{48}{48 + 1} = 0.98 = 98\%$
- Negative predictive value: $\frac{199}{152 + 199} = 0.57 = 57\%$
- Pretest probability (prevalence): $\frac{a + c}{a + b + c + d}$
- Pretest odds: $\frac{0.20}{1 - 0.2} = 0.25$
- Posttest odds: $\frac{0.25 \times 48}{1 - 0.25 \times 48} = 12$
- Posttest probability: $\frac{12}{12 + 1} = 0.92 = 92\%$ (or use nomogram)
- Diagnostic accuracy: $\frac{100 \times 48 + 199}{400} = 61.7\%$
### Table 80.5 Critical Appraisal: Diagnostic Study (Continued)

<table>
<thead>
<tr>
<th>CI (95%)</th>
<th>Formula</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive predictive value</td>
<td>( SE = \sqrt{\frac{p \times (1-p)}{n}} )</td>
<td>( SE = \sqrt{\frac{0.98 \times (1-0.98)}{400}} = 0.01 )</td>
</tr>
<tr>
<td>Note: This formula is also used for sensitivity and specificity</td>
<td>95% CI = ( p \pm (1.96 \times SE) )</td>
<td>95% CI = 0.98 ± 0.01=0.97–99</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Value of a Test—Based on Likelihood Ratio (54)</th>
<th>( +) Likelihood Ratio</th>
<th>(-) Likelihood Ratio</th>
<th>Test Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>1.0</td>
<td>Useless (pretest probability = posttest disease probability)</td>
<td></td>
</tr>
<tr>
<td>1.0–2.0</td>
<td>0.5–1.0</td>
<td>Usually none to small</td>
<td></td>
</tr>
<tr>
<td>2.0–5.0</td>
<td>0.2–0.5</td>
<td>Small to moderate</td>
<td></td>
</tr>
<tr>
<td>5.0–10.0</td>
<td>0.1–0.2</td>
<td>Moderate to large</td>
<td></td>
</tr>
<tr>
<td>&gt;10.0</td>
<td>0.0–0.1</td>
<td>Large</td>
<td></td>
</tr>
</tbody>
</table>


by arthroscopic findings. So it is likely that a patient with a positive test does have a rotator cuff tear. When a test has high specificity, a positive result is able to effectively rule in the target condition. This example demonstrates the importance of considering both sensitivity and specificity when selecting a diagnostic test. Because this test has high specificity, a positive result on the clinical tests is a good indication that the patient has a rotator cuff tear; however, due to the low sensitivity a large percentage of cases with negative test results will actually have a rotator cuff tear that is missed.

In this example, the positive likelihood ratio was large (48) and the negative likelihood ratio was useless (0.76). The positive predictive value showed that 98% of study subjects with positive clinical test results had rotator cuff tears and the negative predictive value showed that 57% of study subjects with negative tests did not have rotator cuff tears. However, the predictive values are affected by the prevalence of the condition. The pretest probability (prevalence) for rotator cuff tears for subjects between the ages of 30 and 39 was approximately 20%. After the three clinical tests were conducted and results obtained, the probability of having a rotator cuff tear increased to 92% and the overall diagnostic accuracy was 61.7%. The diagnostic information presented in this study can help clinicians understand the benefits and limitations of using clinical tests in patients with a suspected rotator cuff tear. This information can also be communicated to patients when explaining the clinical findings.

**Evidence Related to Prognosis**

Clinicians directly consider questions about a patient's prognosis all the time. At times prognostic questions are posed by patients such as, “Will I regain my pre-fracture level of function?” Or “How long will it take for me to recover from this injury and return to work?” At other times clinicians will consider prognosis less directly, such as when deciding where to discharge a patient at the end of an acute hospital stay. A question such as, “Is this patient a good fit for inpatient rehabilitation?” is often based on a clinician’s judgment of prognosis. Prognosis reflects a clinician’s estimation of a patient’s likely clinical course over time taking into account the likely complications of their disease, injury, or impairment. There are several different aspects to prognosis: a qualitative aspect (i.e., What outcomes could happen to this patient?); a quantitative aspect (i.e., How likely are the outcomes to occur to this patient?); a temporal aspect (i.e., Over what time period are the anticipated outcomes likely to occur?); and an evaluative aspect (i.e., What is likely to influence the outcomes?).

In rehabilitation, clinicians frequently base their prognostic judgments and advice on their clinical experience and less often on the basis of scientific evidence. Later in this section, we will consider a framework for appraising the validity, importance, and applicability of evidence using similar guidelines applied to evidence regarding diagnosis and treatment.

To evaluate prognostic evidence, the clinician needs to appraise three distinct and important elements: the validity of the evidence, the importance, and the applicability of evidence to the patient(s).

In determining whether prognostic evidence is valid several different factors must be considered. First, the clinician must appraise if the evidence is based on a defined, representative sample of patients assembled at a common point in the course of their disease of episode of care (usually at the start). A second validity issue is the degree to which patient follow-up was complete and sufficiently long. This judgment must
be made in relation to the outcome being considered. A short follow-up with just a select few patients would not be sufficient evidence to use for advising patients. If the follow-up is long enough, the clinician still needs to assess which subjects entered the study but were lost to follow-up since a substantial loss to follow-up can significantly affect conclusions drawn about patient prognosis. For example, if patients who perform poorly are lost to follow-up this would introduce significant bias into any prognostic determination.

The next issue is whether objective outcome criteria were defined and tracked with adequate validity. To minimize the potential effects of bias in outcome measurement, objective criteria should have been established and implemented in the collection of outcome data. In the best case, those making outcome measurements are kept in the dark (i.e., blind) to a patient’s clinical characteristics and prognostic factors to minimize the likelihood of bias creeping into the assessment.

Once the validity of a prognostic investigation has been established, the next issue is to determine its importance and appropriateness by examining the characteristics of the study subjects and evaluating the degree to which the prognostic information is relevant to an individual patient’s situation and/or condition. Finally, CI around the prognostic estimate should be examined. A prognostic estimate with a relatively small CI provides greater assurance that the information will be important and useful to patients.

An example of a critical appraisal of a prognosis study is presented in Table 80-6. The study examined the ability of the Orpington prognostic scale (OPS) to predict functional recovery for patients at 3- and 6-months poststroke. The study met most validity criteria, but it was noted that patients with severe and mild symptoms were not well represented. Therefore, when this evidence is applied to an individual patient it is important to remember that the findings may differ for patients with

### Table 80.6 Critical Appraisal: Prognosis Study

<table>
<thead>
<tr>
<th>1. Are the results of this prognosis study valid?</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Were the study subjects representative?</td>
<td>Yes—the sample consisted of patients with a broad age range with good representation of minorities. Subjects were drawn from 12 area hospitals. Rehabilitation occurred in a variety of settings.</td>
</tr>
<tr>
<td>Did the research study a broad spectrum of patients with the disorder?</td>
<td>No—patients with severe and mild stroke were underrepresented.</td>
</tr>
<tr>
<td>Were patients at a similar point in the course of their disease?</td>
<td>Yes—all subjects had baseline OPS measures completed between 3 and 14 d poststroke. Follow-up functional assessments were at 3- and 6-mo intervals.</td>
</tr>
<tr>
<td>Were researchers recording the outcomes blinded to the prognostic factors?</td>
<td>Not clear</td>
</tr>
<tr>
<td>Were outcomes well-defined and relevant? Were criteria for measuring outcomes preestablished?</td>
<td>Outcome measures were functional activities identified as important to individuals of both genders.</td>
</tr>
<tr>
<td>Were outcome measures established prior to the start of the study?</td>
<td>Yes</td>
</tr>
<tr>
<td>Was follow-up sufficiently long and complete?</td>
<td>Drop out rate was &lt;20%. Reasons were documented. Characteristics of patients who did not complete the 6-mo follow-up indicate that they were more severely involved.</td>
</tr>
<tr>
<td>What was the drop out rate? Were reasons for drop out documented?</td>
<td>No</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>2. Are the results of this prognostic study important?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prevalence of independence in functional activities based on OPS baseline score of 3.6</td>
</tr>
<tr>
<td>Function</td>
</tr>
<tr>
<td>Personal care</td>
</tr>
<tr>
<td>Meal preparation</td>
</tr>
<tr>
<td>Medication administration</td>
</tr>
<tr>
<td>Community mobility</td>
</tr>
</tbody>
</table>

Adapted from Critical Appraisal Worksheets provided by the Centre for Evidence-Based Medicine, http://www.cebm.utoronto.ca/practise/ca/prognosis.

severe or mild symptoms. At 6-months poststroke, patients with initial OPS scores of 3.6 achieve the following levels of functional ability (note: range includes 95% CI): 56% to 71% were independent in personal care; 52% to 68% were independent in meal preparation; 48% to 63% were independent in medication administration; and 30% to 46% were independent in community mobility. This information can be useful when discussing expectations for functional outcomes with individual patients and their families.

Evidence Related to Intervention

Clinicians have many options in determining a course of care for their patients, but too often treatment decisions are based on a clinician’s experience or preference. Many studies demonstrate disparities in clinical care and point to the need for more equitable and evidence-based care (5). The goal of evidence-based practice is to help clinicians work with their patients to select interventions that are likely to produce the desired outcomes, are worth the costs and the efforts of using them and fit with patient preferences. Evidence-based practice provides an objective framework and a systematic approach to guide the clinical decision-making process. Evidence-based practice focuses on determining the importance and clinical significance of research related to interventions so that their relative benefits can be better understood, communicated to the patient, and used to make decisions.

Critical appraisal of intervention studies begins by examining the outcome measure used in the study to determine if it is appropriate for answering the clinical question. Specifically, the dimension of outcome measure (e.g., impairment, functional limitation, disability) should match the outcome identified in the clinical question. It is also important to establish that the outcome measure used is valid and reliable.

The critical appraisal process continues by examining the methods section. Were subjects randomly assigned to groups? Were researchers (single) and, if possible, subjects (double) blind to group? Next, review the baseline characteristics and demographic information for the control and experimental groups to see if they are similar. If differences exist, could they affect the study results? Were subjects who withdrew or did not complete the study accounted for? And were subjects analyzed in the groups to which they were randomized (intention to treat analysis)?

Review the methods section to determine if both groups were treated equally, except for the experimental intervention. Determine if the follow-up was sufficiently long and complete. If the results were not statistically significant, the study should include a power analysis to determine if the number of subjects in the study was sufficient to find a difference if, in fact, one existed (33).

Once the validity of a study is established, the next step is to determine if the results are of a sufficient magnitude to warrant a change in practice. Research studies comparing outcomes for different interventions, typically, report results in terms of statistical significance or p-values. The statistical significance is determined before the study begins and, most often, the significance level is set at \( p < 0.05 \), the actual probability of obtaining the observed result if the null hypothesis (that there is no difference) is true. If the \( p \)-value is less than the identified significance level, the null hypothesis is rejected. Statistical significance is regularly reported in research studies and is the primary factor used to evaluate research findings. However, there is a difference between statistical significance and the magnitude of the treatment effect or clinical significance. Statistical significance conflates the magnitude of the difference between groups and the sample size. Since sample size affects statistical significance, a rather small treatment effect can yield statistically significant results if the sample size is large enough. Evidence-based practice focuses on the magnitude of the effect or clinical significance of research findings so that they can guide in clinical decision-making.

Results from an intervention study where data are reported as proportions can be summarized as the proportion of patients with good or bad outcomes. The experimental event rate (EER) and the control event rate (CER) represent the proportion of subjects who achieve the desired outcome in the respective groups. Using these values, one can determine the relative benefit increase (RBI), or increase in good outcomes in the experimental group. This useful statistic helps quantify the magnitude or clinical significance of the intervention. The EER and CER are also used to calculate the absolute benefit increase (ABI), which is the absolute difference between rates of good outcomes in experimental and control groups. Finally, ABI can be used to calculate the number needed to treat (NNT) to achieve one additional good outcome. The NNT calculation is a useful statistic that captures the effort required to see the effect of an intervention (34). Moreover, the NNT statistic, in conjunction with financial data, can be used to determine cost/benefit ratios.

In evidence-based practice examples, treatment effects are often described in terms of dichotomous variables, such as the proportion or percentage of subjects that lived/died or were cured/not cured. Rehabilitation outcomes are not usually defined in terms of life and death or cures but the following are examples of dichotomous variables that are relevant to the rehabilitation field: returning to work or not, being discharged to home or skilled nursing facility, or achieving a target level on a functional test or not. When dichotomous variables are reported as percentages or proportions, these data can be used to quantify the magnitude or clinical significance of the treatment effect by calculating the ABI and the NNT.

The NNT calculation is not often reported in rehabilitation literature; however, it is a useful calculation. For example, if the effectiveness of an intervention is examined based on the proportion of patients discharged to home versus skilled nursing facility, the NNT calculation can estimate the number of patients that would have to be treated with this intervention to yield one additional discharge to home. The NNT calculation helps answer questions about the efficacy and clinical importance of an intervention. It is reported with CIs to indicate the precision of the estimate. Because
NNT summarizes the effort required to produce positive outcomes; it can be used to calculate the costs and benefits of an intervention (35).

If data in a study are summarized as means and standard deviations, clinical importance is assessed by calculating the effect size, which considers the size of the difference between two groups relative to the variability of the outcome in the patient group without the confounding influence of sample size. The effect size, which is the standardized mean difference between two groups, is not often mentioned in research reports but is relatively easy to calculate (36).

\[
\text{Effect size} = \frac{\text{Mean of experimental group} - \text{Mean of control group}}{\text{Standard deviation (pooled)}}.
\]

Effect sizes are reported with CIs and are used to categorize the magnitude of the treatment effect as follows: large (anything >0.5), moderate (0.5 to 0.3), small (0.3 to 0.1), and trivial (anything <0.1). Knowledge of the effect size is helpful when determining if the evidence is strong enough to change practice. It is also important to consider other aspects of the research design when evaluating the strength of evidence. A small or moderate effect size may be more meaningful in rehabilitation research because the variability among subjects (standard deviation) reduces the effect size.

Evidence-based practice calculators can ease the burden of determining the clinical significance of research evidence. Free, online calculators are available for diagnostic tests, and for calculating the NNT and effect size (see Table 80-2).

An example of a critical appraisal of an intervention study is presented in Table 80-7. The study examines the effect of an individual counseling and self-management program on bed disability days in elderly individuals. The EER shows that subjects with one or more bed disability day at baseline had a 29% increase in the rate of improving the number of bed disability days compared to the control group. According to the ABI calculation, we can be 95% confident that the increase in the rate of good outcomes is between 11% and 47%. For every three elderly individuals who participated in the program, one additional person improved in bed disability days. The NNT calculation indicates that we can be 95% confident that we will see one additional good outcome for every two to nine elderly individuals who participate in the program. These calculations can help determine the importance and economic impact of interventions. This information can also be part of patient communication.

**Step 5: Apply Research Evidence**

Applying research evidence to individual patients is one of the most challenging aspects of evidence-based practice. While specific rules and procedures serve as a guide in the preceding evidence-based practice steps, applying research evidence to individual patient cases requires sound clinical judgment. Indeed, applying evidence to patients assumes that results from research studies can be generalized when, in fact, patients present with a wide range of differences (37). One of the greatest challenges of evidence-based practice is to determine how an individual patient’s age, disease characteristics, or comorbidities, which may differ from subjects in the study sample, affect the applicability of the evidence. For this reason, clinical expertise is a critical component of evidence-based practice.

The eligibility criteria for subjects in clinical trials must be considered when applying research findings to individual patients. For example, effectiveness studies have wide eligibility criteria and participation is extended to a heterogeneous group of subjects. And, because the results were seen in a heterogeneous group of subjects, it is easier to make a case that similar results will be seen in a wide range of patients. In contrast, efficacy studies have narrow eligibility criteria and participation limited to a homogenous group of highly responsive patients (38). Consequently, one must be cautious about applying the results of an efficacy study to the general population that includes a range of individuals who are not likely to be as responsive (39).

Communication with patients to explain the options and determine their preferences is an important part of the decision-making process. The recent increase in consumerism and emphasis on patient-centered care reinforces the need to consider patient preferences when deciding on treatment choices. Patient involvement is a significant component of evidence-based practice and there are many related benefits. Active patient involvement in care is associated with improved outcomes and enhanced quality of life (40,41). Informed patients are more likely to participate in their care and make better decisions (42). Since patients and their families often have access to a wealth of information through the Internet, they frequently bring along the results of their searches and provide the impetus to begin a discussion of diagnostic, prognostic, and treatment options.

Clinicians can effectively communicate with patients by clearly summarizing research evidence, outlining treatment options, and discussing estimates of potential benefits and risks. The discussion should explore the patient’s values about the therapy and potential options (39). It is important to build a partnership and understand the patient’s experience and expectations. The discussion of evidence should include a balanced presentation of options and, after consideration of the clinical evidence and patient’s values, the clinician can recommend an option and explain how the choice is consistent with the patient’s values. Finally, it is important to make sure that the patient understands and agrees with the recommendation (42).

**Step 6: Evaluate the Use of Evidence**

The final step in the evidence-based practice process involves an evaluation of the process and the use of evidence. The evaluation includes a reflection on how well the process worked. This step presents an opportunity to identify areas for improvement and to implement strategies for continued growth in becoming an evidence-based practitioner. The evaluation step also presents an opportunity to investigate the impact of evidence-based decisions to determine if changes in practice had an effect on outcomes.
### TABLE 80.7 Critical Appraisal: Intervention Study

1. **Are the results of this therapy study valid?**

   - Was randomization maintained? Yes
   - Were all patients who entered into the trial accounted for? Yes
   - Were all patients analyzed in the groups to which they were randomized (intention to treat analysis)? Yes
   - Were researchers (single) and patients (double) blind to treatment group? Not specified, but data collectors were not involved in the intervention.
   - Were baseline characteristics of the two groups similar? No, more intervention subjects were women, not married, living alone, had diabetes, were more restricted in activity days at baseline. Baseline status on outcome, gender, and age were adjusted in analyses.

   Except for the experimental intervention, were both groups treated equally? Yes, but difficult to control in this community-based study.

   - Was follow-up sufficiently long and complete? Only 1 y. Previous study showed effects were not sustained after 1 y.
   - Was a power analysis performed? Yes. Individual randomized control trial.

   - Level of evidence? Individual randomized control trial.

   - Dimension of outcome measure? Are the measures appropriate for answering the clinical question? Study includes administrative, impairment, activity, and participation measures; all are appropriate for answering the clinical question.

   - Were the measures valid and reliable and used in other studies? Yes.

   **Interpretation:** This intervention study is valid.

2. **Are the results of this valid intervention study important? How large and precise is the treatment effect?**

   **Number of bed disability days for patients with ≥1 bed disability day at baseline**

<table>
<thead>
<tr>
<th></th>
<th>Improved</th>
<th>Same or Worse</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental</td>
<td>7</td>
<td>a</td>
<td>b 17</td>
</tr>
<tr>
<td>Control</td>
<td>0</td>
<td>c</td>
<td>d 15</td>
</tr>
<tr>
<td>Total</td>
<td>7</td>
<td>32</td>
<td>39</td>
</tr>
</tbody>
</table>

   **Experimental event rate (EER)**

   Proportion of patients in the experimental group who showed a positive effect

   $$EER = \frac{a}{a + b}$$

   $$\text{EER} = \frac{7}{7 + 17} = 0.29 = 29\%$$

   **Control event rate (CER)**

   Proportion of patients in the control group who showed a positive effect

   $$CER = \frac{c}{c + d}$$

   $$\text{CER} = \frac{0}{0 + 15} = 0 = 0\%$$

   **Relative benefit increase (RBI)**

   Proportional increase in rates of good outcomes between experimental and control patients

   $$\text{RBI} = \frac{\text{EER} - \text{CER}}{\text{CER}}$$

   $$\text{RBI} = \frac{0.29 - 0}{0} = \text{cannot calculate}$$

   **Absolute benefit increase (ABI)**

   The absolute difference in rates of good outcomes between experimental and control patients

   $$\text{ABI} = \text{EER} - \text{CER}$$

   $$\text{ABI} = 0.29 - 0 = 29\%$$

   95% CI: 11%–47%

   **Number needed to treat (NNT)**

   The number of patients needed to be treated to achieve one additional good outcome

   $$\text{NNT} = 1 / \text{ABI}$$

   $$\text{NNT} = 100/29 = 4$$

   95% CI: 2–9

95% CI for ABI.

$$95\% \text{ CI} = \pm 1.96 \frac{\text{EER}(1 - \text{CER})}{\text{No. of control pts}} \frac{\text{CER}(1 - \text{EER})}{\text{No. of treatment pts}}$$

95% CI = ±1.96 $\sqrt{0.29(1 - 0.29)} = 0.18$. 

95% CI = 0.29 ± 0.18 = 11%–47%.

95% CI for NNT (calculated from ABI).

100/CI of lower limits of ABI = 100/11 = 9.

100/CI of upper limits of ABI = 100/47 = 2.

Adapted from Critical Appraisal Worksheets provided by the Centre for Evidence-Based Medicine, [http://www.cebm.utoronto.ca/practise/ca/therapyst](http://www.cebm.utoronto.ca/practise/ca/therapyst).

Becoming an Evidence-Based Practitioner

Assessing Knowledge and Skill

The first step in becoming an evidence-based practitioner is to identify individual learning needs. Specific evidence-based practice competencies, such as searching the literature (43) and critical appraisal skills (44–46) are often assessed. Table 80-8 presents a list of evidence-based practice competencies that can be used as a guide to gaining knowledge and skill. Validated tests of evidence-based practice, such as the Berlin Test (47) or the Fresno Test (48), are available but they emphasize general medical cases and examples.

The prospect of developing all of the skills needed to be a competent evidence-based practitioner can seem overwhelming. It is not necessary to be accomplished in all areas of evidence-based practice, but it is important to start and continue to grow. It may also be helpful to consider stages of developing evidence-based practice abilities, as described below.

- Stage 1: I use evidence-based practice guidelines or protocols developed by colleagues but do not search the literature or critically appraise research.
- Stage 2: I seek and apply evidence-based summaries that give a clinical “bottom line” but I am not comfortable appraising and applying findings from primary literature.
- Stage 3: I have expertise in all evidence-based practice knowledge and skills. I am able to locate, appraise, and apply findings from primary literature and systematic reviews.

Considering evidence-based practice skills as stages makes it possible for every clinician to be an evidence-based practitioner, without mastering all areas of knowledge and skill.

Implementation Strategies

The gap between evidence and practice is widely acknowledged (49) but the movement from best evidence to best practice is not simple. Strategies to become an evidence-based practitioner can be developed at three different levels: personal, professional group, and institutional. On a personal level, an individual clinician can identify knowledge gaps and learning issues. These learning issues can be addressed by collaborative efforts among colleagues, attending workshops and courses or by developing a self-study program using...
Evidence-based practice

Evidence-based practice is an approach to practice that utilizes the best available evidence from research studies to make informed decisions about patient care. This approach is important because it helps clinicians provide the most effective and cost-efficient care. There are several steps involved in implementing evidence-based practice:

1. Develop a system to track clinical questions in your setting.
2. Work evidence-based practice into scheduled activities (e.g., rounds, journal club, in-service).
3. Incorporate evidence-based practice into staff education and student/residency programs.
4. Establish an evidence-based practice work group.
5. Make use of online evidence-based practice resources.
6. Establish evidence-based practice competencies for your setting and include assessment in staff performance evaluation.
7. Discuss evidence with patients as a component of patient education.
8. Develop an evidence-based practice folder with CATs.

Table 80.9: Evidence-Based Practice Implementation Strategies

- Develop a system to track clinical questions in your setting.
- Work evidence-based practice into scheduled activities (e.g., rounds, journal club, in-service).
- Incorporate evidence-based practice into staff education and student/residency programs.
- Establish an evidence-based practice work group.
- Make use of online evidence-based practice resources.
- Establish evidence-based practice competencies for your setting and include assessment in staff performance evaluation.
- Discuss evidence with patients as a component of patient education.
- Develop an evidence-based practice folder with CATs.

Compiled from suggestions made by clinical educators at the Boston University Faculty Summer Institutes 2001 to 2003.

The online resources. Working to become an evidence-based practitioner can seem overwhelming for an individual, but working in a group where individuals contribute their expertise eases the burden. Organizing an evidence-based practice journal club is an excellent strategy for gaining knowledge and skill while working to create an evidence-based practice culture. Table 80-9 provides a summary of evidence-based practice implementation strategies.

A randomized controlled trial studying the effects of journal club participation on medical residents’ evidence-based practice knowledge and skills showed the following significant changes: improved ability to critique the methods sections of articles; increased ability to read and incorporate medical literature in practice; and greater skepticism of results and conclusions (50). There are several strategies for creating effective journal clubs. It is important to have identified roles, particularly with respect to the facilitator and specific presenter. Group members can begin by assessing their evidence-based practice strengths and weaknesses. Then, the group can work together to share areas of expertise and capitalize on the strengths of group members (51).

Evidence-based practice work groups can develop systematic approaches to incorporate an evidence-based approach into scheduled activities. The group can identify a common place for recording clinical questions and select one question for critical appraisal and present the results during an in-service. A discussion of research evidence and patient preferences can also be incorporated into patient education activities.

There are also strategies at an institutional level to support creation of an evidence-based practice culture. The following steps are recommended: (a) begin by targeting a problem that is deemed important due to either a high prevalence or high cost; (b) synthesize information about optimal practice; (c) synthesize information about current practice; (d) identify discrepancies; (e) outline a practice improvement plan; (f) assess the size information about current practice; (d) identify discrepancies; (e) synthesize data; (f) assess the size data; (g) decide to implement or not and continue improvement efforts (52).

Conclusions

Evidence-based medicine, a term that was coined in the 1990s by a group of clinicians and epidemiologists at McMaster University in Ontario, Canada, has evolved into an international movement across all health care disciplines, including the disciplines that comprise the rehabilitation field. It is likely that a primary barrier to evidence-based practice—limited time—will be mitigated by advances in technology and the burgeoning evidence-based practice resources available via the Internet. As databases of evidence summaries continue to grow, one can imagine a future where clinicians routinely use wireless, hand-held devices for point-of-care access to relevant information that has been critiqued and summarized (53,54). Another barrier, limited knowledge of clinical epidemiology, will also have less of an impact as these methods are taught to future rehabilitation professionals.

There are several efforts that can support the continued growth of evidence-based practice in rehabilitation. The field would benefit from agreement on criteria for critiquing, evaluating, and grading the quality of rehabilitation research. Rehabilitation researchers should attempt to use evidence-based practice measures of clinical importance—NNT, effect size, likelihood ratios—to summarize data. And, finally, there is a need for better training in evidence-based practice, using relevant patient cases and tools to assess evidence-based practice knowledge and skills with examples from rehabilitation.

The quote by Goethe at the start of this chapter is an apt reflection of the challenge facing rehabilitation professionals. It is important to know about the latest advances in our field; however, knowing is not sufficient—knowledge of research must be applied. It is important to be willing to critically examine the care we provide to include the latest advances, but being willing is not enough—where evidence exists, we must take action to change practice.

Acknowledgments

This chapter includes many insights gained during the Faculty Summer Institute: Integrating an Evidence-Based Approach into Rehabilitation Professional Education. The Institute, sponsored by the Center for Measuring Rehabilitation Outcomes at Boston University and supported by the National Institute of Disability and Rehabilitation Research, was held for three consecutive years (2001 to 2003). The following individuals participated in the Faculty Institute and we acknowledge their contributions: Wendy J. Coster, PhD, OTR, Nancy Baker, ScD, OTR, Stephen M. Haley, PhD, PT, Julie Keyser, PhD, PT, Mary Law, PhD, OT(C), Robert Meenan, MD, MPH, Ken Ottenbacher, PhD, OTR, Hilary Siebens, MD, Patti Solomon PhD, PT, and Linda Tickle-Degnen, PhD, OTR.
REFERENCES

1. Flexner A. Medical Education in the United States and Canada. New York: Carnegie Foundation for the Advancement of Teaching; 1910.


Increasing evidence suggests that neural repair can support some degree of behavioral recovery in many forms of neurological disease. This chapter reviews these concepts, with an emphasis on stroke and on motor deficits. Stroke remains a leading cause of adult disability and illustrates many of the key points. Some degree of spontaneous behavioral recovery is usually seen in the weeks following stroke onset. Animal studies have provided insight into underlying molecular and physiological events. Brain mapping studies in human patients have provided observations at the systems level that often parallel findings in animals. The best outcomes are associated with the greatest return toward the normal state of brain functional organization. Reorganization of surviving central nervous system (CNS) elements supports behavioral recovery, with several cardinal patterns of brain reorganization apparent as contributory. A number of therapies are in development to improve patient outcomes by promoting repair after stroke, and likely many will be useful in other forms of CNS injury such as multiple sclerosis (MS), spinal cord injury (SCI), and traumatic brain injury (TBI). The optimal methods to prescribe such therapies requires further study, with anatomical, functional, and behavioral measures of brain state, as well as measures of injury, likely proving useful toward this goal.

**Insights from Animal Studies**

Studies in animal models of stroke have provided insight into the events underlying spontaneous behavioral recovery after stroke. These studies have particular value because animal studies are able to provide insight at the molecular and cellular levels.

A unilateral infarct is associated with a number of growth- and repair-related molecular events. In many cases, unilateral injury incites a bilateral response (Table 81-1). These have been reviewed in detail elsewhere (13–19). These include structural changes in axons, dendrites, and synapses; increased activation and migration of endogenous neural stems; plus changes in extracellular matrix, glia, and angiogenesis. The brain becomes excitable, for example, showing increased N-methyl d-aspartate (NMDA) receptor binding and gamma-aminobutyric acid (GABA) receptor down-regulation (20,21). Cell cycle proteins and growth factors increase to levels far above the normal level for the adult mammalian brain (22–25). Many of these events are most pronounced in the peri-infarct area.

Insights have also come from studies of cortical representational maps. Nudo et al. (26), in a landmark study, described changes in the hand motor map after stroke in sub-human primates. These authors found that hand motor training after stroke was associated with expansion of the motor cortex hand representation (Fig. 81-1). This reorganization was later found to be related to local synaptogenesis (27). Similar findings have been described in primary sensory cortex (28). Other influences on cortical representational maps, such as demographics (29–31), injury (32–35), medications (36), experience (17), and environment (26,37–40), have also been described.

Magnetic resonance imaging (MRI) studies in animals have also provided insights into the basis of behavioral recovery from stroke. These are of particular value: they examine changes in representational maps noninvasively, and furthermore the metrics extracted can be very similar to the metrics reported in human studies; such direct comparisons might be very useful for effective translation of therapeutic advances. For example, changes in connectivity between sensorimotor cortex and deep gray matter structures in rats subjected to middle cerebral artery (MCA) stroke (41) are similar to findings in human stroke patients (42). Also, serial functional MRI (fMRI) studies examining rat upper limb have described a shift in laterality of activation after stroke such that early...
FIGURE 81-1. Serial findings using intracortical microstimulation mapping in subhuman primates (A) before experimental stroke and (B) after stroke plus a program of rehabilitative training focused on hand. The rehab training prevented loss of hand territory adjacent to the infarct, when compared to animals not receiving this training. In some instances, the hand representations expanded into regions formerly occupied by representations of the elbow and shoulder. This study demonstrated that rehabilitative training can shape subsequent reorganization in the adjacent intact cortex, and suggested that the undamaged motor cortex may play an important role in motor recovery. The long thin arrows point to regions of undamaged cortex in which digit representations have invaded regions formerly occupied by representations of the elbow and shoulder.

TABLE 81.1 Molecular and Cellular Repair-Related Changes After Stroke

<table>
<thead>
<tr>
<th>Inflammation</th>
<th>Increased growth factors and growth-associated proteins</th>
</tr>
</thead>
<tbody>
<tr>
<td>Down-regulation of GABA receptors</td>
<td>Increased cell cycle proteins</td>
</tr>
<tr>
<td>Hyperexcitability, with facilitation of long term potentiation</td>
<td>Increased NMDA receptor binding</td>
</tr>
<tr>
<td>Angiogenesis</td>
<td>Increased synaptogenesis</td>
</tr>
<tr>
<td>Increased dendrite branching and spine density</td>
<td>Increased neuronal sprouting</td>
</tr>
<tr>
<td>Increased cortical thickness</td>
<td>Increased extracellular matrix remodeling</td>
</tr>
</tbody>
</table>

Growth- and repair-related molecular and cellular changes after a unilateral infarct in animal models. In most cases, these changes arise bilaterally. Some of these changes increase with therapeutic intervention, in parallel with enhanced behavioral gains, further supporting their importance as biological targets for promoting repair after stroke (18,407).

after stroke, brain activation during affected paw stimulation is mainly in the contralateral sensorimotor cortex. Later after stroke, activity shifts toward the prestroke (normal) pattern in the ipsilateral sensorimotor cortex (43,44). These findings are closely aligned with findings from serial fMRI exams in human subjects with stroke (45–50).

**Systems Insights from Brain Mapping Studies in Human Subjects**

The molecular measurements from animal studies are difficult to obtain in humans, though occasionally specific questions can be probed with molecular imaging methods such as positron emission tomography (PET). Information regarding the brain events underlying stroke recovery in humans can be obtained, however, from studies of brain systems via brain mapping. A number of functional neuroimaging methods have been used to this extent.

Overall, after stroke in humans, tissue function is reduced within the injured (or for deep strokes, overlying/corresponding) primary neocortex (51–53). Behaviors that rely on the dysfunctional primary cortex suffer. The best spontaneous return of behavior is associated with resolution of such reductions of cortical function. A number of compensatory brain responses also contribute to spontaneous behavioral recovery. These include increased activation in secondary areas that are normally connected to the injured zones through a distributed network, a shift in interhemispheric lateralization toward the contralateral hemisphere, and shifts in representational maps surrounding the infarcted zone. In many cases, the larger the injury or greater the deficits, the more these compensatory events are seen. These compensatory responses are tricks of the desperate, but in patients with injury-related deficits several lines of evidence suggest that often they are better present than absent. However, the best behavioral outcomes after stroke are associated with the greatest return of brain function toward the normal state of organization (49,53–57). These findings take on particular importance when it is realized that the same events that support
spontaneous recovery are likely those to be measured to assist studies aiming to therapeutically improve recovery (58).

**Stroke Injury Reduces Cortical Activity Locally**

When a stroke injures primary cortex and/or its underlying white matter (and sometimes corresponding basal ganglia), cortical function is reduced initially, increasing over time (45–47,49,56,59). Thus, several lines of evidence suggest that final behavioral outcome at the end of the stroke recovery period is related to the degree of activity in primary cortex underlying the behavior under study (60). For example, a study of patients with stroke who had reached a plateau in motor recovery found that the volume of primary sensorimotor cortex activation in the ipsilesional hemisphere during affected hand movement was related to the level of behavioral recovery (53). Similar findings have also been described in language (49,55,61–63) and right hemisphere attentional networks (56). Consistent with this, transcranial magnetic stimulation (TMS) studies of motor cortex and its efferents after stroke have generally found that cortical motor maps are smaller and corticospinal tract physiological integrity is reduced in proportion to severity of clinical deficits (64).

Viable cortical regions that surround an infarct might have special importance to repair and recovery. In animals, this peri-infarct zone often shows the greatest levels of growth-related molecular changes after stroke (14,18,26,38,65–76). Furthermore, amplification of peri-infarct repair-related events in animals by a therapeutic intervention has been associated with improved behavioral outcome (69,71,77,78). In humans, the volume of threatened but surviving peri-infarct tissue is directly related to final clinical outcome (79,80). Brain mapping studies in humans with chronic stroke have often noted peri-infarct activity (55,81–92), though the T2*-weighted MRI signal used to measure brain activation with fMRI is itself altered in the peri-infarct zone, perhaps due to glial scarring (87). This observation complicates application of fMRI to the study of peri-infarct regions in patients with cortical stroke.

**Stroke Injury Incites Increased Cortical Activity at Distant Sites**

Three main forms of reorganization at sites distant from stroke have been described: (a) increased activity in brain regions distant from, but connected to, the stroke zone; (b) increased activity in the contralesional hemisphere, which reduces the extent to which interhemispheric balance is lateralized; and (c) somatotopic shifts within intact cortical regions. Each has been associated with behavioral improvement after stroke.

The first of these is increased activation within cortical areas that, prior to stroke, comprise a distributed network (49,63,81,83,93–102). This has been described in many studies, indeed since the first poststroke functional imaging study by Brion et al. (103). Reaction across a distributed network has been reported across many neurological domains including motor, language, attention, and visual functions.

A second form of reaction to stroke is reduced laterality of brain activity (61,81,93–96). Reduced laterality is a cardinal pattern of brain response to injury, having also been described in other neurological contexts such as epilepsy (104), TBI (105), primary progressive aphasia (106), and MS (107). Note that the degree of brain insult needed to incite reduced laterality might be much lower than previously appreciated (108,109).

The exact function served by increased contralesional activity after stroke remains to be clarified. To some extent, this can be seen as simply a contralesional example of increased brain activation at sites distant from stroke. Another interpretation is that this is a passive event, reflecting a reduced interhemispheric inhibition resulting from the stroke (110–116). Changes in interhemispheric inhibition are the focus of much discussion (114), and possibly of therapeutic value (117–119), though the significance of such findings might differ across patient subgroups (120). Another hypothesis is that the contralesional hemisphere assumes functions that were previously based in the ipsilesional hemisphere. However, the data do not support this, at least not directly and completely. For example, stimulating contralesional motor cortex with TMS does not result in movement of the affected hand (121), and indeed might indicate greater pathology if such a response is seen (122,123). Contralesional hemisphere functions might be more substantial when stroke is prenatal (124).

A third compensatory response to focal injury is reorganization of somatotopic maps. Such maps are normally present in multiple primary cortices including motor, sensory, auditory, and visual cortex, and such somatotopic organization also exists in white matter, basal ganglia, secondary neocortex, and the hemisphere ipsilateral to movement (125–130). Nonhuman primate studies have provided rich characterization of cortical map changes after ischemic injury, providing insight into the effects of rehabilitation as well as infarct size, and also describing novel axonal projections that can arise in this context (13,26,37,38,131,132). This process has been less studied in humans after a stroke, where most, but not all (83,133,134), investigations have focused on the motor system. In the motor system, the large scale features of motor cortex somatotopy that are normally present (135) are preserved though occasional exceptions have been described (136). A shift of the motor cortex hand representation after stroke in the dorsal (137), ventral (53,94,98,136,138), or posterior (82,136,139–142) direction has been reported. Sites assuming such a shift in brain activation can demonstrate increased cortical thickness (134).

The motor cortex map relationships between face and hand may be among the most plastic in the brain, a suggestion that is supported by reports of invasion of hand motor representation into face motor area (53,98,136,138,143–145) and face into hand (146,147), in a variety of settings, first described after stroke by Weiller et al. (94). This process might reflect survival of distinct subsets of corticospinal tract fibers (128,148) that have key axonal connections between hand and face areas (149), and so could theoretically represent an approach to identify a biologically distinct subgroup of stroke patients in whom a particular therapy (one that encourages a ventral hand map shift) might be most likely to succeed (Fig. 81-2).

These forms of poststroke increased activation have certain principles in common. These changes are not epiphenomena,
as suggested by virtual lesion (150–154) and other approaches (155–157), but rather directly contribute to spontaneous behavioral recovery. Increased activity in distant areas (46,98,158) and reduced laterality (45,48,49,54,136,159,160) are time-dependent, increasing in the early weeks after stroke and generally declining thereafter. This decline is greater among subjects with better behavioral outcome, as the degree of persistently increased activity in both instances is generally highest in those with the poorest behavioral outcome (54,136,161). However, note that persistent, increased reorganization is not always helpful, as persistence of poststroke plasticity can be associated with induction of epilepsy (162) or chronic pain (163).

**Diaschisis**

The changes arising in areas distant from injury might in part be related to diaschisis or its resolution. Diaschisis refers to reduced activity, typically measured in terms of blood flow and/or metabolism, in uninjured brain areas that have rich connections with injured brain areas (164–166). Some data suggest that behavioral recovery is facilitated by the resolution of diaschisis, that is, restitution of brain activity in these uninjured areas that are distant from, but connected to, the site of infarct (50,167,168). The optimal method for defining, and thus for measuring, diaschisis remains unclear.

**Time Window Influences Application of Restorative Therapies after Stroke**

The time window that defines a restorative therapy likely varies according to the therapy and biological target. The biological targets for restorative agents vary over time, with brain levels spontaneously rising then falling dramatically in the weeks following a stroke. Some models have broken down poststroke brain changes into three epochs (160,169,170) that overlap at least in part: (a) acute injury, taking place in the initial hours after a stroke; (b) repair, starting in the first days after stroke onset and lasting several weeks. This is a golden period for initiating exogenous restorative therapies, as the greatest degree of spontaneous behavioral recovery is seen, and endogenous repair-related events (see Table 81-1) reach peak levels (Fig. 81-3). Note that exogenous therapies, from pharmacotherapies to behavioral interventions, must be examined in a two-tailed manner, as each can be deleterious.
**TABLE 81.2** Variables Influencing Stroke Recovery

<table>
<thead>
<tr>
<th>Variables</th>
<th>Recovery Factors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Infarct size</td>
<td>Infarct location</td>
</tr>
<tr>
<td>Prestroke medical co-morbidities</td>
<td>Prestroke disability</td>
</tr>
<tr>
<td>Prestroke experience and education</td>
<td>Age</td>
</tr>
<tr>
<td>Severity of initial stroke deficits</td>
<td>Breadth of stroke deficits</td>
</tr>
<tr>
<td>Acute stroke interventions</td>
<td>Medications during stroke recovery period</td>
</tr>
<tr>
<td>Amount of poststroke therapy</td>
<td>Type(s) of poststroke therapy</td>
</tr>
<tr>
<td>Medical complications poststroke</td>
<td>Socioeconomic status</td>
</tr>
<tr>
<td>Depression</td>
<td>Caregiver status</td>
</tr>
<tr>
<td>Genetics</td>
<td></td>
</tr>
</tbody>
</table>

as well as potentially helpful, as discussed below: (c) a plateau, starting weeks-months after stroke, and representing a stable but still modifiable chronic phase. This third phase might have two temporal components (98): one representing the beginning of the chronic phase, with associated loss of treatment windows from phase 2, and the other arising many months-years after stroke onset, associated with late changes and complications. Late changes and complications of a "late plateau" epoch include new dystonias, cognitive/affective changes, and spasticity/contracture.

**Other Factors Influencing Recovery**

A number of general factors can influence outcome after stroke (47,85,148,166,169,171–190) (Table 81-2). The quality and intensity of behavioral experience during the weeks-months following a stroke, and the environment in which this experience takes place, are well established as having a major influence on behavioral outcome after stroke (39,191–200). The nature of the physical and social environment also influences the brain, in health (201) and after stroke (202). Genetic factors might also be important (203,204). Note that a number of factors can also reduce behavioral recovery after stroke. One example is drugs such as neuroleptics or antiepileptics (205–208). Other drugs, such as benzodiazepines, can induce an anamnestic state: a patient can have poststroke deficits, recover spontaneously, then have the deficits reemerge years later upon administration of certain drugs (209–212). The brain events that support behavioral recovery appear to have a vulnerability (213) that lasts a lifetime.

**Restorative Therapies**

Restorative therapies aim to promote behavioral recovery after stroke. Numerous therapies, devices, and interventions for modifying function after stroke have been evaluated and are considered below.

**Small Molecules**

Many chemical events in the poststroke brain are relevant to repair. Not surprisingly, a number of small molecules have been examined as potential therapies in this context. The mechanism for some of these small molecules revolves around direct manipulation of a specific neurotransmitter axis, such as increasing tone at serotonergic (214,215) or monoaminergic (216–218) receptors. Results with the latter example, such as with amphetamine and related drugs (219–225), have been mixed. Other small molecule approaches have been examined (226–233), including immune-based approaches (234).

**Growth Factors**

Growth factors play an important role in the development and spontaneous brain remodeling and thus, might have great promise as therapies to improve function after stroke. A wide range of potential mechanisms exist, including promoting activity-dependent synaptic competition, long-term potentiation, facilitating key protein synthesis and synaptic transmission, and more. Several positive preclinical studies have been published (72,235,236), including those employing a hematopoietic growth factor (237–239). Kolb et al. (240) found that sequential administration of epidermal growth factor and erythropoietin to rats with experimental infarct reduced deficits, in some cases when treatment was initiated 7 days after stroke onset.

In most cases, these growth factors have been initiated 24 hours or longer after stroke onset in a rodent model of experimental stroke, and found to improve long-term behavioral outcome. This long therapeutic time window, the large number of examples of successful preclinical studies, the use of compounds that are endogenous to humans, and reliance on compounds that in many cases have a long history of safe application in humans are all factors that predict future successful application of growth factors to restoring function after stroke in humans.

For large molecules such as growth factors, access to the brain might prove to be a significant challenge. One potential solution to this involves the use of a “trojan horse” approach (241).

**Cell-Based Therapies**

A number of forms of exogenous cells, including induced pluripotent (242,243) and embryonic stem cells (244,245), have been examined in preclinical stroke models (246,247). Mobilization of endothelial progenitor cells (248) might also help repair after stroke (249). Limited data are available on treatment of human subjects with stroke using exogenous cells (246,247). The safety and feasibility of intracerebral transplantation of cultured neuronal cells in humans were established in two trials of Layton BioScience-neurons (250,251), with at best modest results. A small trial in human patients with subacute stroke found intravenous mesenchymal stromal cell (MSC, also referred to as marrow stromal cell) infusion to be safe and possibly effective for reducing disability (252). MSC can be prepared autologously, eliminating the need
for immunosuppression, and might have a therapeutic time window of up to 1 month poststroke (253). MSC might also prove useful as a conduit for therapies targeting specific genes (254–257). A number of issues require addressing whatever the cell used, such as establishing biological identity, activity, purity, and stability over time.

**Electromagnetic Stimulation**

The brain is an electrical organ. Electromagnetic stimulation has thus been used to modulate a number of brain functions; indeed a brain-wide form of electromagnetic stimulation, electroconvulsive therapy, remains the gold standard for the treatment of major depression (258). Several forms of electromagnetic intervention are under evaluation to improve function after stroke, though in most cases studies are at an early phase. One such form of stimulation is repetitive TMS, which, depending on the number of stimuli per second, can have inhibitory or excitatory effects on cortical activity (259). As such, goals can include increased activity in ipsilesional cortical regions that are underactive (260–262), or decreased activity in contralateral regions that are overactive and a source of potentially harmful inhibition (118,119). Transcranial direct current stimulation has also shown promise in initial studies, and might be less focal in its brain effects (263). Epidural motor cortex stimulation might also improve motor function after stroke (264), though a phase III study in 164 patients with chronic stroke found that the effects on motor status of epidural motor cortex stimulation plus rehabilitation therapy did not significantly differ from rehabilitation therapy alone (265).

**Device-Based Therapies**

Numerous devices are under study to improve function after stroke, some employed external to the body and others surgically internalized. Some devices aim to substitute for the injured CNS, such as via a brain-computer interface (266,267). For others, such as via a laser-based device (268), functional electrical stimulation (269), neuroprosthetic, and robotic devices (270), the goal is to improve function in surviving CNS elements. Electrical stimulation and neuroprostheses are further discussed in Chapters 71 and 48, respectively. A number of different robotic devices have been examined (271–275). These devices have been largely focused on motor function, where mild-moderate gains have been documented. Robotic therapy offers potential advantages in that robotic devices can be active without fatigue for very long time periods, perform in a consistent and precise manner, can be programmed, have the capacity to measure a range of behaviors, and are enabled for telerhabilitation. Robotic therapy is further discussed in Chapter 83. Combination approaches employing a robotic therapy alongside a second restorative therapy such as a drug or neuromuscular stimulation have been less studied but might prove fruitful.

**Task-Oriented and Repetitive Training-Based Interventions**

Interventions based on task repetition are important to the repair process for at least two reasons: first, behavioral experience influences effectiveness of other interventions such as pharmacological therapy. Second, task-oriented and repetitive training-based interventions might have significant value as the main therapy of interest (276–278). A number of forms of therapy have been studied, with uncertain results at this time as to which is best. Cognitive-based strategies, such as those based on motor imagery or action observation, have shown promise (197,279–287). Interventions in the motor system have included bilateral training (288), use of a gravity-reduced environment (289), EMG trigger coupling (290), incorporating passive movement (291), or manipulating movement speed (292). The structure (293) and intensity (192,196) of the training program can influence extent of gains.

One such therapy of note is constraint-induced movement therapy, which is based on overcoming learned nonuse of the affected hand after stroke (294). Constraint-induced movement therapy has the nonaffected hand restrained while the affected hand undergoes an intense course of therapy. A recent phase III trial (295) found that 2 weeks of this therapy produced gains (295) that remained significant 2 years (296) after the end of the intervention (Fig. 81-4). Modified forms of this therapy might increase the fraction of chronic stroke patients who are eligible (290). In addition, studies are examining this approach in other neurological domains such as language (297). An important question is the percentage of patients who can derive gain from this and other task-oriented and repetitive training-based interventions.

![FIGURE 81-4. Therapy can produce significant, enduring gains among patients in the chronic phase of stroke, from the EXCITE trial, a phase III clinical stroke trial (6). Wolf et al. randomized 222 patients with moderate deficits due to a first ischemic or hemorrhagic stroke in the prior 3 to 9 months to either usual care or to 2 weeks of constraint induced therapy. These authors found significant gains in several measures, such as the amount of arm use presented in this figure, a full year following the end of therapy. Furthermore, follow-up exam 2 years after end of therapy also found significant retention of gains (7).](image-url)
Utility of Brain Mapping Data for Restorative Therapy Clinical Trials After Stroke

Clinical stroke trials often enroll patients based on behavioral entry criteria. However, behavioral performance can arise on the basis of many different neurobiological states, only some of which are likely to respond to restorative therapy. Several studies suggest that human brain mapping can provide additional insights useful for optimal implementation of restorative therapies. By providing measures of injury and brain functional state, data from fMRI, PET, TMS, magnetoencephalography, electroencephalography, and other methods might be of value to clinical trials.

There are several functions that functional neuroimaging measures might play. One might be to guide particular details of a restorative intervention. If a therapy achieves its effect by targeting a specific functional brain region, such as the hand motor area, anatomical localization might be insufficient because localization of a specific function in the brain often has an inconsistent relationship with brain anatomy (304–309). Serial measurement of brain function might also guide therapy dose or duration (310,311). This approach has been demonstrated in clinical (312,313) and research (264,314) settings. A second role for functional neuroimaging measures might be to guide dose of therapy, for example, via measurement of fMRI laterality measurements (310) or motor-evoked potential by TMS (311).

An additional potential value of functional neuroimaging methods is to serve as entry criteria. Having a measure of CNS functional status and injury might assist usual behavioral and demographic measures (191,196) to identify patients most likely to respond to a given intervention. Functional measures of brain state have been incorporated into early phase restorative stroke trials and found to predict response to therapy (310,311,315,316). A recent study by Stinear et al. (317) achieved good success at predicting arm motor gains from therapy using measures of demographics, CNS function, and injury. This model, and so many others, requires validation and further investigation, for example, of clinometric patterns. However, likely some measures beyond those obtainable at the bedside will likely prove useful for optimal prescription of restorative therapies after stroke.

Functional neuroimaging measures might also one day serve as surrogate endpoints in the evaluation of restorative therapies. Surrogate markers can be of considerable value to early phase trials. A proof of principle functional neuroimaging study in a small target population could be very useful when considering proceeding to a late phase trial (318). However, important concerns exist in the use of functional neuroimaging measures as a surrogate endpoint for restorative stroke therapies. Many more studies are needed to examine the performance of such measures in this context. A number of issues, such as the extent and specificity with which a surrogate marker captures treatment effect, require examination (319,320).

Summary

Studies in animals and humans are providing new insights into the biological basis of spontaneous recovery after stroke. On the backbone of this knowledge, a number of restorative therapies are entering human trials to promote improved outcomes. Measures of CNS injury and function might be useful toward these efforts. Further study is needed on a number of questions to optimally prescribe restorative therapies and reduce disability after stroke.

MULTIPLE SCLEROSIS

MS is also a common source of disability, and motor deficits are also common in MS (Chapter 41). These deficits are acquired over a much longer time in MS as compared to stroke. For example, in one study of a broad range of subjects with MS, the median time to reach irreparable limited walking ability for more than 500 m without aid or rest was 8 years, to walk with unilateral support for not more than 100 m without rest was 20 years, and to walk for not more than 10 m without rest while leaning against a wall or support was 30 years (321). Upper extremity motor deficits, such as those related to ataxia and paresis, are also a common source of disability in MS (322).

Brain plasticity is an important determinant in MS over two time scales. First, steady destruction of myelin and axons over the years results in disability. During this period, reorganization of brain function can reduce the impact that such injury has on behavioral decline. Second, over weeks-months, approximately 85% of patients with MS have a relapsing, remitting course (323) in which a relapse peaks over 1 to 2 months and then improves over a similar time period. The resolution of these MS flares has been attributed to a number of brain events, such as neurological reserve and resolution of inflammatory insult (323), and a number of studies suggest that brain plasticity is also important (58). Note too that there are numerous asymptomatic brain lesions for each symptomatic one in most patients with MS, a fact that might further support the importance of brain plasticity in the maintenance of behavioral status in this condition.

Brain plasticity, thus, is likely important to motor status in MS, by minimizing the debilitating effects of MS injury accrual over the long-term period and by promoting recovery from silent or symptomatic MS flares over the short-term period. A number of studies have provided insights into the
PART VII | SCIENTIFIC ADVANCES

brain events important in this regard, with substantial overlap as compared to findings in patients with stroke. This information gains importance in the current discussion because events important to maximizing behavioral status in the natural course of the disease are likely to be many of the same measures whose measurement can guide optimization of therapy-derived recovery.

fMRI studies of brain plasticity in MS have found that, early in the course of the disease, brain activation is larger and more widespread as compared to healthy controls. Later in the disease, laterality of activation is reduced (i.e., activation is more bilateral) (324–326), akin to stroke patients who have larger infarcts or greater deficits (54,136). Bilateral sensorimotor cortical regions are activated to a greater extent in the setting of MS-related white matter injury (327,328). This increased degree of bilateral organization persists to the greatest extent in subjects with persistent deficits after an acute MS relapse and returns to a normal, lateralized (i.e., contralateral-predominant) form of organization in subjects with the least degree of persistent disability (329,330). The pattern of brain activation during performance of a simple motor task in subjects recovered from stroke has been considered similar to the pattern seen in healthy subjects during performance of a complex task (81); a similar analogy has been made in subjects with MS (331).

A number of repair-based therapeutic strategies are also being explored in MS, though considerably less knowledge exists on this approach in MS as compared to in stroke. Many of these overlap with approaches described for stroke, above, such as stem cells and growth factors. Manipulation of myelin growth inhibitory molecules such as Nogo, as well as other myelin components, might also be an important avenue to promote repair in this demyelinating disease. A range of immune-based approaches, which constitute the centerpiece of acute MS therapies, might also have a role in promoting repair (332–337).

A handful of studies have examined the utility of brain mapping data for restorative therapies in patients with MS. Parry et al. (338) tested the effects of increased cholinergic tone on the pattern of fMRI activation during performance of the Stroop test, a cognitive task. At baseline, patients with MS and moderate disability had similar behavioral performance as compared to controls, but on fMRI showed increased left medial prefrontal, and decreased right frontal, activation. Treatment with the cholinesterase inhibitor rivastigmine normalized both of these fMRI abnormalities in patients, but had little effect on a small cohort of healthy control subjects. Mainero et al. (339) found that administration of 3,4-diaminopyridine to patients with MS was associated with reduced intracortical inhibition and increased intracortical facilitation. Though no behavioral correlates were noted in relation to these physiological changes, a favorable effect on behavior would be expected by analogy (189,263,340).

SPINAL CORD INJURY

Though SCI can be associated with a range of injury patterns, motor deficits are generally a prominent feature (Chapter 47). At the time of discharge from initial SCI, the most frequent neurologic category is incomplete tetraplegia (31.2%), followed by complete paraplegia (28.2%), complete tetraplegia (17.5%), and incomplete paraplegia (23.1%). Less than 1% of persons experience complete neurologic recovery by hospital discharge (341).

Persons with SCI generally show modest spontaneous sensory and motor improvement in the first 3 to 6 months following injury (342,343), although significant spontaneous improvement beyond the first year post SCI is uncommon (344). Motor deficits are thus common and persistent after SCI, and these impact a number of health, quality of life, and other issues in subjects with SCI (345–347). For example, by 10 years after SCI, 68% of persons with paraplegia, and 76% of those with tetraplegia, remain unemployed (348). Concerns over sexual and bowel/bladder function are also prevalent (349).

There has been limited study of the changes in the function of the human CNS after SCI. Some studies have found a broad decrease in brain activation (350–352), particularly in primary sensorimotor cortex, whereas others have found supranormal brain activation (353). The basis for these discrepancies remains unclear but could be due to differences in age or injury pattern of the population studied, years post SCI at time of study, or the nature of the task used to probe motor system function, some uncovering deficient processing and others emphasizing supranormal efforts to compensate (351,354). A commonly described feature is a change in somatotopic organization within primary sensorimotor cortex, with representation of supraspinal body regions expanding at the expense of infrasplinal body regions (102,355–358). Spontaneous changes in laterality, so prominent in studies of stroke or MS, as above, are generally not prominent after SCI, perhaps due in part to the fact that injury typically affects the CNS bilaterally or perhaps due in part to the fact that SCI spares brain commissural fibers whose integrity helps maintain normal hemispheric balance. As such, a laterality index is unlikely to be a useful variable in brain mapping studies of subjects with SCI. Widespread abnormalities in the limbic system have also been described (359,360), the full impact of which has yet to be appreciated. The temporal sequence of brain changes after CNS are just beginning to be understood (351). Also, the direct study of spinal cord function and plasticity at the level of the spinal cord is an emerging approach that might provide novel insights after SCI and after other forms of SCI (361,362).

A small number of studies have examined changes in CNS function in relation to therapy after SCI. Wolfe et al. (363) found that 4-diaminopyridine can improve central conduction time in subjects with SCI. Winchester et al. (364) studied body weight supported treadmill training in four patients with motor incomplete SCI. These authors compared fMRI during attempted unilateral foot and toe movement before versus after training. This therapy was associated with increased activation within several bilateral areas, including primary sensorimotor cortex and cerebellum, though to a variable extent. The authors observed that, although all participants demonstrated a change in the blood-oxygen level–dependent signal following training, only those patients...
who demonstrated a substantial increase in activation of the cerebellum demonstrated an improvement in their ability to walk over ground, suggesting that this measure in this brain region, at least when examined using this task during fMRI, might be useful as a biological marker of successful treatment effect. A small study suggests that gains from treadmill training correlate with improvements in motor system physiology (365). Similarly, a case report motor cortex hand map expansion in relation to gains from hand sensorimotor physiotherapy (366), though this result was not found in a larger study (367). These initial findings anticipate that many of the principles under discussion for use of brain mapping to improve physiotherapy effects in trials of recovery after stroke, such as use of physiology or brain map measures as entry criteria or as treatment biomarkers, might also pertain to SCI.

Another form of intervention that has been evaluated after SCI is motor imagery. Motor imagery normally activates many of the same brain regions as motor execution, and has been associated with improvements in motor performance (279,368). The effects of 1 week of motor imagery training to tongue and to foot were evaluated in ten subjects with chronic, complete tetraplegia/paraplegia plus ten healthy controls (369). The behavioral outcome measure was speed of performance of a complex sequence. Motor imagery training was associated with a significant improvement in this behavior in a nonparalyzed muscle group, the tongue, for both groups. In both the healthy controls and the subjects with SCI, serial fMRI scanning (before vs. after training) during attempted right foot movement was associated with increased fMRI activation in left putamen, an area associated with motor learning, despite foot movements being present in controls and absent in subjects with SCI. Training effects on brain plasticity can thus be measured independent of behavior effects, a finding that might be important for designing biological markers in trials targeting severely disabled patient populations. Note that this fMRI change was absent in a second healthy control group serially imaged without training. The main conclusion from this study is that motor imagery training improves brain function whether or not sensorimotor function is present in the trained limb. An additional conclusion is that motor imagery, by virtue of its favorable effects on brain motor system organization, might have value as an adjunct motor restorative therapy. Another key point from this study is that brain plasticity related to plegic limbs can be studied in subjects with chronic SCI. The above study examined, as a first step, the effects of a pure motor imagery intervention, but, as with subjects with stroke, motor imagery might have its greatest effect when directly combined with physiotherapy interventions.

**TRAUMATIC BRAIN INJURY**

TBI is a common, heterogeneous condition that is the source of much human suffering. Patients can present with wide-ranging impairments, including cognitive, visual, and motor. Despite a number of advances in neurorehabilitation, substantial morbidity and mortality remain (370). TBI shares some clinical features of other forms of CNS injury. For example, predictors of outcome after TBI, such as severity of initial deficits, are similar to those described in other forms of CNS injury, though social aspects may play a larger role in TBI (371).

Treatment of TBI has been reviewed elsewhere (372,373). In the more chronic phase, physiotherapy regimens targeting TBI may not be as developed as with other conditions (374). More intensive rehabilitation programs are associated with earlier function gains (375,376). There have been limited advances in therapies that promote brain repair after TBI (377), with relatively few studies examining restorative medications in the setting of human TBI (375). Small studies suggest that some of the compounds potentially useful for promoting brain repair after stroke might have utility after TBI (378–381), possibly including brain stimulation (382), cognitive rehabilitation (383), and progesterone (384). Spontaneous increases in growth factor levels that have been observed after TBI might indicate a therapeutic role for this family of restorative therapies (22,385). Cellular therapies might also have promise in reducing disability after TBI (386,387).

Brain imaging is providing new insights into the pathophysiology of TBI effects as well as of the brain repair that arises secondarily (388). Diffusion tensor imaging (DTI) measures injury, particularly in the white matter. DTI sometimes discloses injury that was completely inapparent with standard anatomical MRI scanning (389,390), or discloses evidence of repair (391). fMRI provides information about function that, as with stroke (392), sometimes vastly exceeds what can be determined with behavioral exam (393) or anatomical imaging (394). Insights with fMRI (395–398) and other functional neuroimaging methods (399–402) extend across brain functional systems as well as injury severities (401,403). Functional imaging studies of TBI have been criticized for lack of protocol standardization (404), a concern common to study of many neurological conditions (405).

An improved understanding of the pathophysiology of TBI will be useful to better define and target repair-related processes (404). In parallel, a need exists for advances in functional neuroimaging. Such advances should permit continued growth (406) in the ability to use measures of brain plasticity in humans to improve application of restorative therapies to reduce disability after TBI.

**SUMMARY**

The CNS possesses innate responses to injury. Increased attention is being paid to these. The result is a better understanding of CNS repair in humans. A wide range of therapies are under investigation with the aim being to promote CNS repair. Together, these advances anticipate regular use of restorative therapies to promote CNS repair and reduce disability in patients with stroke, MS, SCI, TBI, and other neurological conditions.
ACKNOWLEDGMENTS

Dr. Cramer is supported by grants from NIH (EB005047 and HD33352), Stem Cell Therapeutics, Northstar Neuroscience, Inc., and GlaxoSmithKline. He is a consultant to GlaxoSmithKline, Merck & Co., Inc., Biotrex, CytRx Corporation, Photothera, Johnson & Johnson, and Allergan Inc.

REFERENCES


195. Legg L, Langhorne P. Rehabilitation therapy services for stroke.
284. Krakauer JW. Motor learning: its relevance to stroke recovery and
285. Page SJ, Levine P, Leonard A. Mental practice in chronic stroke:
286. Johnson-Frey SH. Stimulation through simulation? Motor imagery and
287. Mulder T. Motor imagery and action observation: cognitive tools for
293. Sullivan K, Knowlton B, Dobkin B. Step training with body weight
299. Meinzer M, Elbert T, Djundja D, et al. Extending the constraint-
300. Nakayama H, Jorgensen HS, Raaschou HO, et al. Compensation in
321. Stephen RM, Gillies RJ. Promise and progress for functional and


INTRODUCTION
An innovative interdisciplinary scientific field, termed regenerative medicine, focuses on new approaches to repairing and replacing cells, tissues, and organs and may involve the use of gene therapy (1–6). Derived from the fields of biomedical engineering, developmental biology, nanotechnology, physiology, molecular and cellular biology, and surgery, regenerative medicine holds new promise for previously incurable conditions. Potentially, any disease or condition (acquired or genetic) that results in damaged, failing, or malfunctioning tissue may be amenable to regenerative medicine technologies. These technologies fall into two general categories (Fig. 82-1): (a) a focus on growing tissues and organs in vitro (i.e., tissue engineering) and subsequently implanting them into the patient; and (b) a focus to fully leverage the host’s regenerative in vivo capacity using cellular and/or gene therapies. This review concentrates on the latter approach in regenerative medicine.

In 2004, The Joint Commission on the Accreditation of Healthcare Organizations (JCAHO) warned that failure to address the severe shortage of organs could have ominous consequences for the American population. There is about one death every 30 seconds from organ failure, and complications and rejection of transplanted organs pose significant problems. The United Network for Organ Sharing (UNOS), which maintains a database on transplantation in the United States, reported that approximately 29,000 transplants were performed annually in 2006 and 2007. However, by mid-2008 the number of Americans on waiting lists for donated organs grew to nearly 100,000. These facts point to the importance of developing new therapeutic approaches for this huge unmet medical need. In this chapter, we describe just a few of the new approaches that the emerging field of regenerative medicine uses to address this need for organ and tissue replacement.

Cell Therapy: Regenerating Damaged Tissues In Vivo
Diseased tissue may be regenerated in vivo by transplantation of healthy cells that are able to replicate extensively. Progenitor and stem cells have this intrinsic ability, and are used in regenerative medicine to enhance or restore damaged tissue. According to data from the Centers for Disease Control, as many as 1 million Americans will die every year from disease that, in the future, may be treatable with tissues derived from stem cells (7). Diseases that might benefit from stem cell–based therapies include diabetes, heart disease, cerebrovascular disease, liver and renal failure, spinal cord injuries, and Parkinson’s disease. In the following section, we describe the classes of cells available for regenerative techniques and describe research advances using these cells in selected clinical conditions commonly encountered by rehabilitation clinicians.

Adult Stem Cells and Tissue Progenitor Cells
Adult stem cells are defined as any stem cell population that corresponds to a point in development subsequent to the inner cell mass (ICM), or possibly the slightly later epiblast, of embryos at the blastocyst stage, prior to gastrulation. In other words, adult stem cells are usually defined as stem cells found at any point in development beyond the pluripotent cells of early embryos. Adult stem cells are relatively abundant during fetal development and persist in tissues throughout adult life. Much of our current understanding of stem and progenitor cell biology derives from studies of a particular class of adult stem cells, namely those of the hematopoietic (blood forming) system. In fact, the therapeutic use of adult stem cells dates back to the first bone marrow transplant in 1956 (8). Early suggestions supporting the existence of cells that are capable of reconstituting the blood system came from experience with persons exposed to lethal doses of radiation during World War II. In the early 1960s, James Till and Ernest McCulloch in Toronto, Canada, found evidence for a specific subpopulation of cells in the bone marrow that could restore hematopoiesis after transplantation into irradiated mice. In the spleens of recipient animals, they observed large colonies and showed that these arose as clones from individual stem cells in the donor marrow that both could replicate to generate more stem cells and could give rise to multiple types of mature, differentiated blood cells (9,10). These two characteristics, termed self-renewal and multipotency, remain generally accepted as the defining features of stem cells. In the most complete demonstration of these properties, it has been shown that one hematopoietic stem cell has the capacity to completely repopulate both the lymphoid and myeloid blood cells of a lethally irradiated mouse and that single stem cells from such rescued animals can in turn reconstitute additional irradiated hosts over multiple sequential passages (11,12).
Human hematopoietic stem cells also have been characterized in detail and have similar properties to those in mice (13). They can be enriched using antibodies to surface markers such as CD34 and CD133 and can reconstitute the bone marrow and spleen of irradiated mice (severely immunodeficient mutant strains that do not reject human cells). In addition to bone marrow, umbilical cord blood has proven to be a valuable source of hematopoietic stem cells for therapeutic applications (14,15).

As in the bone marrow, the presence of stem cells in the intestinal mucosa was inferred initially from studies of recovery from radiation damage (16). Labeling experiments showed clearly that these stem cells are located near the base of crypts in the small and large intestines and give rise to the various mature cells of the gut. However, the intestinal stem cells have proven difficult to isolate and characterize outside of their in vivo niche. Recently, improved sorting and laser capture technologies have begun to facilitate the detailed phenotypic characterization of these cells (17,18).

Even under normal conditions, blood cells and intestinal epithelial cells turn over rapidly. The presence of stem cells in such adult tissues (the epidermal layer of skin serves as another example (19)) has become widely accepted. The general model is that stem cells, found in special niches, serve as a self-renewing population responsible for the capacity for tissue restoration. The stem cells give rise to progenitor cells, which can still multiply many times but are unable to indefinitely self-renew or to completely reconstitute a tissue. The progenitors, in turn, are precursors of terminally differentiated cells that generally no longer divide (20,21). Operationally, it can often prove difficult to distinguish rigorously between stem and progenitor cells, as adequate long-term in vivo reconstitution assays are not always available. Both classes are likely to have therapeutic value.

Stem and progenitor cells also have been identified in tissues and organs in which cells normally turn over much more slowly than in the bone marrow, gut, or skin. These include skeletal muscle, liver, kidney, endothelial cells lining blood vessels, and many more (22,23). Remarkably, even specialized cell types that for many years were believed incapable of any degree of restoration in adults, such as neurons of the central nervous system (CNS) and cardiac myocytes, now appear to have corresponding stem cells that persist throughout adult life (24,25).

**Mesenchymal Stem Cells**

While many adult stem cells appear specific to the tissue or organ from which they are obtained, others are more broadly distributed and also may have less restricted differentiation potential. An example with great clinical potential is the mesenchymal stem cell (MSC), a multipotent cell that can give rise to bone, cartilage, adipocytes (fat cells), and fibroblasts. MSCs were first identified as a stromal population in the bone marrow, distinct from hematopoietic stem cells (8,10,13,14). However, similar cells have been found throughout the body, and there are now clear indications that MSCs are perivascular cells associated with blood vessels (26). An accessible, abundant source of such cells is fat tissue obtained by liposuction (27) and also from placental tissue (unpublished observations). Although termed “stem cells,” the MSCs do not display indefinite self-renewal capacity and cannot be routinely expanded in culture beyond approximately seven or eight passages. A single bone marrow donation can yield thousands of potential clinical doses of MSCs but not an unlimited supply. A subset of MSCs from bone marrow appears to have greater capacity to proliferate, perhaps “indefinitely,” and also to differentiate beyond the mesenchymal cell lineages. Verfaillie and colleagues have termed them multipotent adult progenitor cells (MAPC). The MAPCs been reported to differentiate in vitro into numerous tissue types and also to contribute to a wide array of adult tissues and cell types after injection into a developing embryo at the blastocyst stage (14).

Some adult stem and progenitor cell populations have proven difficult to propagate in culture. For example, despite the great effort that has been expended studying hematopoietic stem cells, there has been little success in expanding their numbers in the laboratory. Other stem cell populations, such as satellite cells of skeletal muscle, multiply substantially in culture but then show diminished regenerative capacity when transplanted in vivo (28). Nonetheless, some cell populations that must include adult stem and progenitor cells have been expanded sufficiently for certain clinical applications. Examples include cells from human foreskin utilized for living skin products (29) and urothelial and smooth muscle cells used in tissue-engineered constructs for bladder augmentation (30). Furthermore, certain purified stem cell populations have been grown continuously in culture for many months—at least from a practical perspective, this may be viewed as “unlimited” self-renewal. Keys for expanding a number of adult stem cells have been enrichment of rare cells using selection with antibodies to surface markers, and the use of serum-free defined medium, carefully chosen growth factors, and, in some cases, extracellular matrix components for their expansion. For example, neural stem cells can be isolated from brain tissue by positive selection for CD133 and negative selection for markers found on contaminating cell types (31). When cultured in
defined medium with fibroblast and epidermal growth factors, they grow as "neurospheres" that maintain a significant proportion of multipotent stem cells able to give rise to neurons, astrocytes, and oligodendrocytes (32,33). Recently, conditions were reported for the isolation of human hepatic stem cells, by immunoselection for the epithelial cell adhesion molecule (EpCAM), and their expansion in culture for over 150 population doublings (34–36). These cells are precursors to hepatocytes and bile duct epithelium and may have broader potential to generate other endodermal cell types. Spermatogonial stem cells provide yet another example of a growing list of adult stem cells (albeit for a germ cell fate) that can be propagated extensively in a defined medium (37).

**Embryonic Stem Cells**

In contrast to adult stem cells, embryonic stem (ES) cells show essentially unlimited capacity for self-renewal in culture and should be able to give rise to any of the more than 200 cell types in the body (with the exception of certain extraembryonic cell types). The isolation of ES cells from the inner cell of blastocyst stage mouse embryos was reported in 1981 by Martin (38) and by Evans and Kaufman (39), based on many years of prior work on transformed embryonal carcinoma cells from teratocarcinoma tumors. Successful isolation of ES cells from human blastocysts was not achieved until 17 years later, by Thomson et al. (40).

ES cells have great therapeutic potential, but there are also significant barriers that must be overcome for their implementation in the clinic. ES cells are pluripotent, that is, they have the ability to form tissue representing all three embryonic germ layers—the ectoderm, mesoderm, and endoderm. Thus, the ES cell has the intrinsic ability to give rise to daughter cells that can form virtually any tissue in the body. This was demonstrated convincingly by the finding that mouse ES cells injected into developing embryos can contribute to all adult cell types, including germ cells (40,41). Human ES cells have been induced to generate representatives of all three germ layers in culture and are widely assumed to have the same degree of pluripotency as mouse ES cells.

The great plasticity of ES cells can also represent a drawback to their use, because it may prove more difficult to induce them to yield a single desired cell type exclusively than when starting with lineage-committed adult stem cells. However, significant progress has been made in the directed production of many specialized cells, or at least lineage-specific progenitors, from mouse and human ES cells (42–44). Examples include cells of neuronal, epidermal, cardiac myocytic, hematopoietic, endothelial, hepatic, endocrine pancreatic, and germ cell lineages. Another important problem that must be overcome is that undifferentiated ES cells form teratoma tumors (38–40,44). Therefore, for clinical use it will be essential to differentiate the cells quantitatively before implantation and to rigorously exclude the presence of residual stem cells. Finally, progress is being made in the derivation of human ES cells in the absence of xenogeneic feeder cells and animal proteins that may pose risks for clinical application.

The derivation and certain potential uses of human ES cell lines have engendered a significant ethical debate, especially because embryos, though donated for research, are destroyed in order to isolate the ICM (45,46). Technological advances may offer at least a partial resolution to the ethical controversy. For example, human ES lines can now be developed from single-cell biopsy at the eight-cell stage of development, without compromising the viability of the embryo, similar to a widely used procedure for prenatal genetic diagnosis (47). The development of ES-like cell lines by reprogramming of normal adult cells ultimately may put the debate to rest by providing pluripotent cells entirely without the use of donated embryos or oocytes (see below).

**Cells Derived From Somatic Cell Nuclear Transfer**

Somatic cell nuclear transfer (SCNT) involves the removal of an oocyte nucleus in culture, followed by its replacement with a nucleus derived from a somatic cell obtained from a patient. Activation with chemicals or electricity stimulates cell division up to the blastocyst stage, at which time the ICM is isolated and cultured, resulting in ES cells genetically identical to the patient. In animal studies, nuclear transferred ES cells derived from mouse fibroblasts, lymphocytes, and olfactory neurons are pluripotent and can generate live pups, showing the same developmental potential as fertilized blastocysts (48–50). The resulting ES cells are perfectly matched to the patient’s immune system.

Although promising, SCNT has certain limitations that require further improvement before clinical application. Animal studies have shown that blastocysts generated from SCNT and implanted into a uterus can produce an infant clone of the donor. In 1997, for example, a sheep named Dolly was derived from an adult somatic cell using SCNT (51). This form of reproductive cloning is banned in most countries for human applications. In contrast, therapeutic cloning can generate ES cell lines whose genetic material is identical to that of their source. In therapeutic cloning, blastocysts are grown until a 100 cell–stage where ES cells can be obtained, but the blastocysts are never implanted into a uterus. However, this technique has not been shown to work in humans. Initial failures and fraudulent reports of nuclear transfer in humans reduced enthusiasm for human applications (52–54). However, it was recently reported that nonhuman primate ES cell lines were generated by SCNT of nuclei from adult skin fibroblasts (55,56).

**Cellular Reprogramming**

Reprogramming is a technique that involves dedifferentiation of adult somatic cells to produce pluripotent stem cells without the use of embryos. Cells generated by reprogramming are theoretically identical to somatic cells that would not be rejected by the donor. This method also avoids nuclear transfer into oocytes. Takahashi and Yamanaka (57) were the first to report that both embryonic and adult mouse fibroblasts can be reprogrammed into an induced pluripotent state (IPS).
They identified four key genes required to bestow ES cell–like properties in fibroblasts: OCT3/4, SOX2, c-MYC, and KLF4. Reprogrammed cells were selected via drug resistance. The resultant IPS cells possessed the immortal growth characteristics of self-renewing ES cells, expressed genes specific for ES cells, and generated embryoid bodies in vitro and teratomas in vivo. When the IPS cells were injected into mouse blastocysts, they contributed to a variety of diverse cell types, demonstrating their developmental potential. It was recently reported that reprogramming by transduction of four defined factors can be done with human cells (58). Retrovirus-mediated transfection of OCT3/4, SOX2, KLF4, and c-MYC generates human IPS cells that are similar to human embryonic stem (hES) cells in terms of morphology, proliferation, gene expression, surface markers, and teratoma formation. Others (58) have shown that retroviral transduction of OCT4, SOX2, NANOG, and LIN28 could generate pluripotent stem cells without introducing any oncogenes (c-MYC) (59). An alternative approach would be to use a transient expression method, such as adenovirus-mediated system (59,60). Modified donor cells may not be necessary to generate IPS cells (60).

**Stem Cells Derived From the Amniotic Fluid and Placenta**

The amniotic fluid is known to contain a heterogeneous population of cell types derived from the developing fetus (61,62). Cells from this population have been shown to differentiate into adipocytes, osteocytes, neurogenic, and endothelial cells (63–66). Importantly, a subpopulation of OCT4-positive cells have been detected in the amniotic fluid (64,67). Recently, the isolation of human and mouse amniotic fluid-derived stem (AFS) cells was reported. These cells are capable of extensive self-renewal and give rise to adipogenic, osteogenic, myogenic, endothelial, neurogenic, and hepaticogenic lineages (68). In this respect, they meet a commonly accepted criterion for pluripotent stem cells.

AFS cells represent approximately 1% of the cells found in the amniotic fluid. The same cells can also be found in the placenta. These cells are immunoselected for cells that express the surface antigen c-kit (CD117), the receptor for stem cell factor (69). AFS cells express embryonic markers such as OCT4 and SSEA4 but not other markers of ES cells. They also express mesenchymal and/or neuronal markers such as CD29, CD44, CD73, CD90, and CD105. Clonal analyses using retrovirally marked human lines confirmed that differentiated cells of various types can be derived from a single AFS cell (68).

In addition to demonstrating that differentiated cells express lineage-specific markers, we have shown that AFS cells give rise to specialized functions (68). Cells differentiated down a neuronal pathway secreted glutamine or expressed G-protein–gated inwardly rectifying potassium channels. Cells of the hepatic lineage secreted urea and α-fetoprotein, while osteogenic cells produced mineralized calcium.

The recent discovery of a stem cell population in the amniotic fluid and placenta offers a promising alternative source of stem cells for cellular therapy. The full range of adult somatic cells that AFS cells can produce remains to be determined, but their ability to differentiate into cells of all three embryonic germ layers and their high proliferation rate are two advantages over most adult stem cell sources. AFS cells represent a new class of stem cells with properties somewhere between embryonic and adult stem cell types. However, unlike ES cells, AFS cells do not form teratomas and are easily obtained without destruction of embryos. AFS could be use for both autologous and allogeneic.

**Rehabilitation Patients and Potential Impact of Stem Cell Therapy**

**Spinal Cord Injury**

A target for cellular therapy in spinal cord injury (SCI) arises from the fact that damaged neurons of the CNS are unable to self-repair after damage. Transplantation of MSCs into damaged neurons of the CNS may create an anti-inflammatory microenvironment permissive to outgrowth of the proximal axon stump (70–72). Intriguing studies on olfactory ensheathing cells, specialized glial cells that surround sensory axons in the nose, suggest that when these cells are used to reestablish previously damaged pathways, severed axons have a structural connection onto which they may regenerate (73). According to the pathway hypothesis, olfactory ensheathing cells provide a “bridge” along which nerve fibers, with limited latent growth potential, may regrow. Although functional improvement after SCI has been reported in animal models following cell transplantation (74,75), only a few clinical studies of cell therapy approaches to SCI are reported. For example, a trial of bone marrow cell transplantation into SCI patients (n = 35) reported no safety concerns, yet the data indicated only meager functional improvements (76). Of interest, two SCI patients who received both autologous MSC transplantation and rehabilitation therapy demonstrated significant functional improvement (77). The question of how rehabilitation treatment might benefit SCI patients receiving stem cell transplantation will be an important area for future investigation.

**Brain Injury**

Improvement in neurological function was reported following MSC transplantation in brain injured rats. In one study, 2 × 10⁶ human MSCs were infused intravenously (78) without evidence of tissue rejection. In rat models of traumatic brain injury (79) and cerebral ischemic stroke (80,81), cell chemotaxis, or homing, was observed. Homing refers to the directional migration of cells, regulated by chemokines associated with tissue damage. The ability of stem cells to migrate to diseased/injured cells, tissues, and organs is central to therapeutic utility of MSC transplantation. MSCs are naturally home to areas of tissue damage, and thus offer an attractive source of stem cells for *in vivo* regenerative applications. The mechanisms governing the efficiency of MSC homing are the topic of a number of studies (82–84).
**Myocardial Infarction**

Animal models of myocardial infarction (MI) have demonstrated migration of MSCs to the injured heart with subsequent improvement in cardiac function (85–91). A comparative study of three MSC delivery methods suggested that following MI, local endocardial injection of MSCs might be a preferred technique compared to intracoronary or intravenous infusions. Endocardial injection resulted in better engraftment of cells in damaged cardiac muscle, decreased engraftment in other tissues, and did not impair coronary blood flow (92). A clinical study of post-MI patients showed that intracoronary injection of radiolabeled autologous bone marrow cells resulted in only 1.3% to 2.6% engraftment, with much of the transplanted cells lost to the spleen and liver (93). Migration and subsequent engraftment of MSCs in the heart is mediated by homing/chemokine factors, monocyte chemotactic protein-3 (MCP-3) (94), and SDF-1 (also known as CXCL-12) (95–97). Pretreatment of a stem cell population with these factors may enhance their migratory capacity and improve their utility in therapeutic applications. Interestingly, MSC homing and engraftment into damaged or diseased cardiac muscle does not appear to directly improve cardiac function (98,99) but rather seems to create an environment permissive for regeneration and matrix remodeling by synthesizing and secreting bioactive factors (100).

**Lung Injury**

Two murine studies (101,102) indicate that MSCs, and some bone marrow–derived cells, are retained in the lungs following lung injury and that endogenous MSCs may contribute to the attenuation of pulmonary fibrosis. These observations support a theory that MSCs create regenerative microenvironments in injured tissues by secreting factors that inhibit inflammation/scarring and apoptosis and stimulate angiogenesis and matrix remodeling (103,104). Molecular players in this process include the chemokine receptor CXCR4 (and its ligand, SDF-1) (105). Few human studies are available, but a recent review (106) suggests that the immunosuppressive and regenerative properties of MSCs used in graft-versus-host disease (107) hold promise for application in idiopathic pulmonary fibrosis in patients.

**Bone Repair**

Osteogenesis imperfecta (OI) results from mutations in the gene encoding collagen I that results in defective collagen production and inadequate mineral deposition. This results in brittle, easily broken bones. Two clinical reports suggest that stem cell transplantation can be an effective treatment for OI (108,109). In these studies, children with OI were given a mild course of chemotherapy followed by allogeneic, immunomatched whole bone marrow. Clinical results showed an improvement in growth velocity of 60% to 94% of the predicted median values for age and sex-matched children, within the 6 months post infusion, whereas only 0% to 40% growth velocity occurred in the 6 months pre infusion. MSCs when used in conjunction with engineered scaffolds may be used for large bone defects resulting from fractures (110).

**Muscle Degenerative Diseases**

Bone marrow stromal cells, injected locally into muscles of mdx-nude mice, indicate the tremendous potential for these cells as therapy for muscle degenerative diseases such as Duchenne muscular dystrophy (DMD) (111). Toward this goal, several criteria exist for selection of an ideal cell population (112). Among these criteria is the ability of stem cells to be introduced systemically, because in order to ameliorate the dystrophic pathology, dystrophin will be required in muscles throughout the body—particularly cardiac muscle and limb, diaphragm, and pharyngeal skeletal muscles. Furthermore, restoration of dystrophin in skeletal muscle alone may further exacerbate cardiomyopathy (113). The intra-arterial introduction of mesoangioblasts, a stem cell population associated with blood vessels, resulted in significant improvements in muscle morphology and function in dogs with severe dystrophy (114). This canine model of muscular dystrophy, termed GRMD for golden retriever muscular dystrophy, will be particularly important for preclinical trials of stem cell therapy. Of note, recent evidence suggests that perivascular cells from multiple tissues possess myogenic ability (115) (Crisan et al., 2008) (210) and also present a surface marker profile identical to MSCs. Thus pericytes, mesoangioblasts, muscle-derived stem cells (116), and MAPC (117) may all represent populations related to, or derived from, MSCs. This is clinically promising because it suggests that a cell population fulfilling the criteria (118) for effective cell therapy for muscle degenerative diseases is available from easily obtained sources such as placenta (119) and lipoaspirate (120).

**Application of Rehabilitation Techniques in Stem Cell Therapy**

The translation of stem cell therapy into positive changes for medical practice is clearly dependent on an increased understanding of environmental factors controlling stem cell survival, proliferation, and participation in tissue regeneration. Methods such as gene therapy and direct growth factor injection have been investigated in the laboratory as a means to manipulate the stem cell niche. Yet, each is accompanied by drawbacks, such as cost, feasibility, and safety issues, that limit their clinical application. Toward the end of finding a more practical approach to manipulation of the microenvironment, rehabilitation strategies such as exercise and neuromuscular electrical stimulation (NMES) have been investigated as a means to optimize stem cell transplantation. It was recently shown in murine models that the addition of a treadmill running protocol following the systemic delivery of bone marrow–derived stem cells resulted in enhanced donor cell contribution to myofiber regeneration (121). These changes were evident even after just 1 week of training. In a similar study, transplanted muscle stem cells show a significantly increased myofiber engraftment in dystrophic mice following a skeletal muscle overloading protocol (122). This enhanced engraftment was directly correlated with a loading-induced increase in skeletal muscle vascularity. More importantly, increased myofiber regeneration by transplanted cells was concomitant with...
an increased resistance to overloading-induced weakness of dystrophic muscle.

Functional improvements are the most important criteria for determining effectiveness of cell therapy strategies. While regenerative medicine approaches have been widely investigated for use in populations commonly seen by rehabilitation professionals, rarely have rehabilitation strategies been utilized to maximize regenerative medicine's therapeutic benefit. To realize the potential of cellular therapies for positive functional and clinically relevant outcomes, a paradigm shift may be necessary, and in the future, physical therapeutics may play a critical role in helping to fulfill regenerative medicine's clinical promise.

**Direct Gene Therapy: A Tool in Regenerative Medicine**

Gene therapy is based on the insertion of a normal or corrected copy of a gene into the diseased or damaged somatic cells of the recipient. While some may argue that gene therapy per se does not constitute regenerative medicine, the topic is covered here with the caveat that gene therapy can be one of several tools in the regenerative medicine tool kit. Because gene therapy is not yet mainstream medicine, it is generally considered experimental with associated risks and benefits (123). There are essentially two approaches (124) to direct modification of genes to correct underlying genetic defects or diseased tissue. The first strategy involves *ex vivo* transduction of the patient's own cells in culture followed by reinjection of genetically modified cells back into the same patient. The genetically modified cells subsequently secrete the desired gene product that corrects or mitigates the disease. An advantage to this *ex vivo* approach is that the use of the patient's own autologous cells reduces the risk of rejection compared to transplantation of nonautologous donor cells. Furthermore, cells can be genetically modified using nonviral vectors. Primary, autologous cell cultures also have a low risk of malignant transformation compared with off-the-shelf established cell lines.

A more commonly used alternative strategy, *in vivo* gene transfer, involves the direct delivery of genes, usually carried by viral vectors to the patient’s diseased or injured tissue. An advantage of the *in vivo* strategy is the production of the desired protein by the resident cells, without the risks of unwanted cellular alteration associated with *ex vivo* modification of cells. A disadvantage of direct *in vivo* gene transfer is the risk to the recipient of tissue rejection of the viral vector or the gene product, or both. To address immune response to viral vectors, immunosuppression can be given in conjunction with *in vivo* gene transfer. Combined use of viral vectors with nonimmunogenic preparations (such as liposomes) can address the issue of vector immune response (125).

**Gene Delivery Vectors**

The major obstacles for gene transfer delivery in clinical trials are low transfer efficacy, risk of mutagenesis when using integrating (i.e., retro- and lentivirus) vectors, and immune response to either the vector or the gene product (126).

The leading viral vectors used in clinical trials are lentivirus, adeno-associated virus (AAV), and herpes simplex virus (HSV) (127). The choice of vectors used for gene delivery is strongly influenced by the type of cells targeted with gene therapy. For example, if dividing cells (like cancer cells) are targeted, then vectors that can integrate into the host genome are required to preserve the integrity of transgene replication. RNA viruses (lentiviruses and retroviruses) are typically used for integrative gene therapy in dividing cells. If nondividing terminally differentiated adult cells are targeted, then nonintegrating viral vectors are used to mediate gene transfer. Two such nonintegrating viruses, adenovirus and AAV, are most commonly used in clinical trials. These viral vectors carry genetic material that does not integrate into the host cell's genome and are not replicated at cell division.

**Gene Targeting**

Nucleic acid strategies enable precise targeting of genes that regulate disease and provide an emerging class of therapeutic tools in regenerative medicine. More than 50 clinical studies have used nucleic acid “antisense” strategies in a variety of conditions, including cancer, cardiovascular disease, inflammation, and infection. One such strategy, RNA interference (RNAi), is a natural process that was first observed in plants, when RNA for pigment-producing genes was introduced into petunias to enhance color. Unexpectedly, these petunias lost nearly all of their color. The realization that these effects were mediated by dsRNA was made in *Caenorhabditis elegans* and termed RNAi by Andrew Fire and Craig Mello, recipients of a Nobel prize for this discovery. RNAi works by the following posttranscriptional mechanism: (a) Short interfering RNA (siRNA) of approximately 21 to 23 nucleotides are processed by an enzyme called Dicer and incorporated into an RNA-induced silencing complex (RISC). (b) The sense strand of double-stranded siRNA is cleaved during the formation of the RISC complex. (c) RNA helicases unwind the double-stranded siRNA, and the antisense strand guides RISC to the complementary target mRNA, which is subsequently cleaved by RISC. RNAi technology has been widely applied in research. For example, we have used RNAi to model a type of limb girdle muscular dystrophy in cultured cells (128). Others have used RNAi to model therapeutic benefit in neuromuscular diseases (129). An explosion of applications have resulted from this new technology, particularly in cancer research (130). For clinical use, RNAi meets the criteria of “targeted, less toxic” gene therapy, and will likely expand the numbers of patients who will ultimately benefit from its application (131).

**Antisense Oligonucleotides**

Antisense oligonucleotides (AOs) are single-stranded segments of DNA or RNA usually 15 to 25 bp in length. These short strands of genetic material are thought to work, in part, by their interaction with target mRNA and block translation into protein by preventing normal ribosomal movement along the transcript (132). Vectorless delivery of AOs may offer one solution to vector-induced immune response. Because AOs can
correct the reading frame of abnormal dystrophin transcripts that cause muscular dystrophy, a pilot clinical trial of AO therapy in four patients with DMD was undertaken with promising results (133).

**Rehabilitation Patients and Potential Impact of Direct Gene Therapy**

**Spinal Cord Injury**

An explosive growth in gene therapy research aimed at treatment for SCI has occurred since the mid-1990s. A large number of different strategies have been explored (134–183). Gene silencing strategies to minimize growth inhibitory signals associated with SCI have been explored using RNAi, while most other strategies have been aimed at enhancing regeneration of injured spinal cord axons. The biggest hurdle encountered thus far is the regeneration of axons across the site of SCI. Genetically modifying cells to express neurotrophins or growth factors within the damaged spinal cord increases neuron survival, sprouting, and regeneration. Overexpression of these factors is thought to increase axon growth and possibly mitigate effects of inhibitory molecules surrounding the area of injury.

Return of motor function following injury to the spinal cord is generally poor owing to the “hostile nature” of adult spinal tissue to axonal growth. One approach to this problem is the surgical implantation of a spinal cord graft, because damaged axons can regenerate into and across different types of grafts spanning the area of injury. Use of grafts made from Schwann cells or fetal spinal cord (135) has been explored and appears to enhance the growth of axons in rats (135). Gene transfer of neurotrophic factors (79,138,146,161,162,165,168,170,184–190) has also shown promise to enhance axon growth. For example, genetically modified fibroblasts expressing neurotrophic factors, glial cell-line derived neurotrophic factor (GDNF), NT-3, NT-4/5 and brain-derived neurotrophic factor (BDNF), grafted to spinal cord lesion sites in rodents induced axonal growth and in some cases induced functional recovery after SCI (140,141,166,191–193). Marrow stromal cells transduced to overexpress BDNF have been tested in SCI rats, but without significant motor improvement (168). The effectiveness of techniques used to deliver neurotrophic factors seems to be a rate-limiting step. To address this issue, AAV vectors have been used with some success. For example, improvement in hindlimb function was observed 16 weeks after AAV-mediated neurotrophin gene transfer in SCI rats. AAV was similarly used in rats to transduce BDNF with evidence of axonal growth of descending axons along with significant recovery of hindlimb function.

**Neurodegenerative Diseases**

Gene therapy holds promise as a means to correct the underlying molecular defects that result in neurodegeneration by manipulating gene expression in targeted cells within the CNS. Both gene transfer and gene silencing (194) have been studied in animal models of neurodegenerative diseases (195) with marked improvement in pathologic and functional measures. In the past, therapeutic gene delivery to the CNS has proven an impediment to the development of effective therapy for diseases of the motor neuron. With the discovery that AAV, when injected intramuscularly, is transported from presynaptic terminals to motor neuron nuclei in the spinal cord, came the ability to selectively target spinal motor neurons affected by motor neuron diseases. For example, intramuscular injections of AAV carrying IGF-1 prolonged the survival of the SOD-1 mouse, a preclinical model of amyotrophic lateral sclerosis (ALS) (159). SOD-1 mice injected with AAV-IGF-1, either before or after the onset of symptoms, exhibited higher muscle mass and increased motor neuron preservation compared to controls. Direct brain injections of AAV for gene delivery of the neurotrophic factors BDNF and GDNF were tested in a rat model of Huntington’s disease (196). AAV-mediated gene delivery appeared to provide neuroprotection to certain types of striatal neurons that are selectively vulnerable to degeneration following exposure to an excitatory amino acid. A line of research (197–201) has shown that neurotrophic factors offer significant neuroprotection in animal models of Huntington’s disease, thus supporting the preclinical relevance of AAV-mediated gene transfer strategy for patients with this disease. Neurotrophic factors delivered by viral vectors have been evaluated in similar models of neurodegeneration. For example, a rat model of Parkinson’s disease has been used to evaluate potential benefits of viral-mediated delivery of neurotrophic factors (202–205). In vivo rat experiments as early as 1998 demonstrated that viral-mediated GDNF could protect rat dopaminergic neurons from progressive degeneration (205). Clinical trials are underway to test AAV delivery of neurotrophic factors in selected Parkinson’s patients.

**Rheumatoid Arthritis**

Several viral and nonviral vectors are being used for direct gene transfer or gene silencing (206) in animal models of rheumatoid arthritis (RA). Gene therapy trials for RA are underway. New vectors inducing long-term and regulated gene expression in specific tissue are under development, resulting in more efficient gene transfer using viral vectors such as AAV (124).

**Muscle Degenerative Diseases**

A major obstacle for effective gene therapy in degenerative diseases of skeletal muscle is systemic delivery of vectors into diseased skeletal muscle and cardiac tissue. Both rodent and canine models have been used to develop and test intravascular approaches to efficiently transduce marker genes carried by AAV into the majority of skeletal muscle in animals, including the diaphragm and the heart (207–209). Pressurized venous infusion using vascular tourniquets or balloons allows for viral vectors to escape the intravascular space and penetrate into the surrounding skeletal muscle. A surgical technique allows for cardiac muscle transduction by placing the heart on cardiopulmonary bypass, clamping the aorta and pulmonary artery followed by AAV infusion into a cardioplegic heart at cold temperature. This method has shown unequivocal evidence of global gene expression in a significant proportion of cardiac myocytes.
Potential Impact of Regenerative Medicine Technologies on the Practice of Rehabilitation Medicine

Given the rapid technological progress in regenerative medicine, future rehabilitation treatment paradigms may significantly shift. We have experienced such paradigm shifts in recent years with advances in trauma care. For example, rehabilitation clinicians now care for patients with previously unimaginable recovery from massive multitrauma orthopedic injuries. With continued advances in regenerative medicine, rehabilitation clinicians may similarly see a number of patients with previously unimaginable survival achieved by organs or tissues regenerated with engineered constructs. The rehabilitation needs of such patients are unknown but will certainly require a holistic multiteam approach to restore full functionality. Through advances in cell therapy, clinicians may acquire new skill sets using cell or gene therapy injections for the treatment of chronic disabling conditions. Veterinary medicine is currently experiencing a paradigm shift in this respect, with some clinicians incorporating the use of autologous stem cell injection therapy for their canine and equine patients with arthritis and lameness. Advances in the understanding of biological processes driving tissue regeneration will certainly drive the field of rehabilitation in new directions that remain to be explored.

CONCLUSIONS

The interdisciplinary field of regenerative medicine uses stem cell and gene therapy technologies to replace or repair injured or diseased tissue in vivo. Stem cells and direct gene therapy have been explored in animal models and in a few clinical settings to regenerate damaged or diseased tissues in SCIs, brain injuries, MI, bone and joint diseases, and degenerative muscle diseases. The influence of rehabilitation techniques on the recipients of cellular and gene therapy are under study.

REFERENCES


Rehabilitation Robotics

INTRODUCTION

Rehabilitation robotics may be organized under four major mechatronic clusters, namely, (a) robotic-assistants, (b) prosthetics, (c) orthotics, and (d) therapeutic robotics. Robotic-assistants comprise devices that perform a task in lieu of the person. Prosthetics comprises devices that substitute for limb function. Orthotics aims at augmenting weak or paralyzed limbs by supporting loads and assisting or resisting relative motion between body segments. Finally, therapeutic robotics encompasses an emerging class of interactive robots that support and enhance the clinician’s task of facilitating recovery, delivering therapy, and evaluating patient progress. There are other ways to organize and characterize the scope of rehabilitation robotics. For example, another traditional way separates the devices into two broader categories: the intent to assist the patient in coping with the environment (assistive technology) and the intent to assist the clinician in delivering therapy and facilitating recovery (therapeutic technology). The first three mechatronics clusters mentioned earlier, robotic-assistants, prosthetics, and orthotics, fall under the assistive technology scope while the fourth, therapeutic robotics, falls under its namesake category. More recently, some researchers have organized rehabilitation robotics under a mixed classification system which is aggregated according to user group characteristic, to the source of inspiration for the mechatronics design, and finally according to the goal of the mechatronics system. For instance, many recent conferences are organized under three tracks, namely gerontechnology (user group characteristic), biorobotics (source of design inspiration), and neurorobotics (harness neurorecovery or neuron connectivity). It must be noted that part of the rationale for the recent spurt of neologism around rehabilitation robotics has to do not only with a sense of excitement and perceived opportunities of growth but also with the strong desire to distance itself from old labels and clearly delineate a new crossdisciplinary, biomedical and engineering endeavor.

THE PROBLEM

There is reason for this optimistic view for the rehabilitation robotics field. The demand for rehabilitation services is significant and is expected to grow apace with the graying of the population. For instance, consider cerebral vascular accidents (CVA) or stroke: according to the World Health Organization (WHO), every year 15M persons have a stroke, with a third of them left permanently disabled (1). We estimate that the number of persons who have survived a stroke at any given time is nearly ten times the yearly survival incidence. While there is no reliable data, one can estimate an annual need for rehabilitation services on the order of 50M patients worldwide. Of course, these figures are just “back-of-the-envelope” estimates and must be checked against accurate statistics and local variability. Nevertheless, these coarse estimates highlight the sizeable need and opportunity to deploy technologies such as robotics not only to promote recovery but also to assist a person to cope with residual impairment and disability. Robotics can revolutionize rehabilitation medicine by harnessing technology to assist, enhance, and quantify recovery and to ameliorate the quality of life and independence of the elderly and disabled. In this chapter, we focus exclusively on therapeutic robotics (for more information on assistive technology, orthotics, and prosthetics see (2,3)).

Therapeutic robots are a new class of clinical tools for evaluating patients and delivering meaningful therapy that enables an overdue transformation of rehabilitation clinics from preindustrial manual operations to technology-rich activities. Applications of robots to therapy are fairly recent, with the premier engineering journal in the field, the IEEE Transactions on Rehabilitation Engineering, coming to life only in 1992 (renamed in 2001 to IEEE Transactions on Neural Systems and Rehabilitation Engineering). In the last 10 years, the field of rehabilitation robotics has undergone vigorous growth. Borrowing an appropriate neologism, we “googled” Google Scholar with the words “rehabilitation robotics” and tracked yearly growth of publications in the field (Fig. 83-1). These data should be viewed with appropriate caution, particularly during the early years. But recent robust growth shows no sign of slackening and reflects a real phenomenon that is far from abating.

To keep pace with this fast-growing field is a daunting task, and our goal in this chapter is not to present a survey of the different existing devices, which would be doomed to obsolescence because of the pace of change. That task would better be achieved by tracking publications in the top journals in the field and the meta-analysis of results (4–9). Here, we attempt to highlight some basic robotic differences, present some limited outcomes employing both upper and lower extremity robotic devices, discuss some common concerns with the technology, and finally discuss potential mechanistic views of why it might actually work.
DIFFERENT ROBOTS: WHAT ARE THE DIFFERENCES?

This increase in activities has resulted in a multitude of different designs, different devices, and different choices for clinicians. Rather than make an exhaustive comparison of different designs, our goal herein is to discuss the critical differences among therapy robots. However, it is very important to stress that ultimately only exhaustive clinical results can actually determine what works, what does not, what works better, when to deliver therapy during the course of the illness or recovery, and, most importantly, for whom.

**End-Effector Versus Exoskeletal Designs**

The whole gamut of devices can be parsed into two broad categories of “end-effector” or “exoskeletal” designs (Fig. 83-2). Care must be exerted to understand the design trade-offs. Exoskeletal devices allow us to estimate with reasonable accuracy the kinematics of the human joints, but this advantage comes at a cost as it takes considerable time to properly adjust the lengths of the device links to each patient’s limb lengths. Note that proper collocation of the device and human joints might be necessary in some designs to prevent injury. Another advantage of an exoskeletal design is that it allows for a smaller package when therapy requires large human joint angle displacement. End-effector designs for very large human joint angle displacement require larger devices. Conversely, because the actuators (motors) are typically located at the joints of an exoskeletal device which must carry the weight of each actuator, large transmission reduction ratios are typically employed to reduce the weight of these motors, which lead to intrinsically large impedance (see below for more information). The end-effector based robots operate essentially by “handshaking” with the patients. While there may be more uncertainty about the kinematics of patients’ limb motions, donning and doffing times are minimal with this class of devices. Furthermore, because the motors are typically mounted on the base of the end-effector robots and not carried on the robot joints, the devices can employ larger actuators with smaller transmission reduction ratios (including direct drive designs) leading to intrinsically lower impedance devices—a very important consideration to enhance robot performance and stability (see for example MIT-Manus). As a general rule, we typically consider exoskeletal designs when the human joint must rotate by more than 60 degrees (e.g., MIT’s wrist robot); otherwise we prefer end-effector designs.

**Intrinsically High Impedance Versus Low Impedance Design**

Exoskeletal or end-effector designs employ motors which have not only weight and mass to be accounted for in any design but also static and viscous friction. A comparison of the ratios of energy to weight for human muscle and for an electrical motor demonstrates the impressive competence of muscles. Muscles generate two to six times more energy for the same weight (typical electrical motors 50 W/kg and muscle 300 W/kg). In many robotic devices for which weight is a critical limitation (e.g., exoskeletal robots like MIT’s Anklebot), speed-reducing transmissions are employed to increase the torque output without excessive increase in weight.* An unintended consequence

---

*From the conservation of energy, torque × angular velocity at the motor output shaft is the same as the torque × angular velocity at the transmission reduction output shaft (assuming no losses). Therefore, to increase torque output, we must reduce the angular velocity at the transmission reduction output shaft.
of employing this reduction to boost the actuator torque is augmenting intrinsic impedance, especially friction. The perceived actuator static friction† increases by the reduction ratio, and the viscous friction§ by the square of the reduction ratio. To illustrate, by employing a gear reduction 200:1 to reduce actuator weight, we increase the perceived static friction by 200 and the viscous friction by 40,000. Therefore the resulting robot is likely to have an intrinsically high impedance; that is, if the robot is in the “off” state, it requires very high force to move it. Conversely, an end-effector, direct-drive robot that can employ larger actuators which are fixed at the base and are not mounted on the robot arms, has very low intrinsic and end-point friction and can easily be moved. The resulting robot is likely an intrinsically low impedance robot (e.g., MIT-Manus).

For an interactive experience, the robot must allow even weak patients to express movement and the two distinct designs require very distinct control strategies to achieve this goal. An intrinsically high impedance robot must employ an aggressive control strategy to modify its behavior and reduce its inherent impedance to allow patients to express movement, no matter how weak. An intrinsically low impedance robot allows even weak patients to express movement, and simple control strategies can be employed to increase the robot’s inherent impedance whenever therapy requires guided or restricted movement. It is important to note that in practice, intrinsically high impedance robots suffer from a significant limitation: we cannot arbitrarily increase the gains of the controller to reduce the perceived impedance. Eventually, the robot stability, and therefore safety, is compromised.

Number of Degrees of Freedom
The optimum number of degrees of freedom is a difficult and critical question that requires considerable clinical insight supported by clinical results. The human body has many more degrees of freedom than any robotic device likely to be built. How many of those degrees of freedom should be actuated in a therapy robot? For instance, the human hand has 22 degrees of freedom and it is very unlikely that we will see any therapy robot in the near future that can interact in a controlled fashion with all of these degrees of freedom. This observation is valid for both upper and lower extremity. For the lower extremity, the ankle is a critical joint for gait. How many ankle degrees

† Friction is the force resisting relative movement between two objects. Static friction is the force between two objects not moving relative to each other. For example, static friction prevents a garbage can from sliding down a slope.

§ Viscous friction is the force resisting movement which is proportional to the relative velocity of two objects or fluid layers.
of freedom do we need to actuate during gait rehabilitation? These questions can only be answered via exhaustive clinical trials based on a clinical understanding of the target user population. For example, MIT’s new hand module was designed to train cylinder grasp (10). We opted for such an apparent limited design from our clinical understanding and experience with a previous eight degrees-of-freedom hand robot: the user of MIT robotic systems is likely a severe-to-moderate stroke patient with significant hemiparesis of the hands and fingers. For these patients, we simplified the mechatronic design considerably by limiting our goal and providing a tool to facilitate the training of grasp. We made a compromise to get a relatively low cost, light, and compact module that can easily be integrated with the shoulder-and-elbow MIT-Manus or wrist robots. Another case of design compromise is the Lokomat, designed to actuate the hip and knee in the sagittal plane but with no actuation at the ankle (11). Both robots embody reasonable compromises based on the best available expert insight at the time, but can we do better? Ultimately, all these designs must undertake rigorous and exhaustive clinical trials to determine their adequacy. This is a critical aspect that clinicians must demand and technology developers cannot ignore and must be prepared to undertake. This critical requirement makes the deployment of rehabilitation robotics the antithesis of rapid-fire demonstrations observed in many other technology arenas. Therapy robots and control algorithms must pass through a set of painstaking exhaustive clinical trials to determine, as noted previously, what works, what does not, what works better, when to apply the robotic tools and under what protocol, and for whom.

**CLINICAL RESULTS**

**Upper Extremity Rehabilitation and Stroke**

As discussed earlier, the biggest hurdle faced in therapeutic robotics is not the technology per se but its clinical validation. But every challenge is also an opportunity: robots provide an ideal platform for objective, reproducible continuous measurement and control of therapy. For a glimpse at the clinical results, meta-analyses are an outstanding source for an overall picture and comparison basis; there are at least three recently published articles that include a range of devices with more to come (4–8). Herein, we will summarize some of our results for the shoulder-and-elbow transport of the arm collected during a long series of clinical experiments involving well over 400 participants. We will then conclude reviewing the main findings of two meta-analyses. We will not include in this chapter a discussion of results with upper extremity distal robots (12).

**Subacute Rehabilitation**

Volpe et al. reported the composite results of robotic therapy with 96 consecutive stroke inpatients admitted to Burke Rehabilitation Hospital in White Plains, NY (13). These patients were on average 2 weeks after an acute first stroke. All participants received conventional neurological rehabilitation during their participation in the study. The goal of the trial was to amass initial evidence to test whether robotic therapy had a measurable impact on recovery. Hence, we provided one group of patients with as much therapy as possible to address a fundamental question: does goal-oriented movement therapy have a positive effect on neuromotor recovery after stroke?

Patients were randomly assigned to either an experimental (robot-trained) or control (robot-exposure) group. Individuals in the robot-trained group were seen for five 1-hour sessions each week and participated in 25 sessions of sensorimotor robotic therapy for the paretic arm. Patients were asked to perform goal-directed, planar reaching tasks that emphasized shoulder and elbow movements with their paretic arm. MIT-MANUS’ low impedance guaranteed that the robot would not suppress any attempts to move. When a patient could not move or could not be deviated from the desired path or was unable to reach the target, the robot provided soft guidance and assistance dictated by an impedance controller (14). This robot action (which we dubbed “sensorimotor” therapy) was similar to the “hand-over-hand” assistance that a therapist often provides during conventional therapy. It is interesting to note that this form of “assistance as needed,” which has been a central feature of our approach from the outset (and a challenge for our robot designs), has recently been adopted and promoted by other groups (15,16).

Individuals assigned to the robot-exposure (control) group were asked to perform the same planar reaching tasks as the robot-therapy group for only 1 hour per week. However, the robot did not actively assist the patient’s movement attempts. When the subject was unable to reach toward a target, he or she could assist with the unimpaired arm or the technician in attendance could help complete the movement. The robot supported the weight of the limb while offering negligible impedance to motion. For this control group, the task, the visual display, the audio environment (e.g., noise from the motor amplifiers), and the therapy context (e.g., the novelty of a technology-based treatment) were all the same as for the experimental group, so this served as a form of “placebo” of robotic movement therapy.

The study was “double blinded” in that patients were not informed of their group assignment and therapists who evaluated their motor status did not know to which group patients belonged. Standard clinical evaluations included (a) the MRC motor power score for four shoulder and elbow movements (MP, maximum score = 20) and (b) the motor status score which is divided into two subscales, one for shoulder and elbow movements (MS-SE, maximum score = 40) and a second for wrist and hand abilities (MS-WH, maximum score = 42) (17–20). The MS-SE and MS-WH have met standards for interrater reliability, significant intraclass correlation coefficients, and internal item consistency for inpatients (21).

Although the robot-exposure (control) and robot-treated (experimental) groups were comparable at admission, based on sensory and motor evaluation, on clinical and demographic scales, and the fact that both groups were inpatients in the same
stroke recovery unit and received the same standard care and therapy for comparable lengths of stay, the robot-trained group demonstrated significantly greater motor improvement (higher mean interval change ± SEM) than the control group on the MS-S/E and MP scores (Table 83-1). In fact, the robot-trained group improved twice as much as the control group by these measures. Notably, these gains were specific to motions of the shoulder and elbow, the focus of the robot therapy. There were no significant between-group differences in the mean change scores for wrist and hand function.

Our initial result supported our working model indicating that, to promote recovery, we need to harness the correct sequence of events much like the Hebbian rule for neural plasticity, which is colloquially summarized as “neurons that fire together wire together.” We believe that, in order to promote maximum recovery, events must occur in a proper sequence: the target is displayed, the patient attempts to move, movement then happens (either self-generated or assisted by the robot), and the sensory information comes back to the brain. Though this was a modest beginning, it provided unequivocal evidence that robotic therapy of the kind that might be delivered by a robot had a significant positive impact on recovery.

**Follow-up**
We were able to recall 31 of these 96 patients 3 years after discharge from the subacute hospital. The goal was to determine if the advantage conferred on the robot-trained group over the control group was sustained. It is possible that the short-term additional gains of the subacute phase were temporary and that, ultimately, patients in both the experimental and control groups would achieve the same clinical “plateau.” Such a result is quite important because it contradicts epidemiological studies which suggested that, after 12 weeks poststroke onset, there is little opportunity for further impairment reduction (22,23). It also demonstrates that the clinical “plateau” has little support from biology and there is an opportunity to impact recovery in persons with chronic impairment due to stroke, perhaps many years after the injury.

Bolstered by this result showing the positive and measurable impact of robotic therapy on neurorecovery, we are systematically investigating the multitude of variables that may influence outcome to determine their independence, their interaction, and their actual impact on outcomes in acute, subacute, and chronic phases of stroke. Our goal is the “holy grail” of robotic therapy, to determine how best to customize the treatment protocol to meet each individual patient’s needs.

**TABLE 83.1**  
Burke Inpatient Studies (N = 96) Mean Interval Change in Impairment and Disability (Significance p < 0.05)

<table>
<thead>
<tr>
<th>Between Group Comparisons: Final Evaluation Minus Initial Evaluation</th>
<th>Robot Trained (N = 55)</th>
<th>Control (N = 41)</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Impairment measures (±SEM)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fugl-Meyer shoulder/elbow (FM-SE)</td>
<td>6.7 ± 1.0</td>
<td>4.5 ± 0.7</td>
<td>NS</td>
</tr>
<tr>
<td>Motor power (MP)</td>
<td>4.1 ± 0.4</td>
<td>2.2 ± 0.3</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Motor status shoulder/elbow (MS-SE)</td>
<td>8.6 ± 0.8</td>
<td>3.8 ± 0.5</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Motor status wrist/hand (MS/WH)</td>
<td>4.1 ± 1.1</td>
<td>2.6 ± 0.8</td>
<td>NS</td>
</tr>
<tr>
<td>Disability evaluation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Functional independence measure (FIM)</td>
<td>32.0 ± 5.0</td>
<td>25.5 ± 6.5</td>
<td>NS</td>
</tr>
</tbody>
</table>

Mean ± standard error Motor Power scores (maximum score 20) of 96 patients on admission before rehabilitation, at discharge after rehabilitation and robotic training or control, and at follow-up evaluation approximately 3 years after stroke. Robot-trained patients maintain the motor improvements. *P < 0.05, versus control.

**FIGURE 83-3.** Change in motor power in patients recovering from stroke.
Continuous Passive Motion

Kwakkel has shown that a task-oriented training of high intensity is perhaps the most important aspect of a comprehensive stroke rehabilitation program (24). Volpe et al. tested whether intensity alone would suffice (25). To that end, they recruited subacute patients at the Burke Rehabilitation Hospital, White Plains, NY, and compared 25 minutes of additional conventional therapy with 25 minutes employing a CPM device (Continuous Passive Motion—Shoulder 600, Orthologic, Tempe, AZ), while patients’ attention was held by TV or social conversation. A CPM machine is an intrinsically high impedance device that employs position control and moves the patient’s shoulder or elbow passively. Patients cannot freely express any movement. Of note, there was no neurological advantage conferred on the CPM group. Patients in both groups showed comparable improvement in the neurological scales. This result reinforced our Hebbian view that, in order to improve outcomes, therapy must elicit the proper sequence of events and patients must participate actively and attempt to move. It is important to mention that while no neurological benefit was observed, the CPM group exhibited an apparent benefit in terms of shoulder stability and none of the patients in this group developed pain. Therefore, it became standard of care at Burke Rehabilitation Hospital to assign any stroke patient starting to show signs of shoulder pain to a daily session of movement therapy administered via the CPM machine.

Intensity-Matched Rehabilitation

The neurorehabilitation process is labor intensive, relying on therapy and evaluation procedures that are typically administered by a clinician working with a single patient. This one-on-one interaction characterizes much of the practice of clinical neurology. The repetitive nature of therapy makes it amenable to administration by properly designed robots. A robotic therapist can act as a modern, effective and novel tool, one that delivers a highly reproducible motor learning experience, quantitatively monitors and adapts to patients’ progress, and ensures consistency in planning a therapy program. Indeed, our research suggests that robotics can afford a much more intense experience than traditional therapy. For example, Kahn et al. noted that in a typical conventional therapy session, patients attempt to move their impaired arm circa 60 to 80 times (27). In a typical robot-mediated therapy session, patients must attempt to move 1,024 times. To confirm the importance of task-oriented training and intensity, we conducted a clinical trial including three groups of stroke patients: usual-care, robot-mediated therapy, and human-administered intensive-matched therapy. The last group was an attempt to match human-administered therapy and the intensity of typical robot-mediated sensorimotor therapy (1,024 attempts to move in each session). Though this latter form was somewhat artificial and impractical, results demonstrated that there were no significant differences in outcomes between robot-mediated and human-administered intensive-matched therapy groups; furthermore both groups improved far more than the usual-care group (28). These results further support Kwakkel’s assertion of the importance of a comprehensive rehabilitation program encompassing task-oriented movements of high intensity (24).

Motor Learning as a Model for Motor Recovery**

Given the apparent importance of high intensity and a patient’s active participation in therapy, we revised our robot control algorithm to test whether continuously challenging a patient would enhance recovery. The revised algorithm differs from our earlier sensorimotor therapy in three important ways (29).

First, during our earlier clinical trials, robotic therapy took the form of fixed, repetitive reaching exercises cued by a video display. An impedance controller with constant stiffness and damping made the therapy interactive: the force exerted by the robot varied continuously as a function of the deviation of the patient’s motion from a minimum-jerk trajectory of constant duration that connected the start position to the goal position. This system suited patients with limited motor ability for whom it provided assistance; however, it would also impede patients who moved faster than the nominal trajectory. Our revised algorithm used nonlinear impedance control to implement a “virtual slot” that extended between the start and goal positions and defined the appropriate coordination. Lateral deviation from the desired path was discouraged by the stiffness and damping of the slot sidewalls. Desired motion was assisted by moving the back wall of the slot along a minimum-jerk virtual trajectory so that the slot progressively “collapsed” to a “virtual spring” centered on the goal position. However, motion along the “virtual slot” (well aimed and faster than the nominal desired trajectory) was unimpeded.

A request for the subject to move was signaled by a target in the visual display changing color. If the patient failed to initiate movement within 2 seconds, the robot began to act (i.e., the back wall of the “virtual slot” closed on the goal position). This mode of triggering the robot encouraged even severely impaired patients to participate actively, rather than passively allowing the robot to drive the arm.

Second, the revised algorithm continuously monitored the patient’s performance. By combining records of the kinematics of actual patient motion and the kinetics of mechanical interaction between robot and patient, five performance measures (PMs) were computed: PM1 graded patients’ ability to initiate movement, PM2 measured movement range or extent, PM3 measured the power delivered by the patient to complete the movement, PM4 measured the smoothness of the movement, and PM5 measured aim or deviation from the straight line connecting the center and outbound target. We used these measures to adjust the parameters of the controller during a therapy session. For the first five cycles of movements to the eight goal positions and back to the center position, the time allotted for a movement (the duration of the nominal minimum-jerk trajectory) and the stiffness (impedance) of the

---

**Motor Learning as a Model for Motor Recovery**

Given the apparent importance of high intensity and a patient’s active participation in therapy, we revised our robot control algorithm to test whether continuously challenging a patient would enhance recovery. The revised algorithm differs from our earlier sensorimotor therapy in three important ways (29).

First, during our earlier clinical trials, robotic therapy took the form of fixed, repetitive reaching exercises cued by a video display. An impedance controller with constant stiffness and damping made the therapy interactive: the force exerted by the robot varied continuously as a function of the deviation of the patient’s motion from a minimum-jerk trajectory of constant duration that connected the start position to the goal position. This system suited patients with limited motor ability for whom it provided assistance; however, it would also impede patients who moved faster than the nominal trajectory. Our revised algorithm used nonlinear impedance control to implement a “virtual slot” that extended between the start and goal positions and defined the appropriate coordination. Lateral deviation from the desired path was discouraged by the stiffness and damping of the slot sidewalls. Desired motion was assisted by moving the back wall of the slot along a minimum-jerk virtual trajectory so that the slot progressively “collapsed” to a “virtual spring” centered on the goal position. However, motion along the “virtual slot” (well aimed and faster than the nominal desired trajectory) was unimpeded.

A request for the subject to move was signaled by a target in the visual display changing color. If the patient failed to initiate movement within 2 seconds, the robot began to act (i.e., the back wall of the “virtual slot” closed on the goal position). This mode of triggering the robot encouraged even severely impaired patients to participate actively, rather than passively allowing the robot to drive the arm.

Second, the revised algorithm continuously monitored the patient’s performance. By combining records of the kinematics of actual patient motion and the kinetics of mechanical interaction between robot and patient, five performance measures (PMs) were computed: PM1 graded patients’ ability to initiate movement, PM2 measured movement range or extent, PM3 measured the power delivered by the patient to complete the movement, PM4 measured the smoothness of the movement, and PM5 measured aim or deviation from the straight line connecting the center and outbound target. We used these measures to adjust the parameters of the controller during a therapy session. For the first five cycles of movements to the eight goal positions and back to the center position, the time allotted for a movement (the duration of the nominal minimum-jerk trajectory) and the stiffness (impedance) of the

---

**Parts of this section are similar to (30).**
“virtual slot” sidewalls were adjusted approximately to match the patient's current performance and need for guidance. This was important because patient performance typically declined between the end of one therapy session and the beginning of the next. For every subsequent five cycles of reaches to and from the eight goal positions, the controller parameters were adjusted based on the patient’s performance. The intent was to challenge the patient to improve.

As patients aimed better, the stiffness of the “virtual slot” sidewalls was decreased (and vice versa). As patients moved faster, the time allotted for movement was decreased (and vice versa). Again, this was intended to encourage active participation of even the most impaired patients, yet continuously challenge patients as they recovered.

Third, to provide motivation, positive reinforcement, and knowledge of results, our revised algorithm provided specific, movement-related feedback in the form of a simple graphical display that consisted of five vertical bars to reflect recent patient performance. The height of each of the five bars was determined by the five PMs, expressed as a percentage.

This performance-based progressive therapy (PBPT) algorithm provided support for patients to progress from complete plegia to normal arm movement. PM1, which measured ability to initiate movement, was probably most important for severely impaired patients and helped to ensure appropriate timing of afferent and efferent signals, which may be important for reestablishing the excitability of corticospinal projections (13,31–34). PM2 measured active ROM, an important clinical measure of function, and also rewarded hypertonic patients for relaxing their arms, which allowed the impedance controller to move their hands closer to the target. PM3 measured active power delivered by the patients, probably important for less severe patients that could reach most but not all targets. PM4 and PM5, which respectively measured movement smoothness and the aim or deviation from the straight line connecting the center to the outbound target, quantified the trade-off between speed and accuracy that is characteristic of unimpaired movement and was probably most important for patients with mild-to-moderate impairment.

The PBPT protocol was evaluated by a clinical study of 28 chronic stroke patients between the ages of 39 and 81 with chronic motor impairment following a single stroke that had occurred between 8 and 95 months before the initial assessment (35). All patients were evaluated six times: three times in a 2-month period before the start of therapy to assess baseline performance, at the midpoint and discharge from robotic therapy (18 one-hour sessions of robotic training, three times a week for 6 weeks), and finally at a follow-up evaluation session 3 months after robotic training. Evaluating therapists were different from treating therapists (Table 83.2).

The first three evaluations showed no significant changes on any of the impairment scales, which verified that subjects were indeed at the chronic phase of their recovery. Subsequent evaluations showed that the PBPT protocol evoked a statistically significant improvement in motor performance that was maintained at the 3-month follow-up. This result confirms earlier studies of chronic-phase patients (36–38) and shows that amelioration of chronic neurological impairments long after the expected period for recovery following stroke is possible. More important for our understanding of recovery, the magnitude of the improvement achieved with PBPT was many times greater than that achieved with our previous robotic therapy. Of course, a large randomized clinical trial (RCT) must follow to confirm this finding. Nevertheless, it is worth stressing that the only change was the therapy protocol: the same robot assisted with the same set of reaching movements. A treatment protocol that adapted to the patient’s motor ability and presented a continuous challenge substantially enhanced recovery.

**Meta-Analysis**

While the gold standard in clinical research is the RCT, meta-analysis compiles results from multiple groups and affords a map of the land. Here we summarize two meta-analyses (7,8) (Fig. 83-4).

Prange et al. conducted a systematic review of the literature between 1975 and August 2005 (7). They identified articles on the use of different robots to deliver therapy to stroke patients including MIT-Manus, Assisted Rehabilitation and Measurement (ARM) Guide, Mirror Image Motion Enabler (MIME), Bi-Manu-Track, Gentle/s, Neurorehabilitation Robot (NeReBot), REHAROB, Arm Coordination Training 3D (ACT), ARMIn, Exercise Machine for Upper Limbs (EMUL), Rutgers Master II (RMII), and Mechatronics system for Motor recovery after Stroke (MEMOS). They identified

<table>
<thead>
<tr>
<th>Severity</th>
<th>Impairment Measure (Mean ± SEM)</th>
<th>F-M SEC (Max = 42)</th>
<th>% Change</th>
<th>MP (Max = 70)</th>
<th>% Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moderate</td>
<td>Before treatment</td>
<td>17.0 ± 1.3</td>
<td>37.2 ± 2.5</td>
<td></td>
<td>52.0 ± 2.2</td>
</tr>
<tr>
<td></td>
<td>After treatment</td>
<td>22.5 ± 1.3</td>
<td>45.4 ± 1.7</td>
<td>46.5 ± 1.9</td>
<td>22.0 ± 2.2</td>
</tr>
<tr>
<td>CNS &gt; 4; NIHSS &lt; 15</td>
<td>Follow-up (3 mo)</td>
<td>24.5 ± 0.9</td>
<td>46.5 ± 1.9</td>
<td></td>
<td>25.0 ± 2.2</td>
</tr>
<tr>
<td>Severe</td>
<td>Before treatment</td>
<td>8.2 ± 0.7</td>
<td>17.3 ± 1.8</td>
<td></td>
<td>17.3 ± 1.8</td>
</tr>
<tr>
<td></td>
<td>After treatment</td>
<td>10.9 ± 0.9</td>
<td>17.3 ± 1.8</td>
<td>23.7 ± 2.0</td>
<td>37.0 ± 2.0</td>
</tr>
<tr>
<td>CNS &lt; 4 NIHSS &gt; 15</td>
<td>Follow-up (3 mo)</td>
<td>12.5 ± 0.9</td>
<td>26.3 ± 2.2</td>
<td>26.3 ± 2.2</td>
<td>52.0 ± 2.2</td>
</tr>
</tbody>
</table>

<sup>a</sup> Denotes significant change, p < 0.001.

F-M SEC, Fugl-Meyer, shoulder-elbow component; MP, motor power; CNS, Canadian Neurological Scale; NIHSS, National Institutes of Health Stroke Scale.
17 clinical trials, of which only 4 ultimately were included in the final analysis. These studies involved three robots: MIT-Manus, MIME, and Arm Guide; with an aggregate of 178 experimental subjects and 50 control subjects. Prange et al. concluded that “robot-aided therapy of the proximal upper limb can improve short- and long-term motor control of the paretic shoulder and elbow. This conclusion is supported by a quantitative analysis of short-term pooled FM data in chronic stroke patients and indicates that increased motor recovery of chronic patients is possible after robot-aided therapy. However, no consistent effect on the improvement in functional abilities has been reported. Restoration of motor control appears greater after robot-aided therapy than conventional therapy. We could not establish which aspects of robot-aided therapy (e.g., increased intensity of movements, most effective training modalities) were most responsible for the beneficial influence on recovery. The clinical relevance of our findings is that robot-aided therapy is a promising new approach to rehabilitation of upper-limb motor control after stroke. For both subacute and chronic stroke patients, robot-aided therapy can improve motor control of the hemiparetic upper limb, perhaps even more than conventional therapy.”

Kwakkel included studies published by October 2006 (8). A computerized literature search was conducted in MEDLINE, CINAHL, EMBASE, Cochrane Controlled Trial Register, DARE, SciSearch, Doconline, and PEDro and it returned 173 hits. Contrary to Prange’s meta-analysis, Kwakkel et al. opted to include only papers that compared robot training against a control group. Excluded were many studies included in Prange’s meta-analysis that compared different forms of robotic therapy and studies on chronic stroke that compared discharge values with admission values. Ultimately, ten studies involving 218 patients achieved methodological quality and were included in further analysis. These studies involved four devices: MIT-Manus, MIME, Bi-Manu-Track, and Arm Guide. Analysis of the entire set indicated nonsignificant effect, possibly due to heterogeneity. Kwakkel et al. eliminated the outliers in further analysis (Kahn 2001; Hesse 2005) and concluded that “sensitivity analysis showed a significant improvement in upper limb motor function after stroke for upper arm robotics. No significant improvement was found in ADL function. However, the administered ADL scales in the reviewed studies fail to adequately reflect recovery of the paretic upper limb.”

Gait Rehabilitation and Stroke
Restoration of gait is a major aspect of stroke rehabilitation. Three months after stroke, 25% of surviving subjects remain wheelchair-dependent and in 50%, gait speed and endurance are considerably reduced (39). Traditional treatment concepts in Europe put emphasis on tone-inhibiting and gait preparatory maneuvers, assuming a transfer of skill acquisition from one motor task to the other. Gait itself is practiced very little, rarely more than 50 to 100 steps per session. Not surprisingly, a large outcome study failed to show a relevant improvement of gait function and symmetry in 160 ambulatory subjects following 4 weeks of Bobath treatment (40). Modern concepts of motor learning favor a task-specific, repetitive approach, which for gait, we translate into “who wants to regain walking has to walk.”

Treadmill training with partial body weight support was introduced in the early 90s to intensify gait practice (41,43).
It employed a harness not only to support patients with poor balance but also to unload the paretic leg, while the treadmill enforced locomotion. The repetitive nature of gait therapy makes it amenable to administration by properly designed robots. Such robots relieve the strenuous repetitive effort of therapists during gait rehabilitation of wheelchair-dependent subjects on a treadmill. Although various robot-mediated gait training devices are available, the limited number of clinical studies have largely used the Lokomat (Hocoma, Zurich, Switzerland) and Gait Trainer I (GT I, Reha-Stim, Berlin, Germany).

The Lokomat combines a motorized, body-weight supported treadmill (BWST) with a robot-driven gait orthotic (44). During training, lower extremity movements are controlled at the hip and knee, while the pelvis and thorax are stabilized. If necessary, the foot is dorsiflexed with lifter straps to prevent the forefoot or toe box of the shoe from catching, should foot drop be present. The robot-driven hip and knee movements are synchronized with the treadmill. A number of gait variables can be regulated. The primary variables are body-weight support (0% to 100%) and treadmill speed (0 to 3.2 km/h). The range of motion at the hip and knee can be increased or decreased, and the trunk can be positioned in slight flexion or extension. These adjustments enable fine-tuning of subtle issues such as toe clearance, stride length, or heel strike placement in transitioning from swing to stance phase.

The GT I applies the principle of movable footplates, where each of the patient’s feet is positioned on a separate footplate whose movements are controlled by a planetary gear system, simulating foot motion during stance and swing (45). Cadence and stride length can be tailored to the individual patient’s need. The vertical and horizontal movements of the center of mass are controlled via ropes attached to the harness. The device affords bilateral and distal guided gait training. Patients’ knees are not fixed. Therapists have easy access to patients which allows minor corrections of the knee motion if needed. Contrary to the Lokomat, the GT I imposes minimum number of constraints on natural hip and leg motion. Tight attachment of the leg to a robot with a technically limited number of degrees of freedom results in a more constrained leg motion.

Following are some results on stroke recovery with the electromechanical GT I, shown in Figure 83-5 and the Lokomat.

DEutsche GAngtrainer Studie

Four centers recruited 150 (155) nonambulatory first-time stroke patients, who could sit and hold the trunk straight at the edge of the bed for 5 minutes and had no severe arthritis or contractures. Patients were randomly assigned to two groups. Group A received 20 minutes GT I plus 25 minutes of standard physiotherapy (PT) while group B received 45 minutes of standard PT. Therapy was administered daily except during weekends and lasted for 4 weeks. Responders were those patients who were able to walk (Functional Ambulation Category (FAC) 4 or 5) or who reached a Barthel Index (BI) of at least 75 (45).

At the end of the treatment, significantly more patients in group A had become responders: 41 versus 17 with respect to FAC ($p < 0.0001$, Fig. 83-6) and 44 versus 21 with respect to the BI 75 ($p < 0.0001$). At the 6-month follow-up, the superior gait ability persisted (54 vs. 28, $p < 0.0001$), while the BI responder rate did not differ. For all secondary variables, including gait velocity and gait endurance, the experimental group had improved significantly more ($p < 0.0001$) during the treatment period but not at follow-up. For a subgroup of patients ($N = 60$), heart rate intensity was recorded during the training with patients in group

![Electromechanical Gait Trainer GT I](image)
A exercising longer at the target zone heart rate than patients in group B (16.1 vs. 5.3 min, \( p < 0.001; \ (46) \)).

While DUeutsche GAngrtrainer Studie (DEGAS) is presently the largest RCT on robot-assisted gait rehabilitation, there are other randomized trials using the GT I (47, 48). Studies with the Lokomat and stroke involved smaller sample sizes, with the largest study involving forty eight (50) chronic stroke survivors receiving 12 sessions of 30 minutes of either Lokomat training or therapist-assisted training at similar speeds (49). The results favored the therapist-assisted training over robotic training particularly for the less severe survivors. While more studies are needed, this study clearly exemplifies the importance of corroborating clinical evidence and of the understanding of the different variables that might influence outcome and that in most cases can be controlled via software modification.

The most up-to-date review and meta-analysis of the clinical results employing robotic-assisted gait training on stroke, including the GT I and Lokomat (Hocoma), are included in a recent Cochrane review. Mehrholz pooled all available data, including abstracts, and concluded that patients who receive electromechanical-assisted gait training in combination with PT after stroke are more likely to achieve independent walking than patients receiving gait training without these devices (9). Besides the GT I and Lokomat, Mehrholz considered the UC Irvine’s PAM-POGO, MIT’s Anklebot, and the Autoambulator (Healthsouth) among others, but these devices presently lack the clinical results to substantiate their use (50).

**Gait Rehabilitation and Multiple Sclerosis**

We have described application of robot-mediated therapy in persons with upper and lower impairments due to stroke, but the potential of this technology is not limited to stroke. Here, we describe initial results on the use of the technology to gait rehabilitation in multiple sclerosis (MS), but there are other applications of the robotic technology in spinal cord injury, cerebral palsy, traumatic brain injury, Parkinson’s disease, postsurgical rehabilitation following orthopedic trauma, etc. In MS, our evolving understanding of the disease process, particularly pathophysiology as it relates to physical activity, has translated into the development of new training protocols.

MS differs from stroke in a number of ways, including disease characteristics and sites of neurological injury. These differences in pathophysiology, etiology, site of injury, chronicity, and variable course can lead to additional treatment challenges for MS. MS is both an autoimmune inflammatory process as well as a disease process involving the loss of neurons and axons along the central nervous system (CNS) (51). In MS, neurological sites of pathology are widespread and multifocal and include the white matter tracts, the gray matter of the cerebral hemispheres, the subcortical structures, the cerebellum, and the spinal cord. In addition to primary motor, sensory and cerebellar deficits, fatigue is a very prevalent and prominent issue that can limit physical activity (52, 53). The disease process for MS is not time restricted; rather, MS is an active disease, generating ongoing acute and chronic neurological impairments. The clinical course for MS is characterized by relapses and remissions superimposed over progressive disability.

In the past, because of the ongoing inflammatory process in MS, there was the fear of possibly triggering an exacerbation, and clinicians have historically been cautious in recommending exercise (54). Nevertheless, this concern has not been substantiated by the literature on exercise and rehabilitation (55), and presently, the general recommendation is that persons with MS should participate in exercise and physical activity. This recent shift in exercise philosophy has been accompanied by growing interest in pursuing rehabilitation studies for MS. Since disability in MS accumulates over a period of time, there is an opportunity for early intervention to alter the natural course of functional decline.

Although various robot-mediated gait training devices are available, the limited number of studies conducted in MS populations has largely used the Lokomat (Hocoma, Zurich, Switzerland).

Beer et al. recently published the results of an RCT comparing Lokomat robot-assisted gait training to conventional walking training in MS patients with Kurtzke Expanded Disability Status Scale (EDSS) scores ranging from 6.0 to 7.5 (56). The protocol consisted of daily 30-minute sessions of either robot-assisted training (\( N_1 = 19 \)) or conventional training (\( N_2 = 16 \)) for a total of 15 sessions. Gait training of either type was supplemented with 16 hours per week of general physical therapy or occupational therapy. Primary study outcomes included walking velocity (20 m walk), 6-minute walk (6 MW), stride length, and knee extension strength. Results showed that both groups improved on walking velocity (28% and 29% for robot-assisted and conventional training group, respectively), whereas the robot-assisted training group improved slightly more on distance (9.5% compared to 4.5% for the conventional training group) (Table 83-2).

In another study, one of us conducted a crossover study, comparing robot-assisted gait training with BWST training in 13 subjects with either relapsing-remitting or primary progressive MS (EDSS 4.9 ± 1.2) (55). Subjects received a total of twelve 45-minute training sessions, either starting with BWST alone or BWST with robotic assistance, and then crossing over to the alternative intervention. Gait outcomes included walking velocity (20 m walk), 6-minute walk, EDSS, double support time (DST), and step length ratio (SLR). At midpoint (prior to crossing over), median percent changes for gait outcomes were actually greater for the treadmill alone group compared to the treadmill with robot-assisted group for T25FW (37.6% vs. 15.9%) and DST (−5.9% vs. −1.9%), whereas robot-assisted group had slightly greater change for 6MW (22% vs.19%) and SLR (2.0% vs. 0.4%). However, none of these differences achieved statistical significance.

These pilot studies show that robot-assisted rehabilitation is feasible and safe in the MS population. While these pilot studies did not demonstrate therapeutic advantages of robot-assisted therapies over more conventional therapies such as treadmill training, the labor-intensive nature of the latter confers a significant advantage to the employment of robot-mediated therapy.
DISCUSSION

Neurological Considerations for Motor Neurorecovery

After a brain injury, the neurophysiology that underlies recovery depends on the flexibility of remaining neurons and the altered synapses in the multiple circuits to reorganize the sensory motor processing that had developed over a lifetime. Nineteenth-century neuroscientists thought about recovery using concepts like “diaschisis”—resolution of the depression of functionally connected but anatomically distant circuits, so that recovery required the unmasking of functionally inactive connections. Another consideration of recovery was called “vicariation”—a particular brain circuitry that had been dedicated before the injury to contributing an aspect of sensory motor function and which might, after the injury, contribute to the control of a new function. In the 20th century, imaging and detailed electrophysiological data supported these ideas (58–60). In the 21st century, the promise of recovery has moved into the molecular age and it appears more to depend on learning mechanisms that may in turn be facilitated by task-specific intensive human/machine interactions (29), the effect of transcranial direct current (61–63) or magnetic stimulation (63–65), the potential of persistent neurogenesis (66,67), the possibilities of transplantation of neurons or viral factories for the production of transmitters, growth and transcription factors, and a variety of other signaling molecules that amplify the learning conditions (68). As a result, recovery can be revised by operationally testable terms like reorganize, as in train networks that underlie behaviors (69–71); replenish, as in increase depleted dendritic fields and their synaptic spines (72); replace, as in transplant cellular or molecular slurries of neurons and glia or deliver molecules that abet the formation of new connections (73,74), support or repair damaged connections (75–77). All of these modern mechanisms play a role in recovery after neurological injury.

Neurological injury fills textbooks with its diversity. However, for the purpose of this discussion we restrict ourselves to stroke which plays out over minutes, hours, and days, but recovery processes play out over months, perhaps many months and years. Recovery depends on the severity (78), details of the location (79,80), and the type of injury and concurrent medical condition (81). Treatment for stroke also proceeds in stages that require multimodal protocols. Stabilization of the crisis by clot dissolution or removal, vessel repair, or blood pressure control gives way to the beginning of recovery, decreasing the risk of a second event. Therapy for recovery proceeds with general efforts to encourage the patient to move and to exercise in a wide array of training protocols and now with an equally broad array of devices. Because the currency of the coordinated activity of neurons is behavior, it is no surprise that there is much information about the effect of environment and, more particularly, the nurturing aspects of environment on the development of motor behavior. Early studies of the effect of an enriched environment on the developing nervous system inform current models of recovery that have expanded to include task-specific training (82–84). It is this background that has led trained therapists into the sick room of the patient with stroke, first to passively move the paralyzed limbs and then actively guide the performance.

Physiological Considerations for Motor Neurorecovery

Traditional therapies do not recognize the long-term detrimental effects of reduced mobility, deconditioning, and muscular wasting. Investigators have long recognized that energy demands of paretic gait are 1.5- to 2-fold higher. It has also been shown that a mean VO$_2$ of 10 mL/kg/min is required to maintain slow walking effort. Yet, Macko et al. have shown in a large cohort of persons with chronic stroke that their peak VO$_2$ was just 13.8 mL/kg/min, which is 50% lower than sedentary age-matched healthy individuals; it indicates a VO$_2$ consumption of 70% of their peak capacity just to walk very slowly (85). This finding reveals a detrimental combination of poor fitness and energy inefficient gait potentially leading to immobility. Hence, the rationale for exercise to improve fitness even without neurological improvement is solid.

Fitness programs might include many modalities including robotic training. As described earlier, for a subgroup of 60 patients of the DEGAS study, the heart rate intensity during training with the robotic-assistant was higher than for patients receiving standard therapy (45). Furthermore, robotic devices can be used in many fashions including resistive exercises, which might be even more beneficial from the fitness viewpoint. It is important to note that older literature on exercise therapy for individuals with conditions resulting in “upper motor neuron” neurological damage (principally stroke, traumatic brain injury, spinal cord injury, MS, and cerebral palsy) raised theoretical concerns about causing increases in abnormal muscle tone, especially with resistive exercises (86). More recently, several studies have been conducted to clarify the impact of resistance training on spasticity and found that it has little effect (87–91).

While resistive exercises for the paretic limbs appear safe after stroke and might improve fitness, their impact on neurorecovery remains uncertain. Studies of strength training have not found improved motor control or function in the upper limbs (92–94), but have found modest benefit in the lower limbs (87–91).

Robotic studies of the MIT-Manus/InMotion2 robot also reveal the same stereotypical pattern of flaccidity followed by increase in tone and spasticity for subacute patients (17), but interestingly, both for subacute as well as for chronic strokes, robot training appears to have no impact on tone and spasticity when compared against standard therapy. In fact, there is a trend toward reduced spasticity (35–37), with the incorporation of resistance training into robotic therapy having no measurable impact (38). A study of children with hemiplegia from cerebral palsy employing the same device found small but statistically significant reductions in spasticity. This occurred primarily at the elbow after a course of training with the InMotion2 robot (95). Likewise, a pilot study of an EMG-controlled powered elbow brace found significant reductions in spasticity at the elbow (96).
Similar results were observed for the lower extremity. Individuals with spinal cord injury undergoing an 8-week training program with the Lokomat found no change in the Ashworth scale, though reductions in extensor spams were seen on the Spinal Cord Assessment Tools for Spasticity scale (97).

In summary, there is a strong evidence-based consensus recommending regular exercise to improve fitness and cardiovascular health for all ages and that it is even more acute for persons with chronic disability. This recommendation can be implemented employing many modalities, including robots programmed to challenge cardiovascular fitness. In fact, it appears that robot-assisted exercise may have a modest but measurable unique effect on spasticity, with the exact mechanism(s) yet to be fully elucidated. The effects observed to date have been primarily in the limb segments exercised, rather than in other upper-limb segments, suggesting that the effect is related to the performance of robot-assisted exercise rather than a physiological change, nonspecific effect of observation or measurement.

Neuroscience Considerations of Motor Neurorecovery

Since the evidence supporting the benefits of a cardiovascular fitness program is unequivocal, should robotic training optimize cardiovascular fitness? Common sequelae of neurological injury include spasticity, abnormal tone, disrupted or unbalanced sensory pathways, and muscular weakness. These deficits appear to involve the peripheral nervous system and might suggest that muscles should be the focus of therapy.

Should robotic training optimize muscle strength? The answer is likely more complex, and there is a significant body of evidence supporting an alternate assumption that robot-mediated therapy can also help patients “relearn” motor control. Though intuitively sensible, this notion may need to be refined, as normal motor learning does not have to contend with the neuromuscular abnormalities already discussed. Our results for upper extremity show that, although the process of recovery may share some features of motor learning (such as specificity), the relationship between learning and recovery may be subtle. While movement is beneficial, movement alone is not sufficient. Active involvement of the patient is important. Despite the fact that repetition may be beneficial, repetition alone is not sufficient; the benefits of robotic therapy do not exclusively derive from the high “dosage” of movement delivered. Our working model suggests that, in order to promote recovery, we need to include all of the above elements presented in a proper sequence of events.

ACKNOWLEDGMENTS

This work is supported by NICHD-NCMRR Grant no. 1 R01-HD045343, by the VA Veterans Affairs Grants no. B3688R, B3607R; the NYSCORE. A.C.L. is supported by grants from the Department of Veterans Affairs Rehabilitation Research and Development Service (B4145K and B54031). H.I.K and N.H are coinventors in the MIT-held patent for the robotic device used to treat patients. They hold equity positions in Interactive Motion Technologies, Inc., the company that manufactures this type of technology under license to MIT. Reha-Stim, Berlin, Germany, is owned by Dr. Beate Brandle-Hesse, the spouse of the author SH.

REFERENCES


Abductor pollicis longus (APL) tendons, 27
Abductor pollicis brevis, 23
Abduction, 956, 957
Abdominal muscles, 34
Abdomen, 13
Abandonment, 52, 1998

Abductor tendon, thumb, 1861–1862, 1862
ABLEDATA, 276, 382
See ABRP.

Above-knee amputation, 130–131
Abnormal gait, 131–132

Note: Page numbers followed by f indicate figure. Page numbers followed by t indicate table.

Activities of daily living (ADL), 381, 382
Actiq, 1769
Acetaminophen (Tylenol), 1743, 1745
Acetylcholine, 1793
Acetylsalicylic acid (ASA), 1025

Adductor magnus, longus, brevis, 35, 35f, 52f
Adductor pollicis, 23f, 32–34, 33f, 51f
Adhesive capsulitis, 913–914, 914f
Adjutant analgesic drug, 1765
Adjutant analgesic medications for chronic pain, 1298, 1299f, 1300

Adjuvant therapeutic modalities:
electrical stimulation (ES), 1366–1367
electromagnetic therapy, 1368
hydrotherapy, 1366
negative pressure wound therapy (NPWT), 1367
therapeutic ultrasound, 1368

ADL. See Activities of daily living

Adolescents: scoliosis, 899
with TBI, 587–588

Adult, childhood-onset disability
achondroplasia management, 1512–1513
cerebral palsy management, 1507–1508
definition of, 1503–1504
Medical Home Initiative components, 1504
neuromuscular disease management, 1510–1511
ostogenesis imperfecta management, 1511–1512
spina bifida management, 1508–1510
spinal cord injury management, 1505–1506
Title V, CSHCN program, 1503

Adult motor neuron disease. See Motor neuron disease (MND)

Adult motor neuron disease (MND)
amyotrophic lateral sclerosis epidemiology, 717–718
etiology, 717
clinical trials, 723–724
diagnostic evaluation
adult nonhereditary form, 722–723
clinical presentation, 719–721, 720f, 721f
differential diagnosis, 720f, 721–722
Kennedy’s disease, 722
neuroimaging, 723
spinal muscular atrophies, 722
pharmacologic management
antiapoptosis drugs, 725
anti-inflammatory drugs, 725
antioxidants, 725–726
cannabinoids, 726
growth factors, 725
neurodegeneration, 726
riluzole, 724–725
rehabilitation and palliative care, 726–736
Adult motor neuron disease (MND) *(Continued)*

spinal and spinobulbar muscular atrophy, 718f
adult-onset form, 718

Aerobic exercise

Aerobic capacity

cardiac output (CO), 1081–1082, 1082f
heart rate (HR), 1081, 1081f
myocardial oxygen consumption (MVO₂), 1082, 1082f
stroke volume (SV), 1081, 1082f

Aerobic exercise

Aerobic capacity assessment, 1638–1639
for low back, 993
physiologic adaptations, 1637–1638, 1637f
principles of, 1638, 1639f
Aerobic training, 772, 1083–1084, 1084f
Age effects, 1550
Ageism, 1558

Agency for Health Care Policy Research (AHCPR), 1713, 1732

Aging

active life expectancy, 1546–1547, 1547f
acute hospitalization effects, 1562
ageism, 1558
common geriatric syndromes
falls, 1573
incontinence, 1571
pain, 1572–1573
sleep disorders, 1571–1572
cumulative functional sequelae, 1559

Agnosia, 422

AIDS for daily living, 2001f
Air pollution, 1657
Akathisia, Parkinson’s disease, 658
Alcohol use disorders identification test (AUDIT), 400
Alcohol-related myopathies, 770–771
Alcohol-related myopathies, 770–771
Analgesic agents pharmacology

Acetaminophen, 1807–1808
COX II inhibitors, 1807
ketorolac, 1808
Saliclylates and propionic acid derivatives, 1805–1807
pharmacologic actions

cellular effects, 1810
CNS effects, 1810
peripheral effects, 1810
tolerance and physical dependence, 1810–1811
toxicity, 1811

American Board of Rehabilitation Psychology (ABRP), 388
American Board of Rehabilitation Psychology (ABRP), 388
American Federation for the Blind (AFB), 2005f
American Spinal Injury Association (ASIA), 2044–2045, 2044f, 2045f
anonymous, 1811
Anesthesia dissociative, 1760
Angina pectoris, 1090, 1090f
Angiotensin II receptor blockers, 1779, 1780f
Angiotensin-converting enzyme inhibitors (ACEIs), 1779, 1780f
Ankle

injuries

achilles tendinosis, 1463–1464
impingement, 1462–1463, 1463f, 1463f
nerve block techniques
dep deep peroneal nerve blockade, 1842–1843
inframascular nerve blockade, 1843
saphenous nerve blockade, 1843
superficial peroneal nerve blockade, 1842
sural nerve blockade, 1841–1842
tibial nerve blockade, 1841, 1842f

Ankle-brachial index, 1179, 1180f
Ankle-foot orthosis (AFO), 634, 2071–2072

Ankle-foot orthosis (AFO), 634, 2071–2072

Ankle-foot orthosis (AFO), 634, 2071–2072

Ankle-foot orthosis (AFO), 634, 2071–2072

Ankle-foot orthosis (AFO), 634, 2071–2072

Ankle-foot orthosis (AFO), 634, 2071–2072

Ankle-foot orthosis (AFO), 634, 2071–2072
Athletic shoe, 2068/
Atlantoaxial instability, 185
Atomoxetine (Strattera), 1793
Attention, TBI, 595–597, 596c
Audio feedback, 1939
Audiometric test interpretation, hearing loss
degree and configuration of, 436–437, 436c
types of, 435–436, 435f, 436f
AUDIT. See Alcohol use disorders identification test
Auditory brainstem response (ABR), 435
Auriculotherapy, 1728
Auditory function, 431
Avascular necrosis (AVN), HIV, 1236
Autonomic innervation, 1288
Auriculotherapy, 1728
Behaviocal assessment of the dysexecutive syndrome (BADS), 600
Bezoalzepines, 1326–1327
Bernhardt-Roth syndrome. See Meralgia paresthetica
Beta-blockers, 1779–1780, 1782
Biochemical pain theory, 1275–1276
Biofeedback (BF)
Botulinum toxin therapy
Brain imaging
Bladder
Blood-brain barrier (BBB), multiple sclerosis, 625
Body
Blood pressure
Bone
Botulinum toxin therapy
Bowling, physical activity, 1484
Bowel. See also Neurogenic bowel dysfunction
dysfunction, MS, 635–636
dysfunction with TBI, 594
evacuation mechanics
gastrointestinal problems after spinal cord injury, 1381–1382
medical and rehabilitation management, 1382–1384, 1382c
neurogenic bowel function, 1380–1381
rectodynamics, 1379–1380
Boston process approach (BPA), 399
Botulinum toxin therapy
clinical uses of, 1332–1333
side effects, 1333–1334
varieties of, 1335
Brain imaging
complications, injury, 198, 198c
contusions, 196, 197c
diffuse axonal injury, 196–197, 197c
epidural hematoma, 194–195, 195c
gray matter diseases, 200, 200c
head injury, 192–194
intraparenchymal hemorrhage, 196
penetrating trauma, 198
stroke
cerebral ischemia, 186–191, 186f, 190f
definition, 185
hemorrhagic, 191–192, 193f, 194f
subarachnoid and intraventricular hemorrhage, 192, 194f, 195f, subdural hematoma, 195–196, 196f, white matter diseases, 198–199, 199f, Brain lesions, HIV, 1237–1238, Brain mapping, stroke, 2156–2157, 2161, Brain tumors benign brain tumors, 1171, malignant brain tumors, 1171–1172, rehabilitation principles, 1172, Brainstem auditory evoked responses (BAERs), 583, Breast cancer, 1167–1168, Breathing chronic obstructive pulmonary disease (COPD), 1102, factors regulating, 192f, glossopharyngeal, 1115–1116, 1116f, 1116r, treatment, sleep-disordered, 1117, Broca's aphasia, 421, Bromocriptine (Parlodel), 1793, Brown-Sequard syndrome (BSS), 676, Burger's disease, 1192, Bupivacaine (Marcaine, Sensorcaine), 1757, Buprenorphine, 1770, Buprenorphine, 1770, Burn rehabilitation acute medical management burn wounds, acute care, 1128, dysphagia, 1129, grafting, 1128–1129, 1129f, bone and joint changes bone growth, 1137, contractures, 1136, 1136f, in electrical burns, 1137, heterotopic ossification, 1137, osteophytes, 1137, scoliosis and kyphosis, 1137–1138, septic arthritis, 1138, subluxations and dislocations, 1138, complications, 1134, epidemiology, 1125, 1126f, exercise, 1132–1133, facial burns, 1144–1145, 1145f, gait, 1133, geriatric burns, 1144, hands, 1132, 1132f, metabolic complications catabolic state, 1141, temperature regulation, 1141, neurologic injuries localized, 1134–1135, 1135r, mononeuropathies, 1135, peripheral polyneuropathy, 1135, pruritus, 1135–1136, pain complications, 1138–1139, pathophysiology classification of, 1125–1126, 1126r, normal skin, 1125, 1126f, skin regeneration and scarring, 1128, thermal injury effects, 1127–1128, pediatric burns, 1143–1144, positioning, 1130, 1130f, 1131r, psychosocial issues community integration, 1142–1143, depression, 1142, posttraumatic stress, 1142, sleep disturbances, 1142, skin complications blisters, 1140, hypertrophic scarring, 1139–1140, 1139r, malignancy, 1141, UV sensitivity and skin pigmentation, 1140–1141, splitting, 1130–1132, 1131f, surgical reconstruction, 1133–1134, Burns. See Burn rehabilitation BWSTT. See Body weight-supported treadmill training Byetta, 1789–1790, C, CABG. See Post-coronary artery bypass grafting Cachexia epidemiology, 1918, pathophysiology, 1918–1919, systemic interactions, 1919f, treatment, 1919–1920, CAGE questionnaire, 400, Calcium channel blockers, 1780–1781, 1782, Calcium channel blockers, 1782, Cardiac rehabilitation acute medical management cancer-related fatigue (CRF), 1158–1159, 1158r, demographics of, 1151, 1152s–1153r, exercise prescription, 1159–1160, 1660, fracture risk, 1157, 1157r, Karnofsky scale, 1153, 1153r, lymphedema, 1163–1165, 1164f, 1164r, National Comprehensive Cancer Network (NCCN), 1158, neurologic complications metastatic brain disease, 1160–1161, paraneoplastic neuromuscular disorders (PNND), 1161, polyneuropathy, 1161–1162, 1162r, spinal cord involvement, 1161, nutrition, 1165–1166, pain, 1154, nonpharmacologic pain management approaches, 1156, opioid strategies, 1155–1156, 1155s, radiation therapy (RT), 1162–1163, 1162r, radiation-induced tissue damage, 1162–1163, 1162r, rehabilitation, 1151–1153, 1153r, consultation during acute care, 1154, expectations, 1151, historical background, 1151, inpatient, 1153, outpatients, 1153–1154, precautions, 1154, sexual dysfunction, 1166, tumor necrosis factor (TNF), 1160, tumors brain tumors, 1171–1172, breast cancer, 1167–1168, gastrointestinal malignancies, 1171, head and neck cancer, 1168–1170, 1168f, 1169f, hematologic malignancies, 1170–1171, lung cancer, 1171, prostate cancer, 1173–1174, sarcomas of bone and soft tissue, 1172–1173, Cancer pain, 1304–1305, Cancer-related fatigue (CRF), 1158–1159, 1158r, Candidiasis, 1233, Capasico (Zostrix), 1773–1774, Carbamazepine (Tegretol), 1751–1753, 1753r, Carbipoda, 1793, Cardiac abnormalities, with TBI, 590, Cardiac arrhythmias, 1093, Cardiac complications management, 772–773, Cardiac disease exercise prescription, 1658, 1658r, prevalence and incidence, 1075, types of, 1075, Cardiac disease, stroke, 568–569, Cardiac glycosides, 1781, 1783, Cardiac output (CO), 1081–1082, 1082f, Cardiac rehabilitation abnormal physiology, 1084–1085, 1085r, aerobic capacity cardiac output (CO), 1081–1082, 1082f, heart rate (HR), 1081, 1081f, myocardial oxygen consumption (MVO2), 1082, 1082f, stroke volume (SV), 1081, 1082f, cardiac rehabilitation training, 1083–1084, 1084r, anatomy of, 1078–1079, 1079f, 1079r, 1080r, heart disease epidemiology, 1075, metabolic equivalents (METs), 1083, 1084r, physiology of, 1079–1081, 1080f, post-myocardial infarction (MI) patient acute Phase (phase I), 1086, 1086r, inpatient rehabilitation phase (phase IB), 1086–1087, 1087r, maintenance phase (phase III), 1089 training phase (phase II), 1087, 1088r, 1089r, primary prevention, 1076, 1076r, rehabilitation programs, 1089, angina pectoris, 1090, 1090f, cardiac arrhythmias, 1093, cardiac transplant surgery, 1091–1092, cardiomyopathy, 1092–1093, physically disabled, 1093–1094, 1094r, post-coronary artery bypass grafting (CABG), 1090–1091, 1090r, post-percutaneous transluminal coronary angioplasty (PTCA), 1091, valvular heart disease, 1093, risk factors cigarette smoking, 1078, diabetes, 1076, hypercholesterolemia, 1077, hypertension, 1077, metabolic syndrome, 1077–1078, 1077r, obesity, 1078, Cardiac transplant surgery, 1091–1092, Cardiomyopathy, 1092–1093, Cardiorespiratory fitness, 1443, 1443r.
Cerebral palsy, disabled children.
Cerebral palsy (CP)
Cerebral palsy, 1941–1942
Cell-based therapies, stroke, 2159–2160
Cell therapy, regenerative medicine, 2173
Caudal lumbar epidural injection, 1880–1881,
Catecholamines, 1791
CAT techniques, Computer adaptive testing
Cartilage, physiology, 1415
Cervical disc
Cervical cord neuropraxia, sports, 1420–1421,
Cerebral perfusion pressure (CPP), 579
Carpal tunnel syndrome (CTS), 929–930,
Cardiovascular risk, 1764–1765
Cardiovascular medications, 1778–1779
Cardiovascular and peripheral hemodynamic
quality of movement, 1482–1483, 1482
prognosis for ambulation, 1484
prenatal and postnatal period, 1482
mental retardation, 1484
malnutrition, 1485
gross motor delay, 1482
differential diagnosis, 1483
dental problems, 1484–1485
classification, 1483
behavior disorders, 1485
abnormal vision, 1484
behavior disorders, 1485
classification, 1483
dental problems, 1484–1485
differential diagnosis, 1483
gross motor delay, 1482
hearing loss, 1484
hemiplegia and diplegia, 1483
laboratory examination, 1483
malnutrition, 1485
management therapies, 1485–1486
mental retardation, 1484
prenatal and postnatal period, 1482
prognosis for ambulation, 1484
quality of movement, 1482–1483, 1482/
reflex abnormalities, 1483
respiratory infections, 1485
seizures, 1484
spasticity medications, 1486–1577
surgical intervention, 1487–1488, 1487/
triplegia and quadriplegia, 1484
Cerebral perfusion pressure (CPP), 579
Cervical cord neuropraxia, sports, 1420–1421,
1420f/1421f
Cervical disc
herniation, 172–173, 174f/175f
injury, sports, 1419–1420
Cervical myelopathy, childhood-onset disability, 1508
Cervical orthoses (COS), 2082r
biomechanics, 2082–2083
cervicotorticorhaphy orthosis (CTO), 2085–2086
halo orthosis, 2086–2088, 2086f/
hard cervical collars, 2083–2085, 2085f/
soft cervical collars, 2083, 2083f/
Cervical spine disorders
anatomy and biomechanics, 811–813,
812f/813f
axial pain and symptom referral
epidemiology, 813–814
intervertebral disc, 814, 814f/
and radicular pain, 826–830, 827f/830f/
yggyphopyleal joints, 814–815, 815f/
collars, 828–829
manipulation, 828
surgical therapy, 827–828, 827f/
mobilization, 828
modalities, 829
myelopathy
history of, 823–824
natural history, 824
pathophysiology, 821–822, 822f/
physical examination, 824
radiologic findings, 822–823, 823f/
surgical treatment, 824–825
physical therapy, 830–831
radiouleapathy
electrodiagnostics, 817–818
epidemiology, 815
history of, 816, 818f/819f
nonsurgical treatment, 820–821, 820f/
pathophysiology, 815–816, 816f/
physical examination, 818–819
radiologic findings, 816–817, 817f/
surgical treatment; 821
rheumatoid arthritis, 825–826
stabilization, 829–830
traction, 829
work ergonomics, 830, 830f/
Cervical spine injury, sports, 1421
biomechanics, 1421
clinical presentation, 1420
classification, 1420
mechanisms of action, 1331–1332, 1331f/
with phenol and alcohol, 1332
Chemical use assessment, 399–400
Chemosensor trigger zone (CZT), 1914
Chord-wall mobility, 1110
Chiari malformation
disabled children, 1496–1497
Children
amputation, 2044–2045, 2044r, 2045r/
with TBI, 587–588
Children and youth with special health care
needs (CYSHCN), 1503
Choke syndrome, amputation, 2043
Chondromacia patella, MRI, 159, 160f/
Chorea, Parkinson's disease, 658
Chronic alveolar hyperventilation, 1109, 1109r
Chronic inflammatory demyelinating polyneuropathy (CIDP)
HIV, 1238
Chronic LBP (CLBP)
clinical presentation, 863
definition and pathogenesis, 853–855
diagnosis, classification, and prognosis,
855–856
rehabilitative approach, 845, 857–859
treatment approaches, 856–857
Chronic obstructive pulmonary disease
(COPD)
airway secretion elimination, 1102–1103,
1103f/
breathing retraining, 1102
counseling and general medical care,
1101–1102
inspiratory resistive exercises, 1104
medications, 1100–1101
nocturnal negative pressure body ventilator
(NPBV), 1104
nutrition, 1102
patient evaluation, 1099–1100, 1100r
physical aids, 1107
positive expiratory pressure (PEP),
1102–1103
pulmonary rehabilitation, 1107–1108
reconditioning exercise, 1106–1107
respiratory muscle rest, 1104–1105
sample therapeutic prescription, 1100, 1101r
supplemental oxygen therapy, 1105–1106
upper extremity exercise, 1107
Chronic pain
anesthetic procedures, 1300–1301
behavioral treatment of, 1295
clinical research in, 1301
complications of
environmental, 1278
physical, 1278
psychological, 1278
diagnostic and clinical evaluation
history, 1288–1289
physical examination, 1289
diagnostic procedures
anesthetic, 1290
electrodiagnosis, 1290
imaging, 1289
Indicators for, 1289f/
laboratory testing, 1289
psychological evaluation, 1289–1290
epidemiology of, 1273
etiology of, 1274–1278
multidisciplinary approach, 1290–1291
pain pathways, 1274–1275, 1274f, 1275f
pharmacologic intervention for, 1295–1300
physical modalities, 1291, 1293–1295
resolution of, 1277
socioeconomic factors in
litigation, 1278
worker’s compensation, 1278

Clinical evaluation
individually muscle strength test
genitourinary symptoms, 9

comprehensiveness of, 9
dermatologic symptoms, 10
endocrine symptoms, 10
constitutional symptoms, 9

socioeconomic factors in
resolution of, 1277
pharmacologic intervention for, 1295–1300
pain pathways, 1274–1275, 1274f
etiology of, 1274–1278
cardiovascular symptoms, 9
syndromes, 1301–1309

Clinical practice, ethics
social history, patient profile
setting and purpose, 4–5
respiratory symptoms, 9

problem list, 46–47, 46f
pronator teres, 23f, 29–30, 29f, 50f
quadriceps femoris, 37, 37f, 51f
rhomboids, 23, 23f
serratus anterior, 23–24, 23f, 24f
subscapularis, 25–26, 26f, 49f
supinator, 27, 27f, 50f
supraspinatus, 23f, 24, 24f, 48f
teres major, 23f, 25, 26f, 49f
trapezius, 22–23, 22f
triceps, 23f, 26–27, 27f, 49f
interdisciplinary nature of, 4
musculoskeletal symptoms, 9–10

psychological and psychiatric history, 10
symptoms, 5, 5f
illness, 5, 5r

comprehensive, 9
comprehensive examination, 46–47

clinical implications, 482–483
Clinical practice, ethics
duties and rights, family members, 507
goal setting, 504–505

patient-practitioner relationships
competency and recommendations, 505–506
contractual model, 505
patients selection, 503–504
policy dilemmas, 503, 504r
professional and team issues
moral problems, 506
recommendations, 506–507
quality of life, 507–508
termination of treatment, 507–508

Clinical practice standards
medicolegal evaluations

CMS. See Centers for Medicare and Medicaid Services
CMV. See Cytomegalovirus
CNAPs. See Compound nerve action potentials
CO. See Cardiac output
Coccygeal junction, injection technique, 1864–1865, 1865f
Coccyx injections, 1898
Codeine, 1768
Cognitive behavioral approach, 847–848, 858–859
Cognitive dysfunction, HIV, 1237
Cognitive rehabilitation, TBI, 605
Collagen, 1258
Coma with TBI, 575
Communication

disabled children, 1479

Communication impairments

after right hemisphere damage, 423–424
after traumatic brain injury (TBI) assessment, 425
rancho los amigos scale—revised, 424–425
rehabilitation team, 425–426
TBI, 600–601
Communication pathology, 413–414
Community participation and environment

Americans with Disabilities Act (ADA), 482
assessment of life habits (LIFE-H), 477
clinical implications, 482–483
Crag handicap assessment and reporting technique (CHART), 477
depictions, 473–474, 474f

late-life disability instrument (LLDI), 477
participation survey/mobility (PARTS/M), 477
scientific literature linking, 480–482, 481f
theoretical assessment

environment layout, 478–479
Index

Community participation and environment (Continued)
environment-participation interaction, 475–476, 476f, 477f
Global positioning systems (GPSs), 478–479
measurement considerations, 476–478, 478f
observational approach, 479
participation, 474–475, 475f
self-report assessment, 479–480
Community-based physical activity, 1437, 1438f
personalized exercise program (PEP), 1458–1440, 1439f
tailoring programs, 1437–1438
types of, 1440–1441, 1440f
Complementary and alternative medicine, 2120
biological therapies
anti-inflammatory botanicals, 2126–2127
anti-oxidants, 2126
avo-cado/soybean unsaponifiables, 2126
S-adenosylmethionine (SAMe), 2126
CAM techniques, uses, 2130
energy therapies
prayer and spiritual healing, 2128
Qigong, 2128
Reiki, 2128–2129
therapeutic touch (TT), 2128
functional medicine, 2129–2130
manipulative and body-based therapies craniosacral therapy, 2127
roling, 2127
tragery and psychophysical integration, 2127–2128
mind-body interventions
hypnotherapy, 2122
meditation, 2122
mind-body movement therapies, 2122–2124, 2123f
sports medicine, 2129
whole medical systems
acupuncture, 2120–2121
ayurveda, 2121–2122
homeopathy, 2121
naturopathy, 2122
oriental medicine, 2119–2120
Complex regional pain syndrome, 947–948, 948f, 1302–1303
Compound muscle action potentials (CMAPs), 92, 92f
Compound nerve action potentials (CNAPs), 90–92
Compression stockings
for posterior tibialis tendinitis, 957–958
Computed tomography (CT)
spinal cord, degenerative disorders atlantoaxial instability, 185
lumbar stenosis, 178–180, 179f, 180f
spondylolisthesis, 180–181, 180f–182f
trauma, 181–185, 183f–185f
Computer adaptive testing (CAT) techniques, 236
Concussion, sports, 1419
athlete evaluation, 1418
paper-and-pencil testing, 1419
postconcussive symptoms, 1418
return to play, 1418, 1419
scales, 1418
seizures, 1418
signs and symptoms, 1418–1419
traumatic brain injury, 1418–1419
Conduction aphasia, 422
Conflicts, team function, 363
Congenital amputation, incidence of, 2017–2018
Congenital myopathies, 769
Congenital scoliosis, 893–894, 896
Connective tissue
disease changes in, 1259
mechanical properties of, 1258
Connective tissue diseases, 452–453
Constraint-Induced Movement Training ( CIMT ), stroke, 566
Consumerism, consumer movement, 465
measurement considerations, 476–478, 478f
Global positioning systems (GPSs), 1181–1182, 1181f
for posterior tibialis tendonitis, 957–958
Consumerism, consumer movement, 465
measurement considerations, 476–478, 478f
Global positioning systems (GPSs), 1181–1182, 1181f
for posterior tibialis tendonitis, 957–958
Continuous wave doppler, 1180, 1181
Continuous passive motion, 2192
Continuing disability review ( CDR ), 292
Continuous passive motion, 2192
Continuous wave doppler, 1180, 1181f
Contractions
anatomical classification of, 1259f
management of, 1261–1263
physical function limitations and, 1261
prevention and management of, 774–775
prevention and treatment of, 1262f
with TBI, 592
Contralaterally controlled functional electrical stimulation ( CCFES ), 1956f
Contrast arteriography, 1182
Contrast baths, hydrotherapy, 1694
Contrast venography, 1184
Copaxone.
Corticosteroids, 1210, 1746, 1748–1750, 1750f
Cortical activity, stroke injury, 2157–2158,
2158f
Corticosteroids, 1210, 1746, 1748–1750, 1749f, 1750f
atrophy, 770
injection procedure, 1852–1853
Cosmetic prosthesis, 2041
Costochondral junction, injection technique, 1857, 1856f
Cough, 1923–1925
COX II inhibitors, 1764–1765
COX II inhibitors, 1764–1765, 1765f
Cyclosporine, 1027–1028, 1028f
Cyclooxygenase-II inhibitors, 1764–1765, 1765f
Cyclopentolate, 1027–1028, 1028f, 1207–1208
Cyclosporine, 1027–1028, 1028f
Cyclosporine, 1027–1028, 1028f
Cyclosporine, 1027–1028, 1028f
Cyclosporine, 1027–1028, 1028f
Curve angle, 889–890, 889f
Cryotherapy

disease models, 1560–1562, 1561f
underreporting of illness, 1560
variable patterns of, 1560
Cumulative trauma disorders (CTDs), 1302
de Quervain’s tenosynovitis, 928–929
epicondylitis, 927–928
epidemiology and economic costs, 923
ergonomics, 924
flexor carpi ulnaris (FCU), 931
flexor tenosynovitis, 928
forceful exertions, 924
hand-arm vibration syndrome, 929
neck, 925–926
nerve entrapment syndromes
carpal tunnel syndrome (CTS), 929–930
eronomic interventions, 933
low back pain (LBP), 932
pronator syndrome, 930–931
thoracic outlet syndrome (TOS), 932
ulnar nerve entrapment, 931–932
pathophysiology, 923
posture, 924
psychosocial, 923–924
repetition or prolonged activities, 924
shoulder, 926–927
special diagnostic tests, 924, 925f
trigger finger, 928
vs. various neuromusculoskeletal conditions, 924, 925f
vibration, 924
vibration white finger, 929
Curvature, 889–890, 889f
Cycling, recreational sports, 1454, 1454f
Cyclooxygenase-2 (COX-2) inhibitors, 795, 796
Cyclooxygenase-II inhibitors, 1764–1765, 1765f
Cyclosporine, 1027–1028, 1028f, 1207–1208
Cytomegalovirus (CMV), 1212
Dance medicine
ankle injuries, 1462–1464, 1463f, 1463f
causes of, 1459–1461, 1461f
dance training, 1459
Dance medicine
ankle injuries, 1462–1464, 1463f, 1463f
causes of, 1459–1461, 1461f
dance training, 1459
Dance medicine
ankle injuries, 1462–1464, 1463f, 1463f
causes of, 1459–1461, 1461f
dance training, 1459
Dance medicine
ankle injuries, 1462–1464, 1463f, 1463f
causes of, 1459–1461, 1461f
dance training, 1459
Dance medicine
ankle injuries, 1462–1464, 1463f, 1463f
causes of, 1459–1461, 1461f
dance training, 1459

INDEX
femoral neuropathy, 108
median neuropathy at wrist, 104–106, 105f
radial nerve at humerus, 107–108
tibial nerve, 108–109
ulnar neuropathy at elbow, 106–107

EPB. See Extensor pollicis brevis

Epicondyle injection, elbow
lateral, 1858, 1858f
medial, 1858–1859
olecranon bursa, 1859, 1859f
radioulnar joint, 1859, 1860f

Epicondylitis, elbow injury, 1425

Ergonomics, assistive technology, 2007–2008

Ethics

Estrogen

Electrical stimulation

Evidence-based practice, in rehabilitation

Exenatide (Byetta), 1789–1790
Evista, osteoporosis, 999

Extrinsic/extraintestinal neural regulation,

Extrinsic/extraintestinal neural regulation,

Extrinsic/extraintestinal neural regulation,

Extrinsic/extraintestinal neural regulation,

Extrinsic/extraintestinal neural regulation,

Extrinsic/extraintestinal neural regulation,

Extrinsic/extraintestinal neural regulation,

Extrinsic/extraintestinal neural regulation,

Extrinsic/extraintestinal neural regulation,

Extrinsic/extraintestinal neural regulation,

Extrinsic/extraintestinal neural regulation,

Extrinsic/extraintestinal neural regulation,

Extrinsic/extraintestinal neural regulation,

Extrinsic/extraintestinal neural regulation,

Extrinsic/extraintestinal neural regulation,

Extrinsic/extraintestinal neural regulation,

Extrinsic/extraintestinal neural regulation,

Extrinsic/extraintestinal neural regulation,

Extrinsic/extraintestinal neural regulation,

Extrinsic/extraintestinal neural regulation,

Extrinsic/extraintestinal neural regulation,

Extrinsic/extraintestinal neural regulation,
Hip (Continued)
joint injection techniques, 1865, 1865f
MRI
femoroacetabular impingement, 153, 153f
giant synovial cyst, 154, 154f
ischemic necrosis, 152, 153f
thigh splits, 154, 154f
ostearthritis, 785, 785f
osteoarthritis, 1005–1006
sports medicine, 1429–1430
Hip flexion assist orthosis (HFAO), 634, 634f
Hip replacement, 1033
Hip-knee-ankle-foot orthoses (HKAFOs), 2074–2075, 2074f, 2075f
History taking. See Clinical evaluation
HIV. See Human immunodeficiency virus
Homeopathy, 2121
Hormonal disorders, 1269
Hormone replacement therapy (HRT), postmenopausal women, 1533, 1606
Hot tubs, 1685
HR. See Heart rate
HRNB. See Halstead-Reitan neuropsychological battery
HRQOL. See Health Related QOL
Hubbard tank, 1693–1694, 1693f
Human factors, assistive technology, 1998
Human immunodeficiency virus (HIV) associated with myopathies, 771
ATHOS report, 1227
classification, 1228, 1228r
clinical evaluation, 1229
cognitive dysfunction, 1237
demographics, 1228
focal brain lesions, 1237–1238
functional impairment syndromes, 1234f–1235f
Gastrointestinal disease, 1233
neuropathies, 1236–1237
occupational exposure, 1228
pain syndromes, 1236–1237
pathogenesis, 1227–1228
pediatric, 1239
psychosocial issues, 1240
radiculopathies, 1238–1239
respiratory infections, 1232–1233
swallowing disorders, 1253
transmission routes, 1228
treatment strategies, 1229–1232
vaccular myelopathy, 1239
vocational issues, 1240
weakness and fatigue, 1236
Human muscle function assessment
force generation, 69
and functional performance, 83
functional tests
chair stand time, 82
habitual and maximal walking speed, 83
6-minute walk test, 83
stair-climbing power, 82
muscle actions and units
absolute and relative endurance, 75–76, 76f
energetics of, 71–72
2f7f
functional characteristics of, 72–74
72f
International System of Units, 70, 70f
iso kinetic dynamosometer, 70, 70f
muscle endurance and fatigue, 75
muscle power, 74–75
open kinetic chain and closed kinetic chain, 71
physical and biomechanical concepts, 70, 70f
stretch-shortening cycle, 70
single muscle fibers, 81–82
2f
structural characteristics, skeletal muscle fiber size and type distribution, 80–81, 81f, 81f
function, 77, 77f
muscle biopsy techniques, 79–80
muscle size, 78–79
78f, 79f
structure, 77, 77f
Huntington’s chorea, 200
Hyalgan, 801
Hybrid orthosis, 2071
Hydroalg, 1819
Hydralazine, 691
Hydration, palliative care, 1925–1926
Hydralazine, 691
Hydeltra, 1819, 1819f
Hybrid orthosis, 2071
Hydrotherapy, 1366
balneotherapy, 1364–1365
contrast baths, 1694
dexamethasone, 1694
fluidotherapy, 1695
paraffin baths, 1695–1696, 1695f
safety, 1692f, 1695
sitz baths, 1694
water-based exercise, 1694
Whirlpools and Hubbard tanks, 1693–1694, 1693f
Hydroxychloroquine (HCQ), 1027
Hyperbaric oxygenation, for infectious peripheral neuropathy, 748
Hypercalciumia, 1264
Hyperhidrosis, 1332, 2043
Hyperkinetic dysarthria, 416
Hyperlymphoposisis, 873
Hypermethylidympose, stroke, 552
Hyperreflexia, 204
Hyperthermia
cardiac rehabilitation, 1081, 1082f
stroke, 552
Hypertrophic cardiomyopathy (HCM), sports injury, 1416–1417
Hypokalemia, 771
Hypoproteinemia, 1552
Hyporeflexia
aging, 1569
Parkinson’s disease, 650
spinal cord injury, 689–690
Hypophosphataemic dysfunction, with TBI, 390–391
Hyporeninemic hammer syndrome, 1193
Hypoventilation, alveolar, 1109–1110
Hypoventilation, alveolar, 1109–1110
Hypoventilation, alveolar, 1109–1110
Inpatient rehabilitation facility (IRF) cases, 485, 486, 487
Inspiratory resistive exercises, 1104
Instability syndrome, 1554
Institute of Medicine report, Crossing the Quality Chasm, 2137
Instrumental activities of daily living (IADL), with ADL
COPM, 260–261
functional independence measure (FIM), 258–260, 259
AMPS, 256–257, 256
FAL, 258
NEADL, 257–258, 257
Insulin, 1788–1789, 1789
Integrative care pathways (ICPs), 341, 579
Integrative field theories, 404
Integrative medicine, 2119
Intellectual ability tests, 394
Intellikeys keyboard, 2000f
Intellikeys keyboard, 2000f
Intercalary limb loss, 2018
Intercostal muscles, 34
Intercostal nerve blockade, 1820, 1827–1828, 1829
Interdisciplinary team, referrals and orders, 374–375
Interface pressure mapping, 1362
Interferons, 629–630
Interlaminar epidural injection, 1876–1877, 1877f, 1878f
Intermittent pneumatic compression, 251, 252
Interruption of passive pressure ventilation (IPPV), 1111–1114, 1112f, 1113f, 1113r
Internal carotid artery syndromes, stroke, 555
International Classification for Impairment, Disability, and Handicap (ICIDH), 248, 248f
International Classification of Functioning, Disability and Health (ICF), 289–290, 290f, 471
assessment and evaluation, 314, 314f
b130 energy and drive functions, 311, 312r
checklist, 305, 306f
classification and measurement instruments
dehalfreling buildings and robotic units, 304–305measurement instruments, 308–310, 309r, 310r
practical tools, 305–308, 306r, 307f, 307r
principles of, 311–315, 311f, 311f, 312f, 313f, 314r
validity of, 315
core sets
acute hospital and post-acute rehabilitation facilities, 305–306
chronic conditions, 307
conceptual approach, 305
development process, 307–308, 308f
generic ICF core set, 307, 307f
ICF core set project, 305
development of, 302
historical perspective, 301
implementation and application
classification of, 316–317, 317f, 318r
health sciences, 316
human functioning and rehabilitation research, 316f, 317
management, 317, 317f
official implementation, 316
planning and reporting, 318–320
measurement instruments
applications, 308, 309f–310r
linkage methodology, 308–310, 310r
psychological aspects, 388
RASCH scale, 312, 313f
structure of
activities and participation, 303, 303f
body functions and structures, 302–303, 303f
<table>
<thead>
<tr>
<th>Index</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>K</strong></td>
<td>102, 103</td>
</tr>
</tbody>
</table>
| Joint Commission on the Accreditation of Health care Organizations (JCAHO), 366, 2173 Joint contractures. See Contractures Joint laxity disorders, 1469–1470, 1471f Jones Act, 293 Jones fracture, 964 Jones’ fracture, 963–964, 965f Jumper’s knee, 157 Juvenile chronic arthritis, 1061 Juvenile idiopathic arthritis (JIA), 1061–1064, 1062 Juvenile scoliosis, 899 Knee disarticulation, prostheses, 2029 injection techniques aseptic bursa, 1867, 1868f indications, 1866, 1867f metatarsal joint, 1871, 1871f metatarsophalangeal, 1870–1871, 1870f plantar heel fascia, 1870, 1870f retrocalcaneal bursa joint, 1869, 1869f subtalar joint, 1868–1869, 1869f tibial collateral joint, 1867–1868, 1866f MRL, 154–155 chondromalacia patella, 159, 160f collateral ligament injury, 157, 159f criscuate ligament injury, 155–157, 157f–159f meniscal injuries, 155, 155f, 156f osteochondritis dissecans, 158, 160f patellar tendinitis, 157, 159f popliteal cysts, 159–160, 161f–162f nerve block techniques common peroneal nerve blockade, 1839 saphenous nerve blockade, 1839–1840, 1841f tibial nerve blockade, 1839, 1840f pathology, ultrasound knee effusion, 5, 6f muscular injury, 5–6, 7f patellar tendinitis, 5, 7f posterior cruciate ligament (PCL) injury, 6, 8f physical examination biomechanics of, 66–67 inspection, 63 ligament stability anterior cruciate ligament (ACL), 65, 65f collateral ligaments, 64, 65f medial and lateral menisci, 65–66, 66f neurological manual muscle testing, 64, 64r reflex testing, 64, 6hr sensation testing, 64, 65r palpation, 63–64, 63f range of motion, 64 prostheses, 2033r sports medicine, 1430–1432 Knee extension-flexion isokinetic test, 1628 Knee injuries, 1462 Knee orthoses (KO), 2073–2074, 2074f Knee-ankle-foot orthoses (KAFOs), 2072–2073, 2072f, 2073f cerebral palsy, 1485 chronic myopathic disorders, 775 hip, 2074–2075 management, 2072–2073 plastic, 2073f SCI, 685 Knight brace, 2091, 2092f Knight-Taylor brace, 2090, 2090f Knuckle-bender orthosis, 2056, 2057f Kripalu yoga, 2123r KS. See Kaposis sarcoma Kugelberg-Welander disease (KWD), 1490 spinobulbar muscular atrophy, 718 kyphoplasty, 1003–1005 kyphosis cervical, 825 gibbus deformity, 894 hyperkyphosis, 873 spinal curves, 884f vertebral fracture, rehabilitation, 1002 l L1–L5 level, SCI, 685 Labor and delivery, 1526 Labral tears, 913 Lachman’s maneuver, 64–65 Lactic acid, energy source, 1619 Lacunar infarct, 187, 189f, 190f, 556–557 Lacunar stroke, 556–557 Lambert-Eaton myasthenic syndrome (LEMS), 102, 103 Lamicital, 1754 Laminoforaminotomy, 821 Lamotrigine (Lamictal, Lamictal CD), 1754 Language, 419 agnosia, 422 aphasia assessments, 422
Lumbosacral orthosis (LSO), 1003
Lumbosacral spine, 2089–2090
Luminal, 1754
Lung cancer, 1171
Lung transplantation, 1216–1217
Luchka’s joints, 174
Lymphangiography, 1184
Lymphatic disease
chronic lymphedema, 1198
clinical presentation, 1198–1199
lymphatic system, 1197–1198
lymphedema classification, 1198, 1198r
lymphedema treatment, 1199f
antibiotic therapy, 1199–1200
benzopyrones, 1200
compression, 1200–1201, 1200f
diuretic agents, 1199
elevation, 1201
exercise, 1201
skin care, 1200
specialized massage, 1200
surgery, 1201
vasopneumatic compression therapy, 1201
Lymphatic imaging, 1184
Lymphedema, 1163–1165, 1164f, 1164r
chronic, 1198
treatment, 1199f
antibiotic therapy, 1199–1200
benzopyrones, 1200
compression, 1200–1201, 1200f
diuretic agents, 1199
elevation, 1201
exercise, 1201
skin care, 1200
specialized massage, 1200
surgery, 1201
vasopneumatic compression therapy, 1201
Lymphoma, HIV, 1238
Lymphoscintigraphy, 1184
Lyrica, 1755

M
Magnetic resonance imaging (MRI), 201–202, 201f–202f, 1182
Magnetic resonance imaging (MRI)
ankle
ligament injuries, 160–163, 162f–164f
tenosynovitis, 164–165, 165f
brain (or Brain imaging)
elbow, 146–148, 147f–149f
hip
femoroacetabular impingement, 153, 153f
giant synovial cyst, 154, 154f
ischemic necrosis, 152, 153f
thigh splints, 154, 154f
knee, 154–155
chondromalacia patella, 159, 160f
collateral ligament injury, 157, 159f
cruciate ligament injury, 155–157, 157f–159f
meniscal injuries, 155, 155f, 156f
osteoarthritis, 158, 160f
patellar tendinitis, 157, 159f
popliteal cysts, 159–160, 163f–162f
musculoskeletal system, 139–141, 140f
shoulder
impingement syndrome, 142–144, 143f–144f
instability and labral tears, 144–145, 145f
ischemic necrosis, humeral head, 146
tendinitis, 145–146, 146f
spinal cord, degenerative disorders
atlantoaxial instability, 185
cervical disc herniation, 172–173, 174f, 175f
cervical stenosis, 174–176, 175f–177f
disc, 169–172, 173f
lumbar disc herniation, 176–178, 177f, 178f
spine, 168–169, 169f–172f
spondylolysis, 180–181, 180f–182f
thoracic disc herniation, 176
trauma, 181–185, 183f–185f
venous testing, 1184
wrist, 148–149, 149f
carpal tunnel syndrome, 148–150, 150f–151f
triangular fibrocartilage complex evaluation, 150–151, 150f–151f
Magnetic resonance spectroscopy, 203–204
Major depressive disorder (MDD), 393
Malignant brain tumors, 1171–1172
Malleoli, 954
Malnutrition, 1485
Manual muscle testing (MMT)
Manipulation
adverse effects, 1722
articular technique, 1715
cervical and radicular pain, 828
complications, 1723–1724
contraindications, 1723, 1723r
countertrain, 1715–1716
craniosacral treatment, 1717
definition, 1713–1714
duration and frequency, 1722–1723
functional techniques, 1716
indications, 1722
joint play, 1716
mechanism of action, 1717–1719
modes of practice, 1721
myofascial release method, 1716
patient selection, 1720–1721, 1721r
physiatrist applicability, 1720
prescription, 1724–1725
procedural fundamentals, 1725
referral
to licensed nonphysician, 1721–1722
to physician, 1721
research, 1725
risks and benefits, 1719–1720
skills acquisition, 1722
soft-tissue method, 1716
thrust techniques, 1714–1715, 1715f
treatment concept, 1719
techniques, 1714–1715, 1715f
Manpower and body-based therapies craniosacral therapy, 2127
rolling, 2127
trigger psychophysical integration, 2127–2128
Manual muscle testing (MMT)
cervical spine, 56, 56r
Knee, 64, 64r
lumbar spine, 62, 62r
skeletal muscle, 73, 73r
Manual therapy, 790, 849
MAOIs. See Monoamine oxidase inhibitors
Marcaine, 1757
Marfan syndrome, 896
Martin-Gruber anastomosis, 106–107
Massage
cervical and radicular pain, 830
contraindications, 1730
definition, 1726
eastern massage, 1727–1728
efficacy, 1730
indications, 1729–1730
physiatric application, 1730
physiologic effects, mechanical, 1728–1729
psychological effects, 1729
reflexive changes, 1729
variables, 1728
western procedures, 1726–1727
MAST. See Michigan alcohol screening test
Masters and Johnson (M&J) model, 445
Maximal muscular force, 1626f
Maximal voluntary contraction, 73
Maximum medical improvement (MMI), 294–295
Mayo-Portland Adaptability Inventory (MPAI), 582
May-Thurner syndrome, 1197
MCP joint, 946
MDD. See Major depressive disorder
Mechanical insufflation-exsufflation (MI-E), 1114–1115
Mechanical therapy
cervical and radicular pain, 827–828, 827f
Medial collateral ligament (MCL), 157, 159f
Medial epicondylitis, 919
Median nerve, 749
Medicaid, 490–491, 490f
Medical insurance, 509–510
Medical models, referrals and orders, 373–374, 374r
Medical Outcomes Study Short-Form 36 (SF-36), 312, 313
Medical practice, medicolegal system, 521–523
Medical practice, Medicare, 521–523
Medical rehabilitation services, 493
Medical Research Council Scale, 1908r
Medicare program
part A, 488–489, 488f
part B, 489
part C, 489–490
part D, 490
Medicare statutes, 485–486
Medication-associated myopathies, 770
Medications
GL, 1784–1787
impairing cognition, 1793r
reducing gastric acid secretion, 1786, 1786r
Medicolegal evaluations
history, 515
outside records review, 515
physical examination, 516
reports, 516
Medicolegal system
abandonment, 521
clinical practice standards
history, 515
outside records review, 515
physical examination, 516
reports, 516
deposition
day of, 519–520
definition of, 518
scheduling, 518–519
expert witness reviews, 520
independent medical evaluation (IME)
definition of, 516–517
ethics of, 517
option to treat, 518
serial, 518
without treatment, 517–518
malpractice and related complaints, 520–521
medicolegal work remuneration, 520
practice structure issues
antitrust, 522–523
compliance, 522
group practices and employment, 521–522
medical staff/hospital privileges, 521
payment and compensation issues, 522
Meditation, 2122
Medullary lateral reticular formation (MLRF), 1320–1321
Megesterol acetate, for anorexia and cachexia, 1944
Metabolic myopathies, 769
Metabolic equivalents (METs), 1084
Mestinon (pyridostigmine), 1491
Metoclopramide
Metformin
Methylprednisolone, 629
Methylphenidate for fatigue, 1913
Methylphenidate (Ritalin), 1793
Methotrexate (MTX), 1026–1027
Methadone, 1769–1770
Metatarsals, fracture, 965
Metastatic brain disease, 1160–1161
Metabolic myopathies, 769
Metabolic/endocrine peripheral neuropathy, 746–747, 746f
Metastatic brain disease, 1160–1161
Metatarsals, fracture, 965
Methadone, 1769–1770
Methotrexate (MTX), 1026–1027
MethylenePhenidate (Ritalin), 1793
MethylenePhenidate for fatigue, 1913
Methylprednisolone, 629
Methoclopramide
for biccups and cough, 1924
for nausea and vomiting, 1917
METs. See Metabolic equivalents
Mexiteline, 1299f, 1750–1751
MG. See Myasthenia gravis
MI. See Myocardial infarction
Miacalcin, 998r
Michigan Alcohol Screening Test (MAST), 606
Michigan alcohol screening test (MAST), 606
Microdialysis, 76
Microsporidiosis, 1233
Microstomia, 1144, 1145, 1145f
Microwave diathermy, 1700–1701, 1700r
Middle cerebral artery syndrome, stroke, 555–556
Midfoot sprain, classification, 968, 968r
MI-E. See Mechanical insufflation-exsufflation
Migraine headaches, 1494r
Military health care system, 492
Milwaukee type cervical-thoracic-lumbar-sacral orthosis (CTLSO), 901
Mind-body interventions
hypotherapy, 2122
meditation, 2122
Mind-body movement therapies, 2122–2124, 2123r
Mood disturbance, 2122
Mini-mental state examination (MMSE), 561
Minnesota Multiphasic Personality Inventory-2 (MMPI-2), 390–391
Minnesota Multiphasic Personality Inventory-2 Restricted Form (MMPI-2-RF), 391–392, 392r
Mitoxantrone, 630
Mixed agonists-antagonists, 1770–1772
Mixed dystonia, 416
Mixed hearing loss, 436, 436f
MLRF. See Masters and Johnson model
MLRF. See Medullary lateral reticular formation
MMPI-2. See Minnesota Multiphasic Personality Inventory-2
MMPI-2-RF. See Minnesota Multiphasic Personality Inventory-2 Restricted Form
Mobility
activities of daily living and, 1254
impairments
assistive technology, 1999–2001,
1999f–2001f
pediatrics, 2113
Mobilization, cervical axial and radicular pain, 828
MoCA. See Montreal cognitive assessment
Modafinil, 1793
Motor neuron disease (MND)
Amyotrophic lateral sclerosis
epidemiology, 717–718
etiology, 717
clinical trials, 723–724
diagnostic evaluation
adult nonhereditary form, 722–723
clinical presentation, 719–721, 720f, 721f
differential diagnosis, 720f, 721–722
Kennedy’s disease, 722
motor neuron disease, 723
spinal muscular atrophies, 722
pharmacologic management
antiapoptosis drugs, 725
antineurotrophic drugs, 725
anti-inflammatory drugs, 725
antioxidants, 725–726
cannabinoids, 726
growth factors, 725
neurodegeneration, 726
riluzole, 724–725
rehabilitation and palliative care, 726–728
spinal and spinobulbar muscular atrophy, 718f
adult-onset form, 718
neurology, 719
types, 718–719
Motor neuron disease, 116–117
Motor speech disorders
apraxia, 418
assessment, 416–417
dysarthria, 415–416, 415r
fluency, 419
head and neck cancer, 418–419
Motor system neuroprosthesis, 1797f
Motor unit action potential (MUAP), 95–96, 96f, 98–100, 99f
Movement disorders. See Parkinson’s disease
MPAI. See Mayo-Portland Adaptability Inventory
MRI. See Magnetic resonance imaging
MS. See Multiple sclerosis
MTX. See Methyltrexate
MUAP. See Motor unit action potential
Mucolysis, 1791
Multidisciplinary team, referrals and orders, 374
Multimorbidity evoked potentials (MEPs), 583
Multiple sclerosis (MS), 451
baclofen, 1327
benzodiazepines, 1326
biofeedback, 1942
dantrolene, 1328
demographics, 625
Devic’s disease, 417
diagnosis
criteria, 628
etrodiagnostic testing, 629
evoked potentials, 629
lumbar puncture, 629
magnetic resonance imaging, 628
ocular coherence tomography, 629
epidemiology, 625
etiology, 625
exercise prescription, 1660
gait rehabilitation, 2193, 2196
medical management
corticosteroids, 629
glatiramer acetate (GA), 630
interferons, 629–630
nizatine, 630
natalizumab, 630
optic neuritis (ON), 627
pain in, 1307
pathology, 625–626
pathophysiology, 626
planticery and, 2161–2162
Nonoperative functional rehabilitation, sports medicine
control inflammation, 1414–1415, 1414t
neuromuscular control retraining, 1416
proper motor control restoration, 1415
return to sport activities, 1416
strength restoration, 1415–1416

Nonopioid analgesics
for chronic pain, 1295–1296
for mild to moderate pain, 1292

Nutritional peripheral neuropathy, 746
Nutrition, 1361
Nu Step® recumbent stepper, 1444, 1444f
NSAIDs. See Nonsteroidal anti-inflammatory drugs
OA. See Osteoarthritis
OASDI. See Old age and social disability insurance
Obesity, 782–783
achondroplasia, 1513
exercise prescription, 1659
spina bifida, 1510
Objective measurement, with out standardized test
- discrete categorization, 238
- duration recording, 238
- event recording, 237
- interrater reliability, 238–239
- observing and recording behavior procedures for, 237–238
- rational, 236–237
- scores reporting, 239
- rate recording, 237–238

report scores and graph data, 239
- target behavior, 237
- identify, 237
- operational definition, 237, 237f
- time sampling, 238
Occupational therapy (OT), 792–793
Occtreotide, for nausea and vomiting, 1917
Office Of Public Health Service (OPHS), 493
Old age and social disability insurance (OASDI), 292
Older adults
- amputation, 1575
- arthritis, 1574
- assessment of, 1564–1566, 1565f
- health care policy issues, 1567
- hip fracture, 1573–1574
- joint replacements, 1574
- placement issues, 1567
- rehabilitation goal, 1566
- rehabilitation programs prescription, 1566–1567
- spinal cord injury (SCI), 1575
- stroke, 1574–1575
- traumatic brain injury, 1575
- Olanzapine burstin, 919
- Operant conditioning, 401–402
- OPHS. See Office Of Public Health Service
- Opiate receptor, 1808
- Opioid analogs, 1155–1156, 1155f
- Opioids, 1155–1156, 1155f
- Opponents orthosis metal design, 2056f
- Optic neuritis (ON), MS, 627
- Oral hypoglycemics, 1790, 1790t
- Oral medications, 1322, 1322f
- Order writing, 372
- interdisciplinary team, 374–375
- medical model, 373–374, 374f
- multidisciplinary team, 374
- transdisciplinary team, 375
- Organ system aging, 1550–1558, 1553f, 1555f
- Orthopedic rehabilitation, 1943
- Orthopedic shoes, for posterior tibialis tendinitis, 958
- Orthosis. See also Upper limb orthoses
distal femur, fracture of, 2076f
- limb instability, 2077–2078
- tibia, fracture of, 2076f
- Orthostatic intolerance, 1265–1266
- Orthotics
- hallux limitus, 961–962
- lower extremity materials, 2063–2066
- role and use of, 2063
- terminology, 2063
- phalangeal fracture, 965
- Osteoarthritis (OA)
- epidemiology, 782
- joint biomechanics, 783
- nonpharmacologic treatment
- aerobic conditioning and aquatic therapy, 793–794
- arthritis foundation, 785–787
- hand and occupational therapy (OT), 792–793
- hip and knee, 785, 785f
- lower limb orthotic intervention, 790–792, 791f
- patient education, 785–787
- physical therapy (PT), 788–790
- weight loss, 787–788
- pathology of, 781–782
Pharmacotherapy
GI bleeding risk factors, 796, 796f
image-guided injections, 801
injectable treatment, 800–801, 801r
medication algorithm, 794, 794f
NSAID, 796, 797, 799
opioid analogs, 797
renal toxicity, 795, 795f
topical analogs, 794f, 797
risk factors
age and gender, 782
ethnicity, 782
genetic factors, 782
muscle weakness, 783
obesity, 782–783
occupancy, 783
synoval joints, 783–785, 784f
thumb carpometacarpal, 940–941
Osteoarthritis, radiograph, 165, 166f
Osteochondritis dissecans, MRI, 158, 160f
Osteogenesis imperfecta (OI), childhood-onset
disability, 1511–1512
Osteomyelitis, magnetic resonance imaging, 141
Osteophytes, burn rehabilitation, 1137
Osteoporosis
clinical evaluation for bone markers, 987
Patient at Risk, 987–991
quantitating bone mass, 986–987
definition, 979, 979f, 979r, 980f
disse, 1263–1264
electronic medical record systems, 1007
epidemiology, 980–982
eutrology and risk factors, 982–983
female athlete triad, 1006–1007
pathogenesis
- cellular abnormalities, 984
genetic abnormalities, 985–986
hormonal abnormalities, 984–985
tissue abnormalities, 983–984
pediatric onset, adult disability
cerebral palsy, 1508
Duchenne muscular dystrophy, 1511
osteogenesis imperfecta, 1511–1512
spina bifida, 1510
spinal cord injury, 1506
pharmacologic agents, 997f, 998f
anabolic agents, 999–1000
antiresorptive agents, 997–999
bisphosphonates, 1000
nonFDA-approved drugs, 1000
preventive strategies
exercise, 993–994
fall reduction strategies, 994–996
nutritional adjuncts, 991–993
psychosocial considerations, 996–997
rehabilitation management postfracture

O
OA. See Osteoarthritis
OASDI. See Old age and social disability insurance
Obesity, 782–783
achondroplasia, 1513
exercise prescription, 1659
spina bifida, 1510
Objective measurement, with out standardized test
- discrete categorization, 238
- duration recording, 238
- event recording, 237
- interrater reliability, 238–239
- observing and recording behavior procedures for, 237–238
- rational, 236–237
- scores reporting, 239
- rate recording, 237–238
Pain
anatomy of, 1278–1288
burn rehabilitation, complications, 1138–1139
cancer, 1909, 1909r
analgesic approaches to, 1910r
causes of, 1909
epidemiology, 1909
invasive techniques, 1912
management and biofeedback, 1943
multiple sclerosis, 1307
myofascial, 1301–1302
from peripheral neuropathy, 1302
peripheral neuropathy and, 752
phantom limb, 1303–1304
amputation, 2042
poststroke, 1307–1308
residual limb, 1304
amputation, 2042–2043
scar neuroma, 1304–1305
scoliosis, 891–892
in spinal cord injury, 1306–1307, 1306r
syndromes, HIV, 1236–1237
transcutaneous electrical nerve stimulation (TENS), 1954–1955
traumatic brain injury (TBI), 594
treatment centers, 1291
treatment of, 1909–1912
Pain-reinforcing factors, 1277–1278
Palliative care
adult motor neuron disease
comprehensive management, 726–727
depression, 732
dysarthria, 731
dysphagia, 731, 732r
diffuse end-of-life issues, 733–736, 734f, 735r
pain and immobility, 732–733
respiratory system, 728–731, 729f
spasticity, 731–732
weakness and fatigue, 727
anorexia and cachexia, 1918–1920
dyspnea, 1920–1923
exercise, 1926–1927
fatigue, 1912–1914
hiccups and cough, 1923–1925
hydration and nutrition, 1925–1926
nausea and vomiting, 1914–1918
origins, 1905
pain, 1908–1912
relevance to physical medicine and rehabilitation, 1905–1906
symptom assessment, 1906, 1907–1908r
Palpation
cervical spine, 56
knee, 63–64, 63f
lumbar spine, 61
shoulder, 57
Pamidronate for osteoporosis, 1000
Paralytic impairments.
Paralytic impairments. See Neuromuscular disease (NMD)
Paraneoplastic neuromuscular disorders (PND), 1138–1139
Parathyroid hormone, for osteoporosis, 984
Parenteral hypoglycemics, 1789–1790
Parenting skills, disabled children, 1480
Parenting skills, disabled children, 1480
Parenting skills, disabled children, 1480
Parkinson’s disease, 451
aerobic exercise, 654–655
akathisia, 658
anatomy, 645–646, 646f
athetosis, 658
ballismus, 658
chorea, 658
diffuse lewy body disease, 656–657
drug-induced movement disorders, 659, 659r
dystonia, 657
impairments and functional limitation, 647r
bladder abnormalities, 651
cognition, 651
dyskinesia and dystonia, 650
gait dysfunction, 649, 649r
gastrointestinal problems, 650–651
movement initiation, 650
orthostatic hypotension, 650
pain, 652
psychiatric manifestations, 651–652
therapeutic plan, 648
movement disorders, 656, 656r
multiple system atrophy (MSA)
pathophysiology, 645–646, 646f
physical exercise, 654–656, 655r
progressive supranuclear palsy (PSP), 656
psychogenic movement disorders, 658–659
rehabilitation and collaboration, 647
signs and symptoms, 656r
stereotopy, 658
therapy and exercise effect, 655r
tics, 658
treatment
deep brain stimulation, 652–654, 654r
fetal cell transplantation, 654
pharmacological management, 652, 653r	
tremors, 657
Parodel, 1793
Parsonage-Turner syndrome, shoulder injury, 1424
Partial body weight-supported treadmill training (PBWSTT), spinal cord injury, 686
Partial foot amputation prostheses, 2026
Partial hand prostheses, 2039, 2039f
Participation survey/mobility (PARTS/M), 477
PARTS/M. See Participation survey/mobility
Patella tendon-bearing (PTB) socket, 2027
Patellar tendinitis, MRI, 157, 159
Patellar tendon-bearing (PTB) braces for diabetic foot deformities, 972
Pathophysiology
cumulative trauma disorders (CTDs), 923
myelopathy, 821–822, 822f
radiculopathy, 815–816, 816f
Patient education, 785–787
Patient health questionnaire-9 (PHQ-9), 394, 394r
Patient history
family history, 11
functional history and examination
ambulation, 8, 45–46, 45r
bathing, 6, 44
bed activities, 7, 44
communication, 6
dressing, 7, 44
eating, 6, 44
grooming, 6, 44
motor vehicle, operation of, 8, 46
quantitation of, 45f, 46
transfers, 6–7, 44
wheelchair mobility, 7–8, 45
illness, 5, 5r
patient profile (see also Clinical evaluation)
personal history, 10, 10r
social history, 11
vocational history, 11
symptoms, 5, 5r
Patients selection, 503–504
Payer contracts, 522
Payment
definition, 518–519
medicolegal system, 522
PCAs. See Personal care attendants
Pedagogic issues, ethics, 511–512
Pediatric evaluation of development inventory (PEP), 2193
Pediatrics, wheelchair mobility, 2115
psychosocial and family issues, 2113
seating systems, 2112–2113
Pelvis, asymmetry of, 891, 891f
PEP. See Personalized exercise program; Positive expiratory pressure
Performance-based progressive therapy (PBPT), 2193
Perfusion-weighted imaging (PWI), 204
Peripheral arterial disease, in diabetics, 968
Peripheral nervous system. See also Electrodiagnostic evaluation
anatomy and pathology, 742r
degeneration and regeneration, 742–744, 745r
denervation, 1942
entrapment neuropathies, 103–109, 105f
INDEX I-23
Peripheral nervous system. See also Electrodiagnostic evaluation (Continued)

injury
- electrodiagnostic findings after, 751r
- following trauma, repair of, 752–753
- mononeuropathy, diagnosis of, 751r
- motor neuron disease, 116–117
- myopathy, 115–116
- plexopathy, 113–114
- polyneuropathies, 114–115
- radiculopathy, 112–113
- syndromes
  - median neuropathy, 920–921
  - radial neuropathy, 920
  - ulnar neuropathy, 920
- traumatic neuropathies, 109–112, 110f

Peripheral neuropathy, 741–753, 1942
- classification of, 741–742, 742r
- compression/entrapment, 748r
- diabetic, 741, 747, 747r
- hereditary, 745–746, 746r
- infectious, 747–748
- neuroanatomy and neurophysiology, 741, 742
- nontraumatic, 744
- pain, 1302
- treatment of, 751–752, 752r

Peripheral quantitative computed tomography (pQCT) scan, 1958
- 741, 742

Pharmacologic management, 1024, 1025
- hydroxychloroquine (HCQ), 1027
- gold, 1029
- immunomodulatory drugs, 1027–1028, 1028r
- leflunomide, 1027
- methotrexate (MTX), 1026–1027
- nonsteroidal anti-inflammatory drugs (NSAIDs), 1025–1026, 1026r
- sulfasalazine (SSA), 1027

Pharmacotherapy
- GI bleeding risk factors, 796, 796r
- image-guided injections, 801
- injectable treatment, 800–801, 801r
- medication algorithm, 794, 794f
- NSAID, 796, 797, 799
- opioid analgesics, 797
- renal toxicity, 795, 795r
- topical analgesics, 796f, 797

Phenobarbital (Luminal, Solfoton), 1754
- Phenothiazines for nausea and vomiting, 1917
- Photoplethysmography, 1180

PHQ-9. See Patient health questionnaire-9

Physical activities
- community-based physical activity, 1437, 1438f
- personalized exercise program (PEP), 1438–1440, 1439f
- tailoring programs, 1437–1438
- types of, 1440–1441, 1440f
- fitness programm, 1442
- cardiorespiratory fitness suggestions, 1443, 1443r
- cardiovascular exercise equipment, 1444, 1444r
- flexibility exercises, 1446, 1446f, 1447r
- lower-body exercise, 1444, 1444f
- Nu Step® recumbent stepper, 1444, 1444f
- parameters, 1443, 1443f
- safety guidelines, 1444, 1445r
- strength training guidelines, 1445, 1445r
- leisure physical activity, 1447r
- bowling, 1448, 1448f
- dancing, 1448–1449, 1449f
- gardening, 1447–1448, 1448f
- kayaking, 1450, 1450f
- sailing, 1450–1451, 1451f
- swimming, 1449–1450, 1450f
- National Center on Physical Activity and Disability (NCPAD), 1441–1442
- personalized exercise program (PEP) barriers, 1439–1440
- community-based physical activity options, 1440
- health status, 1440
- intervention model, 1438, 1439f
- mobility limitations, 1440
- motivational level, 1440
- physical activity profile, 1438–1439
- PET. See Positron emission tomography
- Pétrissage, western massage, 1726–1727
- Phantom limb pain, 1303–1304
- amputation, 2042

Pharmacokinetics, 1809–1810

Pharmacologic actions, 1810

Pharmacologic management, 1024, 1025f
- acetylsalicylic acid (ASA), 1025
- antihyperuricemic agents, 1029
- azathioprine and mercaptopurine, 1027
- biologic agents-anticytokine therapies, 1029
- combination therapies, 1029
- cyclophosphamide and chlorambucil, 1027
- cyclosporine, 1027–1028, 1028r
- disease-modifying antirheumatic drugs, 1026–1027
- D-Penicillamine, 1027
- glucocorticoids, 1026
- gold, 1029
- hydroxychloroquine (HCQ), 1027
- immunomodulatory drugs, 1027–1028, 1028r
- leflunomide, 1027
- methotrexate (MTX), 1026–1027
- Physical dependence, defined, 1300r
- Physical examination
  - myelopathy, 824
  - radiculopathy, 818–819
- Physical medicine and rehabilitation
  - palliative care, 1905–1906
- Physical medicine and rehabilitation (PM&R)
  - the rulemaking process, 485–486
  - critical current regulatory
  - health care reform and medicare overhaul, 495
  - medical necessity review, 496
  - methods of advocacy and representation, 496–497
  - payment updates, 496
  - rehabilitation providers, 495–496
  - facility definitions and licenses, 494
  - federal budget process, 494–495
  - governmental influence, 485
  - governmental programs
    - military health care system, 492
    - Office Of Public Health Service (OPHS), 493
    - TRICARE, 492
    - veterans administration, 492
    - vocational rehabilitation, 491
    - workers compensation, 491–492
    - health and disability insurance, 493
  - inpatient rehabilitation facility (IRF) cases, 485, 486, 487r
  - legislative issues, 495–497
  - medicaid, 490–491, 490f
  - medical rehabilitation services, 493
  - medicare program
    - part A, 488–489, 488f
    - part B, 489
    - part C, 489–490
    - part D, 490
    - medicare statutes, 485–486
    - medicine laws, corporate practice, 494
    - professional liability insurance, 494
    - professional licensing, 494
  - State Children's Health Insurance Program (SCHIP), 401
  - state legislative and regulatory aspects, 494–495
  - statutory and regulatory developments
    - 75 percent rule, 486–487, 487r
    - medicare recovery audit contractors, 487–488
    - outpatient therapy caps, 487
    - payment system changes, 487
    - payment updates, 487
    - quality of care, 488
- Physical medicine and rehabilitation, biofeedback
- balance rehabilitation, 1941
- cerebral palsy and traumatic brain injuries, 1941–1942
- dystonias and dyskinesias, 1942
- multiple sclerosis (MS), 1942
- orthopedic and prosthetic rehabilitation, 1943
- pain management and, 1943
- peripheral nerve denervation, 1942
- polyneuropathy and peripheral neuropathy, 1942
- spinal cord injury (SCI), 1941
In this page, the text includes various medical and scientific terms related to psychological aspects, psychiatric disorders, traumatic brain injury, and associated treatments and assessments. Here are some key points extracted from the text:

- **Psychological aspects**
  - Beck depression inventory-fast screen (BDI-FS), 394
  - Behavioral management contracts, 402
  - Modification, 402
  - Reinforcers types, 401–402
  - Token economies, 402
  - Chemical use assessment
  - Alcohol use disorders identification test (AUDIT), 400
- **Psychiatric disorders**
  - Traumatic brain injury
  - Proximal interphalangeal joint
  - Prostate cancer
  - Prospective memory
  - Protein for osteoporosis
  - Protease inhibitors for HIV
  - Prosthesis
  - Psychometric parameters
  - Concurrent validity
  - Construct validity
  - Predictive validity
- **Prosthetic fitting**
  - Jointly powered
  - Externally powered
  - Partial hand
  - Thumb
- **Psychometric parameters**
  - Repeated measurements
  - Test-retest
  - Point-by-point agreement ratio
  - Frequency ratio
  - Pearson product-moment correlation
- **Psychosocial issues**
  - Quality of life (QoL)
  - Physical therapy
  - Pain
  - Disability
  - Physical function
- **Quality of life**
  - Pulmonary disorders
  - Radical prostatectomy
  - Radiation therapy
  - Reproductive fitness
  - Psychosocial issues
  - Primary care
  - Pain management
  - Psychologic assessment
  - Social skills training
  - Social interaction changes

The text also mentions various tests and assessments such as the Beck Depression Inventory (BDI), the Mini-Mental State Examination (MMSE), the Minnesota Multiphasic Personality Inventory (MMPI), and the Wechsler Adult Intelligence Scale (WAIS). The document appears to be a comprehensive guide to the assessment and management of various clinical conditions and their psychological implications.
INDEX

Service systems, VR (Continued)
state VR agencies, 467
vocational assessment, 468
Servicemen’s Readjustment Act, 463
Sexual disorders
Sexual dysfunction
sexual dysfunctions, 445–446
sexual dysfunctions, 445–446
sexual dysfunction
multiple sclerosis, 434–435
stroke, 370–371
therapies, 435
erection enhancement, 454–455
female arousal, 455
male ejaculatory, 455
nongenital organic options, 455–456
organic difficulties, 455
sexual drive problems, 454
Sexual function complexity
models of, 445
sexual dysfunctions, 445–446
Sexual neurophysiology
sexual neurophysiology, 446–447
Sexuality and disability
Sexuality and disability
disability-related disruptions
inging, 437–448
bladder and bowel concerns, 449
vascular factors and risks, 448–449, 449f
fertility concerns, 450
impending mortality, 448
mechanisms, 447, 448f
neurological changes, 448
pain issues, 450
psychological factors, 448
MBJ model, 445
rehabilitation framework, 453
sexual dysfunctions therapies, 453
erection enhancement, 454–455
female arousal, 455
male ejaculatory, 455
nongenital organic options, 455–456
organic difficulties, 455
sexual drive problems, 454
sexual function complexity
models of, 445
sexual dysfunctions, 445–446
sexual neurophysiology, 446–447
specific disabilities
arthritis tissue diseases, 452–453
chronic renal failure, 452
connective tissue diseases, 452–453
corony artery disease, 452
diabetes, 452
limb amputation, 451
multiple sclerosis, 451
neuromuscular disorders, 451
Parkinson’s disease, 451
pulmonary disease, 452
spinal cord injury (SCI), 450
stroke, 452
traumatic brain injury, 451–452
Sexuality, disabled children, 451
Shear, 1568
Shiatsu, 2127
Shoulder
Shoulder disorders, 466–468, 1468f
Shoulder orthoses, 2059–2060, 2059f, 2060f
Shoulder pathology
bicipital tenosynovitis, 3, 4f
subacromial-subdeltoid bursitis, 2, 2f
supraspinatus tendon tear, 1, 2f
Shoulder, physical examination
impeing test, 59, 59f
inspection, 57
lbral tests, 59
neurological sensory testing, 59
subcapsularis strength, 58, 58f
supraspinatus strength, 58, 58f
overhead athlete, biomechanics, 61
 palpation, 57
range of motion
internal rotation determination, 57, 57f
normal range, 58f
scapula, stabilization of, 58, 58f
sports biomechanics, 60–61
stability tests, 59–60, 59f, 60f
Shoulder subluxation, 1961–1962
Shoulder-elbow-wrist-hand orthosis, 2060f
Sialorrhea (drooling)
with ALS, 731
Sick role, 404
SII. See Strong Interest Inventory
Siedenahl, 703
Silesian belt, 2030–2031
Similarities, 42
Single muscle fibers, 81–82, 82f
Single-photon emission CT (SPECT), 587, 1893
Single-limb stance (SLS), 122, 132
Sirolimus, 1209–1210
Skin
breakdown, 1509
burn rehabilitation
blisters, 1140
hypertrophic scarring, 1139–1140, 1139f
malignancy, 1141
pathophysiology, 1125, 1126f
UV sensitivity and skin pigmentation, 1125, 1126f
malignancy, 1141
pathophysiology, 1125, 1126f
UV sensitivity and skin pigmentation, 1125, 1126f
normal
anatomy, 1358
pathophysiology, 1358
preparation, injection procedures, 1817
shoulder, physical examination
impeing test, 59, 59f
inspection, 57
lbral tests, 59
neurological sensory testing, 59
subcapsularis strength, 58, 58f
supraspinatus strength, 58, 58f
overhead athlete, biomechanics, 61
 palpation, 57
range of motion
internal rotation determination, 57, 57f
normal range, 58f
scapula, stabilization of, 58, 58f
sports biomechanics, 60–61
stability tests, 59–60, 59f, 60f
Sialorrhea (drooling)
with ALS, 731
Sick role, 404
SII. See Strong Interest Inventory
Siedenahl, 703
Silesian belt, 2030–2031
Similarities, 42
Single muscle fibers, 81–82, 82f
Single-photon emission CT (SPECT), 587, 1893
Single-limb stance (SLS), 122, 132
Sirolimus, 1209–1210
Skin
breakdown, 1509
burn rehabilitation
blisters, 1140
hypertrophic scarring, 1139–1140, 1139f
malignancy, 1141
pathophysiology, 1125, 1126f
UV sensitivity and skin pigmentation, 1125, 1126f
normal
anatomy, 1358
pathophysiology, 1358
preparation, injection procedures, 1817
INDEX

SLAP lesion, 145f
SLE. See Systemic lupus erythematosus
Sleep
apnea, SCI, 689
disorders, 1052, 1571–1572
disturbance, 594–595
exercise therapy, 1637
Sleep-disordered breathing treatment, 1117
Sleeve suspension system, 2028
Sliding-filament theory, 1623
SLS, 122, 132
SMA. See Spinal muscular atrophy (SMA)
Small bowel transplantation, 1215–1216
Small molecules, 2159
Smith-Fess Act, 462
Smith-Hughes Act (PL 65–178), 462
Smoking
cardiac rehabilitation, 1078
cessation, 1597
with stroke, 552
SNAP, 90–92, 90f, 91f
Snapping hip, 1461
SNAPs. See Sensory nerve action potentials
SNARE complex, 1331
Snow skiing, recreational sports, 1452, 1452
SNARE complex, 1331
SNAPs.
Sodium fluoride, for osteoporosis, 1000
7.5% Sodium bicarbonate, 1758
Social skills
treatment considerations, 417–418
Social Security Disability Insurance (SSDI), 292
Social Security Administration (SSA), 289, 292
Social Security Act, 463
social interaction changes, 402
Social history, 11
Snow skiing, recreational sports, 1452, 1452
SNARE complex, 1331
SNAPs.
Spastic quadriplegic, opisthotonic posture in,
Spastic dysarthria, 416
Spas, 1685
SOMI, 2085, 2085
Somatosensory evoked potentials (SEP), 583
SOMI, 2085, 2085f
Spa, 1685
Spastic dystarhria, 416
Spastic gait
with aging, 134
Spastic quadriplegic, opisotonic posture in,
Spasticity management
spasticity-reducing interventions, 133–134
Spastic quadriplegic, opisotonic posture in,
spin echo, 139
Spina bifida management, childhood-onset
disability in adult
bowel dysfunction, 1509
obesity, 1510
osteoarthritis, 1510
renal failure, 1509
sexuality, 1510
skin breakdown, 1509
Spinal articulations, 886–887
Spinal block, accidental, 1821
Spinal cord, 1161
MRI, degenerative disorders
atlantoaxial instability, 185
cervical disc herniation, 172–173,
174f, 175f
cervical stenosis, 174–176, 175f–177f
disc, 169–172, 175f
lumbar disc herniation, 176–178,
177f, 178f
Spinal cord, 1161
MRI, degenerative disorders
atlantoaxial instability, 185
cervical disc herniation, 172–173,
174f, 175f
cervical stenosis, 174–176, 175f–177f
disc, 169–172, 175f
lumbar disc herniation, 176–178,
177f, 178f
spasticity, 139
Spinal bifida management, childhood-onset
disability in adult
Spinal cord (Continued)

spine, 168–169, 169f–172f
spondylosis, 180–181, 180f–182f
thoracic disc herniation, 176
trauma, 181–185, 183f–185f
Spinal cord dysfunction, disabled children 
bowel management, 1498
Chiari malformation, 1496–1497
demographics, 1496
embryology, 1495–1496
etiology 
genitourinary abnormalities, 1497–1498
arthropathy, 1497–1498
embryology, 1495–1496
Chiari malformation, 1496–1497
bowel management, 1498
neurological assessment, 1499–1499f 
team approach, 1499
traumatic spinal cord injury, 1495, 1495f
Spinal cord dysfunction, HIV, 1239
Spinal cord injury 
exercise prescription, 1660
pain in, 1306–1307, 1306f
Spinal cord injury (SCI), 450, 1941, 
2162–2163
age, gender, and race, 665
anatomy, 670–672, 670f
autonomic dysreflexia, 690–691, 690f
baclofen, 1327
benzodiazepines, 1326
bowel management, 692–693
cardiovascular disease (CVD), 704
clonidine, 1328
dantrolene, 1328
death causes, 667
diabetes, 705
discharge placement of, 666
etiology, 665–666
fertility, 703–704
functional evaluation, 676–677
gastrointestinal complications, 691–692
neurological system 
neuropathy, 699–691, 699f
baclofen, 1327
benzodiazepines, 1326
bowel management, 692–693
cardiovascular disease (CVD), 704
clonidine, 1328
dantrolene, 1328
death causes, 667
diabetes, 705
discharge placement of, 666
etiology, 665–666
fertility, 703–704
functional evaluation, 676–677
heterotopic ossification, 689–690
musculoskeletal, 699
neuromusculoskeletal, 700
post-traumatic syringomyelia, 700–701
tetraplegia, 701–702, 702f
treatments, 699–700
prevalence, 665
prognosticating neurological recovery
AIS A conversion and late recovery,
677–678
ambulation, 685–686
assisted technology devices (ATDs), 686
C1–4 level, 680–683
CS–5 level, 683–684
driving independently, 686–687
functional goals, 680, 681f, 682f
home accessibility, 686, 686s
intramedullary hemorrhage, 678
L1–5 level, 685
modalities, 684
nontraumatic injury, 678–679
pressure relief, 684
range of motion, 684
reflex absence, 678
standing exercise, 685
T1–12 level, 684
tetraplegia, 701–702
transfer, 684–685
psychological issues 
anxiety, 703
depressive disorders, 702–703
post-traumatic stress disorder (PTSD), 703
substance abuse, 703
treatment, 703
respiratory failure, 688
sexuality, 703–704
sleep apnea, 689
thromboembolic disorders, 687–688
tizanidine, 1329
vascular supply, 670–672
weaning, 688
Spinal cord stimulation, 1962–1964
Spinal curves, pathologic, 884f
Spinal deformities. See Scoliosis
Spinal deformity, 891f, 901f
clinical assessment of, 887–889
management of, 77%–776
Spinal dysraphism, 1495, 1496f
Spinal fixation instrumentation 
archetypes of, 903f
Spinal fracture site pain, 1307
Spinal injection procedures 
caudal lumbar epidural injection, 
1880–1881, 1880f, 1881f
coccyx injections, 1898
contrast enhancement in, 1882
discography, 1894–1898, 1894f, 
1895f–1897f, 1897f
epidural steroid injections (ESIs), 1875–1876
fluoroscopic guidance
epidural injections efficacy, 1883–1884
radicular pain management, 1884
safety and complications of, 1884–1885
interlaminar epidural injection, 1876–1877,
1877f, 1878f
nerve root block, 1885
sacroiliac (SD) joint, 1892–1893, 1892f, 1894f
timing, frequency, dose, and volume,
1881–1882, 1882f
transforminal epidural injection, 
1878–1880, 1878f–1881f
zygapophyseal joint (Z-joint), 1886–1892,
1887f–1889f, 1890f
Spinal interneurons, 1320
Spinal manipulation therapy. See Manipulation
Spinal muscular atrophy (SMA), 718–719
Spinal nerve root block, 1828
Spinal ostitis, 1046
Spinal orthotics 
action, mechanism of, 2082f
control of motion, 2081
spinal alignment, 2081–2082, 2082f
trunk support, 2081
cervical orthoses (COS), 2082f
biomechanics, 2082–2083
cervicothoracic orthosis (CTO), 2085–2086
halo orthosis, 2086–2088, 2086f
hard cervical collars, 2083–2085, 2085f
soft cervical collars, 2083, 2083f
clinical uses for, 2092–2093
lumbosacral orthoses, 2091–2092, 2091f, 
2092f
sacroiliac orthoses, 2092, 2092f
thoracolumbosacral orthosis (TLSO), 
2088–2091, 2089f, 2090f
Spinal pain, radicular, pathophysiology of,
1875–1876
Spinal rotation, Moe-Nash classification of, 890f
Spinal stability, 669
Spinal stenosis 
cervical, 174–176, 175f–177f
lumbar, 863–865
Spinal trauma, 181–185
Spine. See also Scoliosis
anatomy and physiology, 885–887
aqueatic rehabilitation, 1683–1684
assessing rotation, 890–891
assessing vertebral integrity of, 891
imaging, 167–185
injuries, 1461
mechanical instability of, 1306
radiographic assessment of, 889–891
stabilization technique, aquatic rehabilitation,
1683–1684
SpineCor, 901
Spine-stabilized swimming programs, 1684
Spinobulbar muscular atrophy (SBMA), 718–719
Splint airplane, 2095f
Spondyloarthropathies, 1060
Spondylolisthesis and lumbar spine instability,
865, 867–870
Task-oriented biofeedback application, in rehabilitation (Continued)

research and application on, 1946–1947

robot-assisted retaing, 1946

static biofeedback, 1943–1944

virtual reality (VR), 1944–1946, 1945f

TBI. See Traumatic brain injury

Team function

building and development

communication barriers, 365–366

established team complacency, 365

new team, 364–365, 364f, 364r

personal characteristics, 365, 363r

care providers multiplicity, 360, 360r–361r, 371–372

communication, 375–376, 376r

communication methodology, 366

conflict and disagreement, 363

dynamics, 362–363, 362r

static biofeedback, 1943–1944

robot-assisted retraining, 1946

research and application on, 1946–1947

therapy referrals and order writing, 372

building and development

dynamics, 362–363, 362r

multiple caregivers, 359, 360f

treatment planning algorithm, 378, 378f

regulatory organizations, 370–371, 371f

team meeting management, 371, 371f

transdisciplinary team model, 368–370

medical necessity, 377–378

resources, 383

interactions styles

approach to use, 370

interdisciplinary team model, 368, 368f, 369f

medical model, 366–367

multidisciplinary team mode, 367–368, 367f

regulatory organizations, 370–371, 371f

team members associations, organizations and journals, 359, 360–361r

therapy quality control

Centers for Medicare and Medicaid Services (CMS), 380

diagnosis-related groups (DRGs), 379–380

durable medical equipment (DME) and therapy, 379, 379r

treatment-planning algorithm, 378, 378f

therapy referrals and order writing, 372

interdisciplinary team, 374–375

medical model, 373–374, 374r

multidisciplinary team, 374

transdisciplinary team, 375

Teams

assistive technology, 2009–2010

building and development

communication barriers, 365–366

established team complacency, 365
	new team, 364–365, 364f, 364r

personal characteristics, 365, 363r

dynamics, 362–363, 362r

Tegretol, 1751–1753, 1753r

Telecommunications Act of 1996 (PL 104–104), 466

Telehabilitation, 1946

Temporal arthritis, 1193

Tendinitis, MRI, 145–146, 146f

Tendon injuries

classification of, 908, 909r

functional scale, 908, 909r

structural characteristics, 907, 908f

Tendon surgery, 1032

Tendons

Abductor pollicis longus (APL), 928–929

injuries of classification of, 908, 909r

functional scale, 908, 909r

structural characteristics, 907, 908f

Tendons, physiology, 1414

Tennis, recreational sports, 1453, 1453r

Tendons, physiology, 1414

Tenodesis, 928

stenoing, 942–943, 942f

TENS. See Transcutaneous electrical nerve stimulation

Testosterone, osteoporosis, 999

TGF beta, osteoporosis, 985

Thalidomide, for nausea and vomiting, 1920

Thermoplastic molded thoracic lumbar sacral orthosis (TLSO), 901

Therapeutic exercise

acute physiologic responses

dynamic exercise, 1631–1633, 1632f

static exercise, 1633–1634, 1633f

aerobic exercise, 1637–1639

aerobic metabolism, 1620

anaerobic metabolism, 1619

for chronic pain, 1294–1295

anaerobic metabolism, 1619

energy continuum, 1620, 1620f, 1621f

exercise prescription, 1654–1662

flexibility, range of motion, 1652–1653, 1653f

fuel utilization, 1620–1621, 1621f, 1621r

glycolysis, 1619

health benefits derived, 1634–1635

muscle physiology

functional factors, 1625–1629, 1626f, 1628f

muscular fatigue and endurance, 1630–1631, 1630f, 1631f

optimal phenomena, 1629–1630, 1629f

structure and function, 1622–1625, 1622f

1623f, 1624r

oxygen uptake, 1621f, 1622

proprioception, 1653–1654

strengthening exercise, 1639–1652

Therapeutic physical agents

cryotherapy, 1701–1702, 1701f

diathermy, 1696–1701

electrotherapy, 1702–1707

heat and cold, 1691–1693, 1692r

hydrotherapy, 1693–1696

modality choice and prescription, 1707, 1707r

ultraviolet (UV) radiation, 1702

Therapeutic robots, 2187

Therapeutic shoes, for diabetic foot ulcers, 972, 972f

Therapeutic ultrasound, 1368

Therapy quality control

Centers for Medicare and Medicaid Services (CMS), 380

diagnosis-related groups (DRGs), 379–380

durable medical equipment (DME) and therapy, 379, 379r

treatment-planning algorithm, 378, 378f

Therapy referrals, 372

Interdisciplinary team, 374–375

medical model, 373–374, 374r

multidisciplinary team, 374

Transdisciplinary team, 375

Thermoplast mold hip, 1366

Thermoplast molded thoracic lumbar sacral orthosis (TLSO), 901

Thoracic outlet syndrome (TOS), 749, 932, 1192–1193, 1460, 1460f

Thoracicolumbar support, osteoporotic vertebral fracture, 1003

Three-point control concept, 2053, 2053f

Throat, examination of, 12

Thromboangitis obliterans, 1192

Thrombocytopenic purpura, 687–688

Thrombophilia, 1194

Thrombosis, stroke, 553

Thumb prosthesis, 2039f

spica, 2056f

Thumb carpometacarpal osteoarthritis, 940–941

Thyroid disorders, associated with myopathies, 771

Tiagabine (Gabitril), 1755–1756

Tibial calcaneal fusion, 971

Tibial nerve, 790

Tibial calcaneal fusion, 971

Tibial nerve, 790

Tibial arteries, 954

Tibial nerve, 750

Tibial arteries, 954

Thumb carpometacarpal osteoarthritis, 940–941
Tumor necrosis factor (TNF), 1160

Tumors
brain tumors
benign brain tumors, 1171
malignant brain tumors, 1171–1172
rehabilitation principles, 1172
breast cancer, 1167–1168
gastrointestinal malignancies, 1171
head and neck cancer, 1168–1170,
1168r, 1169f
hematologic malignancies, 1170–1171
lung cancer, 1171
prostate cancer, 1173–1174
sarcomas of bone and soft tissue, 1172–1173
TWWIA. See Ticket to Work and Work
Incentive Improvement Act
Tylenol, 1743, 1745, 1745
Type 2 diabetes mellitus prevention
Ultraviolet (UV) radiation therapy, 1702
Ultrasound
biophysics, 1069
equipment, 1696–1697
indications, 1697–1698
precautions and contraindications, 1698
technique, 1697, 1697f
elbow pathology
lateral epicondylitis, 3, 4f
ulnar collateral ligament (UCL) rupture, 5, 5f
knee pathology
knee effusion, 5, 6f
muscular injury, 5–6, 7f
patellar tendonitis, 5, 7f
posterior cruciate ligament (PCL) injury, 6, 8f
machine and transducer probes, 1, 1f
shoulder pathology
bicipital tenosynovitis, 3, 4f
subacromial-subdeltoid bursitis, 2, 2f
supraspinatus tendon tear, 1, 2f
Ultrasound imaging (USI)
joint, 208
muscle, 204, 206, 206f, 207f
peripheral nerve, 207–208, 207f, 208f
Ultrasound (UV) radiation therapy, 1702
UMNS. See Upper motor neuron syndrome
Unilateral upper motor neuron dysarthria, 416
United States Preventive Services Task Force
(USPSTF), 1601
Universal cuff, 2060f
University of Chicago Fear of Falling Measure
(UIC FFM), 995
University-based training programs
assistive technology, 2010r
Upper back pain, 1471
Upper extremity
amputation
bilateral, 2035–2036
postoperative care, 2034
preprosthetic patient evaluation and
management, 2034
prostheses, 2036, 2036r
by level of amputation, 2039–2041
prosthetic evaluation and control training, 2035
prosthetic fitting, 2034–2035
prosthetic follow-up, 2035
muscle innervation, 1287r
orthoses (See Upper limb orthoses)
rehabilitation
EMG biofeedback for 4–5
rehabilitation and stroke, 2190
Upper extremity ischemia, 1190–1191
Upper limb applications
neuromuscular electrical stimulation, 1955–
1956, 1956f
Upper limb orthoses
anatomical principles, 2054–2055, 2054f
biomechanical principles
lever force, 2053f
materials selection, 2054
three-point control concept, 2053, 2053f
tissue tolerance, 2053
traction angle, 2054f
definitions and nomenclature, 2051
designs
elbow orthoses, 2058–2059, 2059f
finger and thumb orthoses, 2055, 2055f
hand-finger orthosis, 2055–2056, 2056f,
2057f, 2058f
shoulder orthoses, 2059–2060, 2059f,
2060f
wrist-hand-finger orthosis, 2057–2058,
2057f, 2058f
goals, 2051–2053
prescription template, 2052f
technology, 2060
Upper motor neuron syndrome (UMNS)
aquired brain injury (ABI)
baclofen, 1327
benzodiazepines, 1326
clonidine, 1328
dantrolene, 1328
tizanidine, 1329
cerebral palsy (CP)
baclofen, 1327
benzodiazepines, 1326–1327
clonidine, 1328
dantrolene, 1328
tizanidine, 1329
spasticity management
baclofen, 1327, 1334–1335
benzodiazepines, 1326–1327
dantrolene toxin therapy, 1332–1334
chemical denervation agents, 1330–1332, 1333f
clonidine, 1328
cognitive status, 1325
concurrent medical problems, 1325
dantrolene sodium, 1327–1328
distribution of, 1325
electrical stimulation (ES), 1324
etiology, 1325
financial issues, 1325
functional prognosis, 1325
goals for treatment, 1323
intrathecal medications, 1334–1335
massage, 1324
neurosurgical procedures, 1335–1336
noxious stimulation reduction, 1325
oral and transdermal medications,
1325–1330, 1326r
oral medications, 1322, 1322r
orthopedic procedures, 1336–1337
pharmacologic treatment, 1324–1325
physical modalities, 1324
positioning, 1323
sample passive and active goals, 1322,
1322r
stretch and casting, 1323–1324
support system, 1325
surgical techniques, 1334–1337
time since onset, 1325
tizanidine, 1329–1330
transcutaneous electrical nerve stimulation
(TENS), 1324
treatment options, 1323–1324
spasticity pathophysiology
muscle overactivities, 1321–1322
origin of, 1321
spinal cord injury (SCI)
baclofen, 1327
benzodiazepines, 1326
clonidine, 1328
dantrolene, 1328
tizanidine, 1329
Urinary diversions, 1361–1362
Urinary tract infections
asymptomatic urinary tract infections, treat-
ment of, 1371
bacteriuria, 1370
complications of, 1371
definition of, 1369–1370
prophylactic antibiotics role, 1371–1372
pyuria, 1370
sign and symptoms, 1370
symptomatic urinary tract infections, treat-
ment of, 1371
Urinary urethral sphincters, 1346–1347
Urodiynamics
bladder emptying, 1356, 1356f
bladder filling (storage phase), 1356
indications, 1355
multichannel water-fill urodynamic study,
1357, 1357f, 1358f
U.S. disability systems comparison, 291r
Department of Veterans Affairs (VA), 293
FECA, 293
Federal Black Lung Program, 293
FELA, 293
LHWCA, 293
private disability insurance, 292–293
SSDI, 292
supplemental security income, 292
tort claims, 290
workers’ compensation, 290, 292, 292r

VA. See Department of Veterans Affairs (VA)

Vascular myelopathy
HIV, 1239

Vacuum-assisted socket suspension (VASS), 2029f

Valproate (Valproic acid, Depakene), 1756–1757

Vasopneumatic compression therapy, 1201

Vegetative state with TBI, 585–587

Venous disease
chronic venous insufficiency, 1196f
rehabilitation, 1196–1197
surgery, 1197
deep vein thrombosis (DVT)
anticoagulants, 1195
medical treatment, 1194
phlegmasia cerulea dolens, 1196
thrombolytic agents, 1195
vena cava filters, 1195–1196
May-Thurner syndrome, 1197
thrombophilia, 1194

Venous testing
continuous wave doppler, 1182–1183, 1182r
contrast venography, 1184
CT venous imaging, 1184
D-Dimer, 1184
duplex ultrasound, 1184
magnetic resonance imaging, 1184
plethysmography, 1183, 1183f

Vasculitic syndrome, 1191–1192, 1192f
Venopneumatic compression therapy, 1201

Vein
patients selection, 504
goal setting, 505

Ventricular septal defect
1199

Ventricular tachycardia
1199

Vertebral fracture, rehabilitation of, 1000–1005, 1001r

Vertebrae
compression fractures, 1004
vertebral deformities, in adults
back pain
and AS, 873
and hyperkyphosis, 873
clinical presentation, 870–871
definition and pathogenesis, 870
diagnosis, classification, and prognosis,
871–872, 871f
rehabilitative approach, 872
scoliosis progression, 872–873, 872f
surgical treatment, 872

Vertebral fracture, rehabilitation of, 1000–1005, 1001r

Vertebrobasilar syndrome, stroke, 556–557
Vertebralplasty, 1000

Vesicoureteral reflux, 1372

Videofluorography swallowing study
VFSS. See Videofluorography swallowing study

Vibroacoustic accentuation
1405–1408

Visual feedback, 1939

Vitamin D
bone loss, 1000
for osteoporosis, 1000

Vocation
employment outcomes, 470

Vocational rehabilitation (VR)
characteristic, 459–460

Visual feedback, 1939

Videofluorography swallowing study, 1367–1369

Vibration syndrome, 1193

Vibroacoustic accentuation, 1405–1408

Vascular disease
chronic venous insufficiency, 1196–1197
deep vein thrombosis (DVT), 1194–1196
May-Thurner syndrome, 1197
thrombophilia, 1194
venous testing
continuous wave doppler, 1182–1183, 1182r
contrast venography, 1184
CT venous imaging, 1184
D-Dimer, 1184
duplex ultrasound, 1184
magnetic resonance imaging, 1184
plethysmography, 1183, 1183f

Vasculitic syndrome, 1191–1192, 1192f
Venopneumatic compression therapy, 1201

Vegetative state with TBI, 585–587

Venous disease
chronic venous insufficiency, 1196f
rehabilitation, 1196–1197
surgery, 1197
deep vein thrombosis (DVT)
anticoagulants, 1195
medical treatment, 1194
phlegmasia cerulea dolens, 1196
thrombolytic agents, 1195
vena cava filters, 1195–1196
May-Thurner syndrome, 1197
thrombophilia, 1194

Venous testing
continuous wave doppler, 1182–1183, 1182r
contrast venography, 1184
CT venous imaging, 1184
D-Dimer, 1184
duplex ultrasound, 1184
magnetic resonance imaging, 1184
plethysmography, 1183, 1183f

Venous thromboembolism, stroke, 570

Ventilatory support, in myopathies
clinical features of, 773
indications for, 773–774

Modeling and simulation
1001

Visual feedback, 1939

Videofluorography swallowing study
VFSS. See Videofluorography swallowing study

Vibroacoustic accentuation, 1405–1408

Vascular disease
chronic venous insufficiency, 1196–1197
deep vein thrombosis (DVT), 1194–1196
May-Thurner syndrome, 1197
thrombophilia, 1194
venous testing
continuous wave doppler, 1182–1183, 1182r
contrast venography, 1184
CT venous imaging, 1184
D-Dimer, 1184
duplex ultrasound, 1184
magnetic resonance imaging, 1184
plethysmography, 1183, 1183f

Vasculitic syndrome, 1191–1192, 1192f
Venopneumatic compression therapy, 1201

Vegetative state with TBI, 585–587

Venous disease
chronic venous insufficiency, 1196f
rehabilitation, 1196–1197
surgery, 1197
deep vein thrombosis (DVT)
anticoagulants, 1195
medical treatment, 1194
phlegmasia cerulea dolens, 1196
thrombolytic agents, 1195
vena cava filters, 1195–1196
May-Thurner syndrome, 1197
thrombophilia, 1194

Venous testing
continuous wave doppler, 1182–1183, 1182r
contrast venography, 1184
CT venous imaging, 1184
D-Dimer, 1184
duplex ultrasound, 1184
magnetic resonance imaging, 1184
plethysmography, 1183, 1183f

Venous thromboembolism, stroke, 570

Ventilatory support, in myopathies
clinical features of, 773
indications for, 773–774

Modeling and simulation
1001

Visual feedback, 1939

Videofluorography swallowing study
VFSS. See Videofluorography swallowing study

Vibroacoustic accentuation, 1405–1408
INDEX

Voiding dysfunction (Continued)
Parkinson's disease, 1351–1352
peripheral lesions, 1353
sacral lesions, 1353
supraspinal lesions, 1351
suprasacral spinal cord lesions, 1352–1353
Voiding physiology
geriatric, 1349–1350
normal, 1349
pediatric, 1350
Volleyball, recreational sports, 1455–1456, 1456f
Vomiting
center
emesis, 1915f
location of, 1914f
palliative care, 1914–1918
VR, See Vocational rehabilitation

W
Wagner classification, 972
Wagner-O'Day Act, 463
WAIS-III. See Wechsler Adult Intelligence Scale-III
Walking casts, posterior tibialis tendonitis, 958
Wallerian degeneration
of peripheral nerve, 743
Warfarin

Water
buoyancy, 1674–1675
density, 1674
flow characteristics, 1675
hydrostatic pressure, 1674
specific gravity, 1674
specific heat, 1675–1676
thermal energy transfer, 1676
viscosity, 1675, 1675f
Water skiing, recreational sports, 1451–1452, 1452f
Water-based exercise, 1694
Weak calf gait, 131–132
Weakness
disse, 1255
HIV, 1236
Wechsler Adult Intelligence Scale-III (WAIS-III)
14 subtests, 394, 395f
IQ scores, 395, 396f
Weight loss
osteoarthritis (OA), 787–788
Wernicke's aphasia, 421
Western massage
effleurage, 1726
friction, 1727
pétissage, 1726–1727
tapotement, 1727
Wheelchair
antitzippers, 2104
armrests, 2102–2103
caster wheels, 2104, 2104f
clothing guards, 2102–2103
controls, 2108
drive classification, 2107, 2107f
electric, 2109
footrests, 2102
funding/insurance, 2114–2115
future, 2116
gate-aids, 2105
joysticks, 2108
manual
alternative, 2105–2106
attendant-propelled, 2100–2101, 2101f
depot and attendant-propelled, 2100, 2100f
lightweight and ultralight, 2101–2102, 2101f
propulsion technique, 2106, 2106f
selection and setup, 2105, 2105f
pediatrics
mobility, 2113
psychosocial and family issues, 2113
seating systems, 2112–2113
power, 2106–2109
power base, 2106–2107, 2107f
prescription process
documentation, 2099–2100, 2100r
follow-up, 2100
patient history, 2097–2098
physical examination, 2098–2099
team approach, 2097, 2098r
push handles, 2105
push rims, 2104, 2104f
seating and positioning
back support, 2110
custom-contoured, 2110
material properties of, 2109f
power-assisted, 2111, 2111f
pressure-mapping, 2110
recline and tilt-in-space, 2110
scoters, 2111–2112, 2112f, 2112r
setup, 2111
secondary injury/accidents, 2114
sip-and-puff, 2108
standards and durability, 2113–2114
switches and buttons, 2108
tires, 2103–2104
wheel locks, 2103
wheels, 2104
Wheelchair rugby, recreational sports, 1456–1457, 1456f
Wheelchair User's Shoulder Pain Index (WUSPI), 700
Whirlpools and Hubbard tanks, 1693–1694, 1693f
White matter diseases, 198–199, 199f
WIA. See Workforce Investment Act
Wide range achievement test-4 (WRAT-4), 396
WF-III. See Woodcock-Johnson psycho-educational battery-III
Wolf's law, 784
Women health
demographic factors, 1518–1519, 1519f
health care barriers, 1517–1518, 1518f
1518f
Women with disabilities
abuse, 1535–1538, 1535f–1537f
aging, 1534
depression, 1534
health disparities, 1519
menopause, 1533
mothering, 1528–1533
reproductive health care
adolescence
access to information, 1521–1522
maturation management, 1520–1521
personality/self-esteem, 1522–1523
breast self examination, 1520
childbearing age
contraception, 1524–1525
labor and delivery, 1527
menstrual irregularity, 1523–1524
pregnancy and prenatal preparation, 1525–1527
gynecologic examinations, 1520
violence, 1535–1538
Woodcock-Johnson psycho-educational battery-III (WF-III), 396
Work
disability and reasonable accommodation
Americans with Disabilities Act (ADA), 296–297, 297r
ergonomics, cervical axial and radicular pain, 830, 830f
history, 11
team effective
McGregor's characteristics of, 362r
Workers' compensation, 290, 292
Workforce Investment Act (WIA), 465
Workload
oxygen consumption, 1081f
World Health Organization
analytical ladder, 1910f
World Health Organization (WHO)
ICIDH, 301
World War II, 463
Wound care, 1365–1366
with burns, 1128
diabetic, treatment guidelines, 975f
Wound healing, 1358–1359, 1358f
therapeutic electrical stimulation, 1959–1960
Wound measurement, 1363
WRAT-4. See Wide range achievement test-4
Wrist
arthroplasty, 1033
fracture
osteorrhesis, 1006
immobilized
positioning, 2054f
injection technique, 1861, 1861f
joint injection techniques, 1861, 1861f
MRL, 148–149, 149f
carpal tunnel syndrome, 148–150, 150f–151f
triangular fibrocartilage complex
evaluation, 150–151, 150f–151f
rheumatoid arthritis of, 1054
splint
functional, 1044f
sports medicine, 1425–1427
Wrist disarticulation (WD), 2039–2040
Wrist joint, 1861
Wrist units, 2039
Wrist-hand-finger orthosis, 2057–2058, 2057f, 2058f
Writer’s cramp, 948
Writing, 42
Written evaluation, 235–236
Written protocols, 375–380
W-sitting, 1482f
WUSPI, 700

X
Xylocaine, 1757, 1758t, 1781t, 1818r

Y
YMCA
  aquatic exercise regimens, 1683
Arthritis Foundation Aquatic Program (AFAP), 794
Yoga, 2123

Z
Zanaflex, 1765
Zonegran, 1757
Zonisamide (Zonegran), 1757
Zoomtext Xtra, 2005f
Zoster, 1773–1774
Zung Self-Rating Depression Scale, 1907f
Zygapophyseal joint (Z-joint) injection
  anatomy, 1886, 1887f
axial pain and symptom referral, 814–815, 815f
cervical medial branch ablation, 1891
complications, 1891–1892
contraindications, 1888
efficacy, 1888–1890, 1889f
  indications, 1887–1888, 1887f, 1888f
lumbar medial branch neurotomy, 1890–1891
medial branch blocks, 1888
pathophysiology, 1886–1887
prevalence, 1886
radiofrequency neurotomy, 1890
therapeutic effect, 1890, 1890r